The Aggregate Value of Land in the Greater Los Angeles Region[1]

Huiling Zhang[2] and Richard Arnott[3]

August 31, 2014

Abstract: This paper estimates the aggregate value of land in the Greater Los Angeles Region in 2000 using the land parcel database of the Southern California Association of Governments (SCAG), which combines land registry and property tax assessment data from the constituent counties. To our knowledge, the paper is the first to estimate aggregate land value from a parcel database. Aggregate land value is of interest in several contexts: macroeconomic modeling of land and property markets, taxation of land and property, and regional and national accounting. Land parcel databases hold great promise for application in urban and regional policy analysis. Unfortunately, the assessment component of the SCAG database has severe problems with missing and erroneous data. One contribution of the paper is to alert researchers to these problems, which are likely present in other land parcel databases, and to advise them not to trust results reported for any land parcel database unless accompanied by documentation of how these problems were dealt with. We establish a lower bound on the ratio of aggregate land value to regional income of 1.114, which is higher than previous estimates, and argue that the true ratio is likely considerably higher.

[1] This paper was written in conjunction with a research project, "Virtual Co-Laboratory for Policy Analysis in the Greater L.A. Region", funded through the University of California's Multi-campus Research Program and Initiative, grant number 142934. The authors would like to express their gratitude to the Program for its support, and to the Southern California Association of Governments (SCAG) for giving us access to their database. We would like to thank Jifei Ban, Yizhen Gu, Yungtao Guo, and Ramy Mami for their able research assistance on various aspects of the project related to the topic of this paper. Zhang would like to thank participants at 2014 World Symposium on Transport and Land Use Research for useful comments on an earlier draft of the paper. We would both like to thank the referees for constructive and insightful comments on the submitted version of the paper.
[2] Corresponding author. Department of Economics, University of California, Riverside. hzhan017@ucr.edu.
[3] Department of Economics, University of California, Riverside. richard.arnott@ucr.edu

What is the aggregate value of land, absolutely and relative to gross domestic product and the value of national wealth? This question is of interest in at least three broad contexts. First, how important is land as a factor of production? To the classical economists, land wealth was perhaps the most important component of the "wealth of nations", and the determination of land rent and land value was a central element of their economic theories. In contrast, outside location theory, regional science, and urban economics, land now receives virtually no attention in the modern mainstream literature in economics. Does the shift from an economy based primarily on agriculture, to one based primarily on manufacturing, and now one based primarily on services, justify the marginalization of land in current economic theory? Second, how potentially important is land value as a tax base? Henry George advocated a 'single tax' on land value to finance government goods and services, on both equity and efficiency grounds[4]. The mainstream view today seems to be that land taxation is highly efficient because land is inelastically supplied but difficult to implement because land value is difficult to measure. But property sales, including land sales, are typically in the public domain, and an increasingly large proportion is now recorded and publicly available in electronic form. It is reasonable to conjecture that, before long, all land sales will be publicly available in this form, substantially mitigating the valuation problem, which will make greater use of land taxation an attractive policy option. The third is macroeconomic. Shocks from outside the real property market affect the asset value of real property, generating wealth effects, which is a potentially important channel through which shocks can affect aggregate economic performance. Shocks from inside the real property market can also have amplified economic effects on the rest of the economy, as is evident from the Great Recession. To understand these effects better, reliable data on land values are important.

Our paper is the first, to our knowledge, to estimate the aggregate value of land for any geographic area using a land parcel database, which in principle provides reliable data on every parcel of land in a region. We address the aggregate value of urban land in the Greater Los Angeles Region only because, through the project we are working on, we have had access to a complete land parcel database for the Region. The Southern California Association of Governments (SCAG) assembled the database from sales and assessment records provided by its six constituent counties. In principle, for every parcel

---

[4] There is an intriguing Theorem in modern spatial economics called the Henry George Theorem, which Arnott worked on forty years ago in his Ph.D. thesis research. It was discovered independently by Serck-Hanssen (1969), Flatters, Henderson, and Mieszkowski (1974), Starrett (1974), and Vickrey (1977), and elaborated in Arnott (1979) and Arnott and Stiglitz (1979). It states that, if economic activity is efficiently organized over space and if pricing is efficient (at marginal cost), then the confiscatory taxation of urban land rent raises just enough revenue to finance the efficient level of local public services, so that, following Henry George, it is the 'single tax' needed to finance local government. The empirical relevance of the Theorem has been discussed (Arnott 2004) and debated among urban economic theorists. On one hand, Kanemoto *et al.* (1996) applied the Theorem to estimate whether Tokyo is too large. On the other, skeptics argue that pricing and the spatial organization of economic activity are in reality so inefficient that the Theorem is merely a theoretical curiosum.

of land in the Region the database provides accurate information on current land use, and, for parcels that are not yet developed in urban use, on planned land use as well, in addition to current assessed value, and the date and sales price of the parcel on its most recent sale. Thus, the database should provide the basis for accurately estimating the aggregate value of land in the Region. Unfortunately, principle differs from practice. As we shall see, there are numerous problems with the assessment database. Thus, a secondary contribution of our paper is to document the types of problems encountered in working with land parcel databases. Documenting these problems should prove helpful to future researchers when such land parcel databases become available to researchers for other metropolitan areas. The sales price data, while incomplete, are however reliable.

Due to problems with the parcel database, we do not provide a point estimate of the aggregate value of urban land in the Los Angeles Region. As we work our way through our procedure, we shall either make conservative assumptions or investigate alternative approaches. We are therefore confident of our *lower-bound* estimate that the aggregate value of all land in 2000 in the Region was 1.114 times the regional income in that year, and for urban land according to our definition was 0.94 times the regional income. Since this is higher than other estimates in the literature derived using alternative, national accounting methods, we are also confident in asserting that the existing literature underestimates the aggregate value of land, though not by how much. In the next section, we provide some background before starting our numerical work.

Section 1 discusses theoretical issues related to the valuation of land, and also reviews some recent estimates of aggregate land values. Section 2 introduces the 2008 SCAG parcel database. Section 3 estimates the aggregate value of vacant land, Section 4 estimates the aggregate value of developed land, and Section 5 combines the two sets of estimates, obtaining the aggregate value of both developed and vacant land. Section 6 discusses robustness checks, while Section 7 concludes. There are also five appendices, which together provide more detail on the estimation procedures and results.

1. Background

In this section, we review the main theoretical issues involved in the valuation of land. We also review previous work in the literature that estimates the aggregate value of land using national accounting procedures.

1.1 Theoretical Issues in the Valuation of Land

There are seven principal issues in the valuation of land.

1. The concept of "value" comes with a lot of intellectual baggage. It was much debated among the classical economists. There is, for example, the paradox of value. Why is water, which is "priceless", almost free, when diamonds as personal adornment are very expensive, though frivolous and completely inessential? Today, economists distinguish between social value (the shadow price) and private or market value. If the market value of land is taxed, its market value falls, while, with qualifications, its social value remains unchanged. Similarly, zoning a parcel of land restrictively reduces its market value, even though it does not alter the intrinsic properties of the land.

We sidestep altogether these almost metaphysical issues, and instead focus on the *market value of land.*

2. A raw parcel of land becomes more valuable when it is graded (leveled), and even more valuable when it is "serviced" -- connected to a city's water and sewage, telephone, electricity, gas, etc. networks. Should land be valued in its raw, leveled, or serviced state?

We sidestep this distinction too, since the parcel database provides little information on whether a parcel of land is raw, leveled, or serviced. Thus, we measure a parcel's market value at the time it was sold, whatever its state at the time.

3. How should the value of land be measured for a developed site? Shoup (1970) and Arnott (2005) have argued that the appropriate definition is "raw site value" -- "what the land would be worth if it there were no structure on the site, even though in fact there is." With this definition of developed land value, the taxation of land value does not distort the timing and density of redevelopment. Assessors often use an alternative definition, "residual site value", which is measured as property value minus the depreciated value of the structures on the site. To see that this measure distorts the timing and density of redevelopment decisions, consider a property that is about to be redeveloped -- with the current structure demolished and a new structure constructed. The current structure contributes nothing to the value of the property, even though it may, according to the accounting method of depreciation used, have positive depreciated value. Since the pre-development value of land is typically underestimated, residual site value taxation encourages later redevelopment than is efficient.

4. This paper distinguishes between the value of *urban* and *non-urban* land. How should urban land be distinguished from non-urban land? The standard approach is to define land as urban if it is potentially developable in urban use, and this definition is

typically applied based on the topographical, accessibility, and hydrological characteristics of the site (Angel *et al*., 2012; Saiz, 2010). We follow the standard approach in defining land as urban if it is potentially developable in urban use. However, our approach in defining what land is potentially developable in urban use is unconventional. We believe that the market knows best what land is developable, and will value it accordingly. Thus, we define a parcel of vacant land as undevelopable in urban use if its assessed value per unit area is significantly less than the average assessed value per unit area of land in its county. The application of this, or any other method, to the Greater Los Angeles Region is confronted by the problems that two of its constituent counties, Riverside and San Bernardino, are vast, extending almost 300 miles from Los Angeles, all the way to the Arizona border, and that much of the land in the Region is either desert or mountain. Another related issue is the treatment of national and state parks and forests. In England, the Royal Forests used to be much more extensive than they are today. Similarly, over the long term, even national and state parks and forests should not be immune to development. Nevertheless, over the medium term that we consider, we treat the land in them as undevelopable.

We shall present estimates of the both the aggregate value of developable land according to our definition, and the aggregate value of all land.

5. A more technical issue that arises is which should be estimated econometrically, the value of land per unit area, or the value of a vacant parcel, taking into account its area? It turns out that this issue is quantitatively important because, after controlling for the standard measures of accessibility, larger parcels are on average less valuable per unit area. In the literature, this is called the *plattage effect*. In our base regressions, we find that the elasticity of the value of a developable vacant parcel with respect to lot size is only 0.639 and the elasticity of the value of an undevelopable vacant parcel with respect to lot size is only 0.583. On one hand, the economic theory of land rent and land use is developed in terms of land rent per unit area, independent of lot size. On the other hand, the plattage effect is evident in our daily lives. On average, downtown lots are smaller than suburban lots, which are in turn smaller than periurban and rural lots. Furthermore, as cities develop lots tend to be subdivided. We estimate the value of a parcel, taking into account the plattage effect, on the rationale that the plattage effect captures unobserved variation in land quality, in particular aspects of land quality that are not accounted for by the standard accessibility measures.

6. Should the value of land be estimated taking into account neighborhood effects, which include the average quality of structures, the level of public services and taxes, the level of various pollutants, and the income-demographic composition of residents? From an econometric point of view, the answer is yes. The most accurate forecasts are obtained from regression equations in which everything but the kitchen sink is an explanatory variable, however much the explanatory variables are endogenous. Unfortunately, augmenting the parcel database to include all these other variables would have increased the research cost by an order of magnitude. The best we do is to incorporate city dummy variables. A defense of our approach is that neighborhood effects affect the land value surface only locally, and should have little effect on the aggregate value of land -- the

volume under the aggregate land value surface. We know of no theoretical or empirical work that addresses the soundness of this defense.

7. Most economists believe that, in theory, land value provides a better tax base than property value, since, unlike property taxation, land value taxation is neutral with respect to (does not distort) the timing and density of urban development. The standard argument used to be that, while land value taxation is superior to property value taxation in principle, it is inferior in practice since there are few sales of vacant land, and many, if not most, of those are not arm's-length transactions, especially in heavily developed areas. During the era of urban renewal projects, many city governments assembled blighted properties using the right of eminent domain and sold them to a private developer for a nominal price, in return for the developer redeveloping the assembled parcels according to an agreed-upon plan. This may still be the case for certain parts of the downtown area, and perhaps for other blighted areas, but most of the Los Angeles area is healthy. Also, it used to be claimed that many intra-family transfers of land and property are at below-market prices (though the assessments are based on market prices). We identified a field in the 2003 SCAG parcel database, giving the legal form of property sale, and found that only a small percentage of vacant land sales were classified as not being at arm's length. These problems therefore appear to be quantitatively less important than the conventional wisdom suggests. To the extent that these problems remain, they result in the *undervaluation* of vacant urban land.

1.2 Estimates of Aggregate Land Values Using Alternative Approaches

There are two general approaches to estimating aggregate land values. The first is the approach that we have taken, to estimate aggregate land values from vacant land sales. The second draws on information in the national accounts.

The first approach is employed in a pair of companion papers, Nichols, Oliner, and Mulhall (2010, 2013) and is applied to 23 MSA's. The aim of these papers was to develop indexes of commercial and residential land prices from 1995 to 2009. Average land values were not estimated. These papers' method is very similar to that employed in this paper, but drew on a different database, source parcel data vacant land sales obtained from the CoStar Group, Inc. For each MSA and for both land use types, and with single parcel sales as observations, a separate regression was run, with the log of land value per ft^2 regressed against property and transactions characteristics: type of property, condition of the property, intended use of the property, characteristics of the transaction, grid vertex, distance from the population-weighted center of the MSA, and semi-annual time dummies. The indexes were obtained from the time dummies, and are similar to those implied by the time dummies in our regression analysis.

The second approach draws on the national accounts. Davis and Heathcote (2007) and Davis and Palumbo (2008) focused on residential land at the national level. The aggregate value of residential land was determined as a residual, equaling the aggregate market value of residential real estate minus the aggregate replacement cost of residential structures, which were derived from the Flow of Funds accounts published by the Federal

Reserve Board. To estimate the aggregate replacement cost of residential structures, they employed a Census Bureau estimate that residential land accounts for 12.6% of the National Income and Product Accounts gross investment in new residential structures, and then applied a perpetual inventory method, with assumed depreciation rates, to estimate the aggregate replacement cost of residential structures. For the aggregate market value of housing, they employed Census estimates along with the Residential Finance Survey. Having obtained value time series, they decomposed these into price and quantity series. For structures, they employed a price index for gross investment in new residential structures produced by the Bureau of Economic Analysis, and for residential real property, they employed the repeat-sales-based index produced by the Office of Federal Housing Enterprise Oversight. This provided them with the information they needed to estimate a land price index.

The World Bank (2006) estimated per capita wealth, as well as its components, in 2000 for a broad range of countries. This study too drew on the countries' national accounts, but because of data limitations in many countries' national accounting, the estimates were obtained by cruder methods. They treated land as "natural capital", decomposing it into energy resources, mineral resources, timber resources, non-timber forest resources, cropland, pastureland, protected areas, and urban land. Their urban land is comparable to our developed urban land. They estimated the value of urban land as 24% of the value of physical capital, which, applied to the United States, gives an estimate of $15,460 (Appendix 2 of their paper). In our paper, the aggregate land value in 2000 for developed properties in the LA Metro Area was $351.92 billion. With a population in that year of 16.372 million, this translates into a per capita value of $21,495.

In conclusion, our results are broadly consistent with others in the literature, both those that draw on sales transactions in vacant land, and those that draw on the national accounts.

2. *Some Preliminaries*

We proceeded in two steps. In the first, we estimated the aggregate value of land for parcels that were vacant in 2007. In the second, we estimated the aggregate value of land for parcels that were developed in 2007. Section 3 reports on the first step. Section 4 reports on the second step.

The core data for this study are taken from the 2008 Southern California Association of Governments (SCAG) parcel database[5]. The data were assembled by SCAG from land registry and parcel assessment data provided by each of its constituent counties. The SCAG region covers six counties, Imperial, Los Angeles, Orange, Riverside, San Bernardino, and Ventura. The Greater Los Angeles Area, in contrast, is defined to include Los Angeles, Orange, Riverside, San Bernardino, and Ventura Counties. Imperial County is a relatively small, predominantly agricultural county, bordering on Mexico, centered about 200 miles southeast of Los Angeles. We chose to exclude Imperial County from this study, so that when we refer to LA Metro Area we mean the five counties of the Greater Los Angeles Area. Figure 1 provides a map of the LA Metro Area. The reader should keep in mind that much of the LA Metro Area, especially in Riverside and San Bernardino Counties, is desert and mountains, and that the eastern edge of Riverside County is 225 miles away from downtown Los Angeles. Thus, one of the issues that will need to be addressed is the appropriate definition of urban land within the Metro Area.

Of the many fields in the 2008 SCAG parcel database, the main part of this study uses only year of most recent sale and sales price at that time, which were obtained from the county registry offices, the 2007 assessed value and current land use code, which were obtained from county assessment data, and for vacant parcels the plan land use code (which gives the planned land use from the relevant official Plan, as collected by SCAG).

Under law, the county registry offices are required to have comprehensive and reliable registry data for all parcels of land within their jurisdictions, and the county assessment offices are required to have comprehensive and reliable assessment data for all parcels within their jurisdictions, and in principle the SCAG land parcel database provides a consolidation of these databases, using SCAG's standardized land use classification system. Unfortunately, not only are their numerous problems with the assessment data in SCAG parcel database, but also, since the consolidation was undertaken by employees for a private consulting company who did not fully document their procedures and no longer work for the company, there is no expert to consult about these problems. When we started this research, we naively assumed that the data were complete and sound. If we had known then what we know now, we would have proceeded rather differently.

---

[5] Data in the SCAG parcel database were assembled in 2008, and hence the name "2008 SCAG parcel database". But all data, including assessment data and land registry data, are the most recent as of 2007, and hence the name "2007 assessed value".

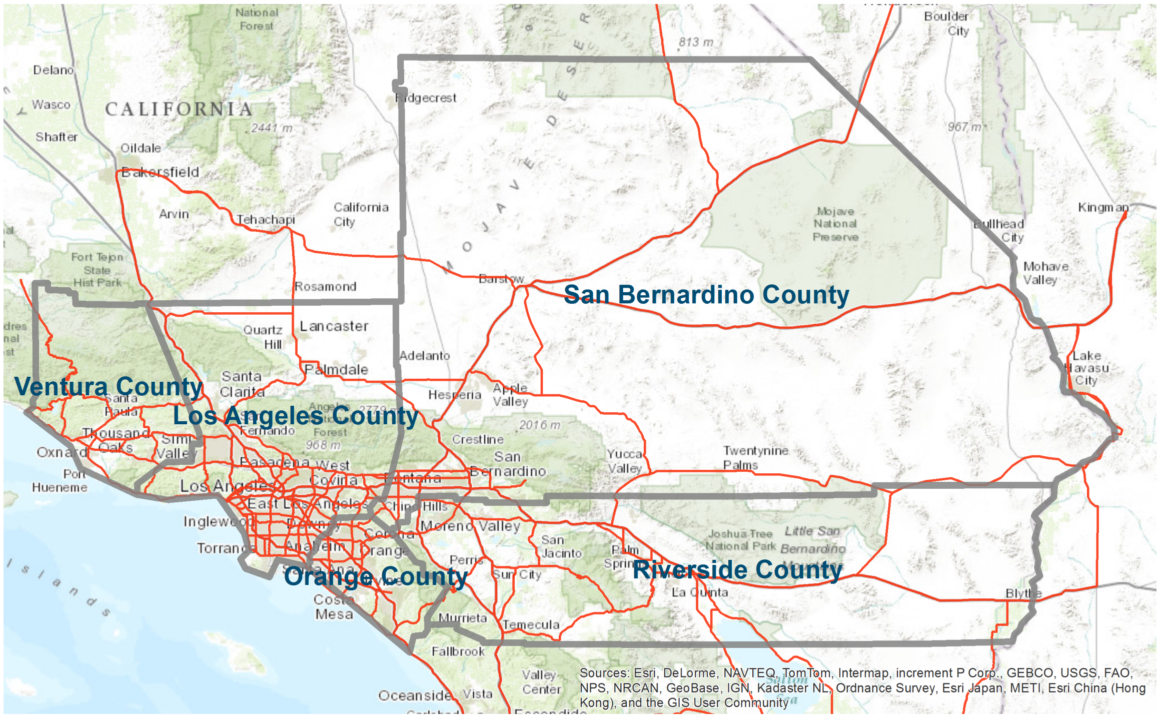

Figure 1: Map of the LA Metro Area. Grey lines denote county boundaries, and red lines denote major road links

A preliminary step in the procedure is to provide an operational definition of vacant land. The concept of "vacant" land employed here is a broad one, and includes all land either without a completed structure or for which the structures are of secondary importance to the land use (notably open space and recreation, and agriculture). Table 1 gives the SCAG 2008 current land use categories that we took to correspond to vacant land. There are five main categories: under construction, open space and recreation, agriculture, vacant, and military (vacant). Later we shall decompose vacant land into developable and undevelopable vacant land. The right-most column of the table is relevant for that procedure.

Main categories	Detailed type	Developability
17** Under Construction	1700 Under Construction	Y
18** Open Space and Recreation	1800 Missing detailed type	N
	1810 Golf Courses	Y
	1820 Local Parks and Recreation	N
	1821 Developed Local Parks and Recreation	N
	1822 Undeveloped Local Parks and Recreation	N
	1830 Regional Parks and Recreation	N
	1831 Developed Regional Parks and Recreation	N
	1832 Undeveloped Regional Parks and Recreation	N
	1840 Cemeteries	Y
	1850 Wildlife Preserves and Sanctuaries	N
	1860 Specimen Gardens and Arboreta	N
	1870 Beach Park	N
	1880 Other Open Space and Recreation	N
2*** Agriculture	2000 Missing detailed type	Y
	2100 Cropland and Improved Pasture Land	Y
	2110 Irrigated Cropland and Improved Pasture Land	Y
	2120 Non-Irrigated Cropland and Improved Pasture Land	Y
	2200 Orchards and Vineyards	Y
	2300 Nursery	Y
	2400 Dairy, Intensive Livestock, and Associated Facilities	Y
	2500 Poultry Operations	Y
	2600 Other Agriculture	Y
	2700 Horse Ranches	Y
3*** Vacant	3000 Missing detailed type	Y/N
	3100 Vacant Undifferentiated	Y/N
	3200 Abandoned Orchards and Vineyards	Y
	3300 Vacant with Limited Improvements	Y
	3400 Beaches (Vacant)	Y
127* Military (Vacant)	1274 Former Base (Built-up Area)	Y
	1275 Former Base (Vacant Area)	Y/N
	1276 Former Base Air Field	Y

Table 1: SCAG 2008 parcel database land use categories classified as "vacant".

Now is a convenient point at which to introduce some aggregate data on the area of land in the five counties that is vacant, developed, and unclassified, and corresponding data on parcels. Table 2 and Table 3 report on the relevant aggregate data.

	Vacant land area (10^9 ft^2)	Developed land area (10^9 ft^2)	Unclassified land area (10^9 ft^2)	Prop of land vacant	Prop of land developed	Prop of land unclassified
Los Angeles	72.14	38.63	8.4	60.54%	32.42%	7.05%
Orange	9.32	13.35	2.6	36.88%	52.83%	10.29%
Riverside	175.69	28.23	4.5	84.30%	13.54%	2.16%
San Bernardino	455.68	186.68	4.1	70.49%	28.88%	0.63%
Ventura	44.19	8.47	0.1	83.76%	16.05%	0.19%
Total	757.02	275.36	19.7	71.95%	26.17%	1.87%

Table 2: Vacant, developed, and unclassified[6] land area, by county, total and proportions.
Notes: 1. Unlike the other counties, Ventura County assigns parcel numbers to streets, which explains why such a small proportion of its land area is unclassified.
2. 1 ml^2 = 0.0278784 X 10^9 ft^2 so that 10^9 ft^2 = 35.87 ml^2.

	Number of parcels	Average land area of a parcel (Acre)	Number of vacant parcels	Average land area of a vacant parcel (Acre)	Number of developed parcels	Average land area of a developed parcel (Acre)
Los Angeles	2047715	1.24	153974	10.76	1893741	0.47
Orange	703656	0.74	23594	9.07	680062	0.45
Riverside	793150	5.90	160346	25.15	632804	1.02
San Bernardino	852460	17.30	206865	50.57	645595	6.64
Ventura	266701	4.53	20021	50.67	246680	0.79
Total	4663682	5.08	564800	30.77	4098882	1.54

Table 3: Number and average land area of parcels, total, vacant, and developed.
Note: 1 acre = 4046.86 sq. meters = 43560 sq. ft.

[6] Unclassified land includes land that is not assigned to a parcel and parcels whose land use is entered as blank or zero. Except in Ventura County, little of the land used for city streets and freeways is part of a parcel in the SCAG database. There are also parcels whose current land use is water, which are not counted as unclassified; they are simply excluded from our calculations.

As casual experience would suggest, the proportion of land that is developed is considerably higher in the two central counties, Los Angeles and Orange, than in the three peripheral counties. The anomaly that San Bernardino County, which is primarily desert, has a larger proportion of developed area than either Riverside or Ventura Counties is accounted for by military bases.

As experience suggests, average parcel size is smaller in the central than in the peripheral counties, and the average parcel size of developed parcels is smaller than that of vacant parcels.

A central decision that needed to be made in the research design was whether to use sales price at the date of last sale, adjusted to the year 2000 using a method we shall explain, or to use 2007 assessed value, adjusted to the year 2000 using a different method we shall explain. 2007 data on assessed values were available for all the counties. Data on the sales price at the date of most recent sale were available for only three of the counties, Los Angeles, Orange, and Riverside, and we did not trust the Los Angeles County data because there were so many common values. Furthermore, for both Orange and Riverside Counties, the data on 2007 assessed values were more complete (fewer blanks and zeroes) than the data on sales prices. The State of California employs the principle of market-value assessment, and when a property is sold the assessed value of a property is, in principle at least, set equal to the sales price. We therefore decided to proceed using assessed values.

One complication with using assessed values is that, due to Proposition 13,[7] a parcel's assessed value in a particular year is not an estimate of its current market value, even though the underlying assessment principle is market value, but is instead determined by a formula that sets a parcel's assessed value at the time of the property's most recent sale equal to the sales prices at that time, and then updates its assessed value annually based on a formula (State Board of Equalization, State of California, 2012)[8]. We shall explain how we deal with this complication. After we undertook our calculations using assessed

---

[7] Enacted in 1978, Proposition 13 is an amendment to the Constitution of California that limits the tax rate applied to real estate. Of particular relevance here is one of its provisions that restricts annual increases of assessed value of real property to an inflation factor, not to exceed 2% per year. It also prohibited reassessment of a new base year value except in the cases of a change in ownership or the completion of new construction. (http://en.wikipedia.org/wiki/California_Proposition_13_(1978)).

[8] Letting t = 0 denote the year of most recent sale

$$\text{assessed value}_0 = \text{most recent sales price}$$

$$\text{assesssed value}_t = \min\{\text{current market value}, \text{assessed value}_{t-1} * (1 + \min[\text{growth rate in CPI}_{t,t+1}, 0.02])\}$$

In the decade prior to 2007, the annual inflation rate was consistently greater than 0.02 and current market values were never less than assessed values. Over that period, therefore, the formula reduces to

$$\text{assesssed value}_t = \text{most recent sales price}(1.02)^t.$$

values, we checked the average sales price adjusted to 2000 with the average assessed value adjusted to 2000 for a subset of parcels in Riverside County for which we had data for both cells. We found that certain categories of properties appear to be systematically underassessed[9]. We shall return to this issue at some length in Section 6.

Since we are primarily interested in the value of *urban* land, we need to provide an operational definition of "urban". We define a parcel to be urban if either it is already developed or it is "developable". Previous work, such as Angel et al. (2012) and Saiz (2010), categorizes parcels as developable or undevelopable primarily on the basis of topographical characteristics. Our general approach instead is to categorize parcels according to official land use and land value per ft^2, on the rationale that the market knows best what land is developable and what is not. Thus, for example, much of the vacant land in the Santa Monica Mountains, though expensive to develop for topographical reasons, is nevertheless classified as developable since it has a high value per ft^2. Missing data, errors in coding, and anomalies were dealt with via a combination of econometric imputation, spatial smoothing, and ground truth tracking[10] using satellite imaging. We provide a summary of the procedure in Appendix C; details are presented in Arnott and Guo (2012).

Here is an appropriate place to note that, throughout the paper, when we have a choice of assumptions, we employ an assumption that errs on the conservative side. We do this with the aim of being able to assert with confidence that the central estimates of aggregate land values that we present provide lower bounds on the true aggregate values.

[9] For example, most states systematically underassess the value of agricultural land. The stated aim of this policy is to preserve the family farm. The way in which this is done is to assess the value of agricultural land, using the income method, on the basis of the income generated by the parcel. Close to cities, however, most of the value of agricultural land is its option value in urban use. At the time a parcel of agricultural land is converted to urban use, some states impose a cumulative retrospective property tax, recouping the implicit subsidy provided, but others do not. California's Board of Equalization makes no mention of taxing agricultural land on the basis of its "agricultural value", instead simply stating that they apply their principle of market valuation on the basis of one of three assessment methods: hedonic assessment, comparable values, and the income method. It appears that the State of California implicitly, though not explicitly, applies a lower effective property tax rate to agricultural land.

[10] Appendix E describes some of the ground truth tracking that was employed to check the accuracy of the SCAG parcel database.

3. Estimating the Aggregate Value of Vacant Land

The general procedure we employ in this section is as follows:

- For all developable vacant land parcels that contain both the year of most recent sale (and if it is no earlier than 1980[11]) and current assessed value, regress the natural log of 2007 assessed parcel value against a set of explanatory variables.
- For the same set of parcels, use the time dummies in that equation to estimate each parcel's market value, and later make an adjustment to reflect the assessment formula so as to estimate each parcel's 2000 market value
- For the complementary set of developable vacant parcels, use the estimated regression equation to impute each parcel's market value, and later make an adjustment to reflect the assessment formula so as to estimate each parcel's 2000 market value.
- Repeat the procedure for undevelopable vacant land parcels.

Unfortunately, many cells for the relevant fields in the SCAG parcel database contain either a zero or a blank. Except for the fact that Ventura County had not got around to converting all their historical data to electronic form, we do not know the causes of the zeroes and blanks, nor whether the zeroes and blanks are random, conditional of observables. Table 4 provides partial information on the zeroes and blanks, by county. Column 1 gives the total number of vacant parcels in each county. Column 2 gives the percentage of vacant parcels for which assessed value or the year of most recent sale is zero or blank or for which the year of most recent sale is before 1980. Column 3 gives the number of vacant parcels for which data on both assessed value and year of most recent sale are available and for which the sale year is after 1980.

[11] We excluded those observations with most recent sale date prior to 1980 since we are unfamiliar with assessment practices prior to Proposition 13, except that each county employed its own assessment procedures.

In what follows, when we say that a vacant parcel has "complete" data on assessed value, we shall mean that it has data on assessed value *and* on year of most recent sale *and* that the year of most recent sales was later than 1980. Similarly later in section 6, where we compare results using sales price with those using assessed value, when we say that a vacant parcel has "complete data" on sale price, we shall mean that it has data on sales price *and* on the year of most recent sale *and* that the year of most recent sale was later than 1980.

	Number of vacant parcels	Percentage of vacant parcels with zero or blank for assessed value, and zero or blank for year of most recent sale or sale year before 1980	Number of vacant parcels with assessed value, and sale year after 1980
Los Angeles	153974	36.0	98607
Orange	23594	89.8	2400
Riverside	160346	58.2	66966
San Bernardino	206865	89.1	22587
Ventura	20021	85.1	2977
Total	564800	65.7	193537

Table 4: Zeroes and blanks in relevant fields of SCAG 2008 parcel database

Because of the large sample size, the high incidence of zeroes and blanks in the data would not be a major concern if, after controlling for observables, they were random, but their pattern is more likely indicative of differences in assessment practice across offices and even officers. The only way to satisfactorily deal with this problem is to collect the missing data for a sample of parcels with zeroes and blanks, and then to adjust the land value estimates taking this new information into account. While we did some limited ground-truth tracking of vacant/developed classification using a combination of satellite imaging[12], Google Streetview, and site visits, we did not have the resources to proceed far in this direction. Thus, throughout the paper, our estimates are based on the untested assumption that, after controlling for observables, zeroes and blanks in the data occur randomly.

To impute 2000 land value to vacant parcels without complete data, we employed data for the set of vacant parcels with complete data (column 3 of Table 4) to econometrically estimate a hedonic land valuation equation. As the dependent variable, we could have used either natural logarithm of assessed value per unit area of vacant land or the natural logarithm of the assessed value for the entire parcel. We considered both, but since the fit was considerably better, present only the results when the dependent variable is the natural logarithm of assessed value for the entire parcel[13]. The independent variables

[12] Another potential source of error are mistakes in the data entered. Registry office data are normally very reliable, but assessment data are of variable reliability. In ground truth tracking, we did detect errors in the classification of parcels according to current land use. The ground truth tracking we undertook to uncover these and other potential sources of error is reported briefly in Appendix E. Yet another possible source of error is the consulting company that integrated the County databases.

[13] This is due to the plattage effect discussed in section 2.1.

include city dummies, sale year dummies, and six accessibility measures[14] (distance to the CBD, distance to the CBD squared[15], distance to the coast, distance to the coast squared, distance to the nearest employment subcenter, and distance to the nearest major highway or freeway), current land use, and plan land use in 2007. The distance measures to the coast capture crudely differences in natural amenities, climate, and pollution. The city dummies capture crudely amenities, both natural and man-made, zoning policies, and residential sorting by income/race/demographic variables. The sales year dummies capture temporal changes in vacant land value at the level of the metropolitan area, as well as the state formula used to compute assessed values on the basis of the year of last sale and sale price at that time. The accessibility measures are standard. The plan land use dummies capture the dependence of the value of a vacant parcel on its allowable developed land uses. No independent variables are included that specifically measure the residential sorting of income/race/demographic groups across the metropolitan area.[16] The regression equation is estimated using OLS, which ignores spatial autocorrelation. We discuss issues associated with spatial autocorrelation in Section 6, which reports on robustness checks. The estimated regression equation is

$$ln(assessed\ land\ value) = constant + b_0\ ln(land\ area) + b_1X_1 + b_2X_2 + u \qquad (1)$$

where X_1 includes six accessibility measures in miles. X_2 includes four dummy variables: city, (vacant) land use as of the assessment date, planned (developed) land use as of the assessment date, year of most recent sale, and u, the error term, which is assumed to be normally and independently identically distributed.

As noted earlier, we chose to generate separate estimates of the aggregate value of urban and non-urban land, and defined urban land to be land that is "developable in urban use". The general approach employed in categorizing a parcel of land as developable or undevelopable was described in section 2.1, and some details are provided in Appendix C. Arnott and Guo (2012), a project technical report, provides even more detail.

---

[14] Distance to the CBD and distance to the nearest employment subcenter are calculated as the straight-line distance between the centroid of the parcel and the centroid of the CBD or the centroid of the nearest employment subcenter (point-to-point distance measures). Distance to the coast and distance to the nearest freeway are calculated as the shortest straight-line distance between the centroid of a parcel and the coast or the major road (point-to-curve distance measures).

[15] Squared distance to CBD and squared distance to coast are included to capture the non-linear effects of distance to the ocean or CBD on land value. For all vacant parcels in the regression sample, the average distance to CBD is around 95 miles and the average distance to coast is around 50 miles, while distance to nearest freeway averages around 2 miles and distance to nearest subcenter averages around 9 miles. We therefore expect linear approximation to be accurate for the effects of distances to freeway and nearest subcenter but not for the effects of distances to ocean and CBD.

[16] Including these variables would have improved the goodness of fit, but our intuition is that their exclusion smoothes the land value surface without having an important impact on the variable of interest --- the volume under the land value surface.

Table 5 below presents a subset of the regression coefficients. The complete set of regression coefficients is given in Appendix A. We refer to the regression equation for developable vacant parcels as 1A, and that for undevelopable vacant parcels as 1B[17]. The full regression results are reported in Appendix A. With the assumed functional form of the regression equation, the coefficient on ln(land area) is an elasticity, the coefficients on the distance variables that enter linearly are gradients[18] (indicating the proportional change in land value per unit distance), and the coefficients on the dummy variables relate to the proportional change in parcel value when the dummy variable switches from zero to one. Thus, for example, *ceteris paribus,* the value of developable vacant land is estimated to fall off by 3.7% every mile further distant from the nearest subcenter, and a larger coefficient on a city dummy implies that land parcel value is proportionally higher for all sets of values of the other independent variables[19]. The signs and magnitudes of all the coefficients in Table 5 are reasonable.

Table 5: Vacant parcel valuation hedonic regression results

	Dependent variable: Log of assessed value ($)	
	Developable Vacant parcels	Undevelopable Vacant parcels
Log of land area (sq. ft)	0.639***	0.583***
Distance to the nearest subcenter	-0.037***	0.007***
Distance to CBD	-0.016***	0.006***
Squared distance to CBD	4.16E-5***	-1.01E-4***
Distance to the nearest major road	0.017***	-0.010***
Distance to the nearest coast	-0.019***	-0.051***
Squared distance to the nearest coast	9.02E-5***	1.90E-4***
Constant	7.243***	7.232***
Number of observations	99169	94368
Adjusted R-squared	0.264	0.405

* p<0.05, ** p<0.01, *** p<0.001

[17] Note that the sum of the sample sizes in the two regressions equals the total of column 3 in Table 4.

[18] Holding all other regressors fixed, the relationship between assessed value and a distance variable is $A = k \exp\{c_1x + c_2x^2\}$, where k is a constant, x is the distance variable, c_1 is the estimated coefficient on the distance variable, and c_2 is the estimated coefficient on the distance variable squared. Then $dA/dx = [dA/dy(x)][dy(x)/dx]$, where $y(x) = c_1x + c_2x^2$, so that $dA/dx = A(c_1 + 2c_2x)$.

[19] More specifically, when a particular city dummy has a coefficient of c a vacant parcel is e^c more valuable than an observationally identical parcel in the base city, which is Adelanto.

Table 6 records the aggregate value of vacant parcels, by county, for those parcels with complete data on assessed value. The assessed values were inflated or deflated, as the case may be, from the sales year to the year 2000 by exponentiating the coefficient of the corresponding time dummies (since 2000 was the base year), which takes into account both the temporal change in the average market value of land and the procedure employed to update the assessed value of a parcel based on its year of last sale and its sales price at that time.

It is important to note that the land value so measured is the estimated 2007 assessed value, conditional on sale in the year 2000. We shall refer to this simply as "market value" We shall continue with this measure of land value until towards the end of the section 5, when we make the appropriate adjustment to aggregate market value so as to obtain the estimated 2000 assessed value, conditional on sale in year 2000, which we shall refer to as the "aggregate 2000 market value".

	Number of vacant parcels	Aggregate area of vacant parcels (10^9 ft^2)	Market value of vacant land per ft^2	Aggregate market value of vacant land ($\$10^9$)
Los Angeles	98607	25.16	0.72	18.12
Orange	2400	1.34	0.92	1.23
Riverside	66966	21.34	0.55	11.73
San Bernardino	22587	5.49	0.63	3.43
Ventura	2977	4.46	0.45	2.02
Total	193537	57.79	0.63	36.54

Table 6: Estimated aggregate land values for all parcels that were vacant in 2007 and for which complete data on assessed value were available, by county.
Note: Market value per ft^2 was obtained by dividing aggregate market value by the aggregate area of parcels.

Parcel values for parcels that did not contain data on year of most recent sale or whose most recent sale was listed as earlier than 1980 or whose 2007 assessed value was absent, were imputed using fitted values for the above regressions (using (1A) for "developable" vacant parcels and (1B) for "undevelopable" vacant parcels). Table 7 is the same as Table 6 but applies to parcels that did not contain complete data on assessed value. For this set of parcels, the market value of vacant land per ft^2 in Orange County compared to the other counties is an extreme anomaly. We should have, but did not, use ground-truth tracking to determine the source of this anomaly.

	Number of vacant parcels	Aggregate area of vacant parcels (10^9 ft^2)	Market value of vacant land per ft^2	Aggregate market value of vacant land (\$$10^9$)
Los Angeles	55367	46.98	0.31	14.68
Orange	21194	7.98	6.27	50.08
Riverside	93380	154.35	0.05	8.38
San Bernardino	184278	450.19	0.02	6.90
Ventura	17044	39.72	0.22	8.84
Total	371263	699.23	0.13	88.87

Table 7: Estimated aggregate land values for all parcels that were vacant in 2007 for which complete data on assessed value were not available, using fitted values from (1A) and (1B).
Note: Market value per ft^2 for a particular county was obtained by dividing that county's aggregate market value by its aggregate area of parcels.

Note that the ratio of the (imputed) market value of vacant land per ft^2 in Table 7 to the corresponding number in Table 6 differs markedly across counties. We suspect that this is due to differences in assessment practice across counties. Even though the counties have been supposed to conform to State guidelines in assessment practices since the State took over administration of the property tax in the early 1980's, substantial differences apparently remain.

Table 8 gives the estimated aggregate land value for all vacant parcels, adding together the estimated aggregate land value for all vacant parcels with complete assessment data from Table 6 and the estimated aggregate land value for all vacant parcels without complete data from Table 7.

	Number of vacant parcels	Aggregate area of vacant parcels (10^9 ft^2)	Market value of vacant land per ft^2	Aggregate market value of vacant land (\$$10^9$)
Los Angeles	153974	72.14	0.45	32.80
Orange	23594	9.32	5.51	51.31
Riverside	160346	175.69	0.11	20.11
San Bernardino	206865	455.68	0.02	10.33
Ventura	20021	44.19	0.25	10.86
Total	564800	757.02	0.17	125.42

Table 8: Estimated aggregate value for all parcels that were vacant in 2007.
Note: Market value per ft^2 for a particular county was obtained by dividing that county's aggregate market value by its aggregate area of parcels.

Do these numbers look reasonable? At least the rankings do to us. Consider the two extremes, Orange and San Bernardino Counties. Orange County is the smallest, most developed, and least mountainous of the counties, and contains no desert. The bulk of San Bernardino County is Mojave Desert. It is not surprising to us, therefore, that San Bernardino County has almost fifty times as much vacant land as Orange County. Nor does it surprise us that the aggregate value of vacant land is higher in Orange than in San Bernardino County. But the anomalously high value of vacant land per ft^2 in Orange County suggests a problem with the data for Orange County that warrants further investigation.

The value of *urban land* within a metropolitan area hinges on the distinction between developable and undevelopable land. In the context of the LA Metro Area, intuitively one does not want to include the hundreds of square miles of the Mojave Desert in San Bernardino County or national and state parks as urban land. But there is also a broad band of grey between developable and undevelopable land. For example, land in the Santa Monica Mountains, which includes the Malibu Hills, might be classified as undevelopable on the basis of its terrain, but might nonetheless have considerable value as sites for the homes of the rich. Also, land that is currently undevelopable because of the unavailability of water might become developable if, due to technological improvements, the cost of drilling deep wells decreases significantly. Table 1 shows which land use categories are classified as unambiguously undevelopable (N) and which as unambiguously developable (Y). The developability of parcels in the remaining land use categories (3000, 3100, and 1275) was decided on a parcel-by-parcel basis using a method based on ground truth tracking, described in Arnott and Guo (2012), and summarized in Appendix C.

Table 9 corresponds to Table 8 exactly, but applies to parcels that were classified as "developable" according to the above procedure. Table 9 gives the estimated aggregate value for all "developable" parcels, using (1A) to impute missing values.

	Number of vacant parcels	Aggregate area of vacant parcels (10^9 ft^2)	Market value of vacant land per ft^2	Aggregate market value of vacant land ($\$10^9$)
Los Angeles	59726	11.06	1.33	14.72
Orange	13260	1.96	2.02	3.97
Riverside	88308	19.05	0.71	13.51
San Bernardino	53054	18.40	0.30	5.52
Ventura	10628	7.06	0.74	5.26
Total	224976	57.53	0.75	42.98

Table 9: Estimated aggregate value for all parcels that were vacant and "developable" in 2007, using fitted values from (1A) to impute land values for parcels with missing data. Note: Market value per ft^2 for a particular county was obtained by dividing that county's aggregate market value by its aggregate area of parcels.

Table 10 corresponds to Table 9 exactly, but applies to parcels that were classified as "undevelopable" according to the Arnott-Guo procedure. Table 10 gives the estimated aggregate value for all "undevelopable" parcels, using (1B) to impute missing values.

	Number of vacant parcels	Aggregate area of vacant parcels (10^9 ft^2)	Market value of vacant land per ft^2	Aggregate market value of vacant parcels (\$$10^9$)
Los Angeles	94248	61.08	0.30	18.08
Orange	10334	7.36	6.43	47.34
Riverside	72038	156.64	0.04	6.60
San Bernardino	153811	437.29	0.01	4.81
Ventura	9393	37.13	0.15	5.60
Total	339824	699.49	0.12	82.44

Table 10: Estimated aggregate value for all parcels that were vacant and "undevelopable" in 2008, using fitted values from (1B).
Note: Market value per ft^2 for a particular county was obtained by dividing that county's aggregate market value by its aggregate area of parcels.

Does our decomposition of land into developable and undevelopable give reasonable results? Broadly we think it does since, across the Metro Area, the market value of undevelopable vacant land per ft^2 is only about 16% that of developable vacant land[20] (recall we classified land as developable or undevelopable partly on the basis of its market value per ft^2), and considerably lower when Orange County is excluded.

[20] Recall that in Table 1 we classified vacant parcels on the basis of their current land use code as Y (developable), N (undevelopable), and Y/N (developability to be decided on a parcel-by-parcel basis). For parcels with Y/N land use code, we applied the Arnott-Guo procedure to estimate a parcel's developability, which used the parcel's assessed value/ft^2 relative to the assessed value/ft^2 of parcels in the same county.

4. *Estimating the Aggregate Value of Developed Land*

The SCAG parcel database contains a current land use code for all parcels. For the project, we aggregated the SCAG land use codes into 14 land use codes. Table 1 listed the vacant land use codes. Table 11 lists all 11 developed land use codes, and indicates how they correspond to the SCAG land use codes. Our developed land use codes are single-family residential, multi-family residential, mixed residential, office, retail, other commercial, public (except for former military bases, which we classified as vacant), warehousing, other industrial, transportation/communication/utilities, and mixed. There are two other land use codes, "water" and "other", which were not included in either vacant or developed land uses. Land use code "other" corresponds to all parcels with SCAG land use code missing. They constitute the "unclassified" parcels, which along with land not assigned to a parcel, constitute the unclassified land area in Table 2.

All parcels classified as developed are classified as developable. Furthermore, the assessed value of all parcels classified as developed is the assessed *property* value, where a property includes both land and structures. Table 12 gives the estimated land values of developed other than single-family residential (NSFR) parcels. Table 13 gives the estimated land values of single-family residential parcels. Table 14 gives the estimated land values for all developed parcels together, and is obtained from summing the corresponding values in Tables 12 and 13.

LA project land use code			SCAG 1993 code		
	Alphabetic	Name	3-digit	Name	Note
1	RS-SF	Single-family residential	111	Single family residential	
2	RS-MF	Multi-family residential	112	Multi-family residential	
3	RS-MX	Mixed residential	110	Residential	Missing more detailed land use
			113	Mobile homes and trailer parks	
			114	Mixed residential	
			115	Rural residential	
4	OF	Office	121	General office use	
5	RF	Retail	122	Retail stores and commercial stores	
6	OC	Other commercial	120	Commercial and Services	Missing more detailed land use
			123	Other commercial	
7	P	Public	124	Public facilities	
			125	Special Use Facilities	
			126	Educational institutions	
			127	Military installations	
8	W	Warehousing	134	Wholesaling and warehousing	
9	OI	Other industrial	130	Industrial	Missing more detailed land use
			131	Light industrial	
			132	Heavy industrial	
			133	Extraction	
10	TCU	Transportation /communication /utilities	14x	Transportation, communications, and utilities	Including 140, 141, 142, 143, 144, 145, 146
11	M	Mixed	150	Mixed commercial and industrial	
			160	Mixed urban	

Table 11: The aggregation of SCAG parcel land use codes into the codes used in the project, for developed land uses.

Table 12 reports on the estimated aggregate land values for developed properties other than single-family residential. These land values were imputed using an equation similar to (1A)[21], which we label (2A). These estimates are based on the assumption that (2A) provides unbiased estimates of the value of land of NSFR developed parcels, even though they were estimated from data on the assessed value of vacant land parcels. One reason to doubt this assumption is spatial autocorrelation, which we shall address in section 6. Another reason is omitted variable bias. Intuition suggests that, holding the value of the regressors in (2A) fixed, the average parcel of developed land is more attractive than the average parcel of vacant land[22]. For one thing, most developed parcels are serviced while most vacant parcels are unserviced. If this intuition is correct, then the average land value for a developed parcel should be higher than that of a vacant parcel with the same values of the regressors, in which case the values in Table 12 are downward biased.

	Number of developed parcels	Aggregate land area of developed parcels (10^9 ft^2)	Market value of land per ft^2	Aggregate market value of land (\$$10^9$)
Los Angeles	625342	27.45	2.49	68.43
Orange	172772	9.08	3.84	34.86
Riverside	154654	22.72	0.60	13.72
San Bernardino	178123	180.52	0.09	16.80
Ventura	72105	6.56	1.87	12.23
Total	1202996	246.33	0.59	146.05

Table 12: Imputed aggregate land values for developed other than single-family residential properties.

[21] Equation (2A) is the same as (1A) but without the current land use dummies, and like (1A) is estimated using OLS. The regression results are available on request from the corresponding author. In the imputation, the plan land use was replaced by the property's actual current land use. Suppose, for example, that a developed property's current land use is "office". Then the land value imputed to it is the fitted land value from (2A) for vacant land with plan land use "office".

[22] In an earlier version of the paper, Zhang and Arnott (2013), we included an appendix, Appendix F, which applied the Arnott-Lewis model of the timing of the transition of land to urban use (Arnott and Lewis, 1979) to investigate whether theory suggests that more attractive parcels will tend to be developed earlier. Under reasonable assumptions, we found that the theory does support the intuition that "more attractive" parcels are developed earlier. The intuition is that conversion of a vacant (non-urban use) parcel to a developed (urban use) parcel occurs when the marginal benefit of postponing conversion equals the corresponding marginal cost. If one parcel is more attractive than another in having better amenities, and hence would command higher rent when developed, the marginal cost of postponing its development is higher. If one parcel is more attractive than another in having lower development cost, the marginal benefit of postponing development (the amortized development cost) is lower.

At least the relative values across counties of the imputed land values per ft^2 appear reasonable. Orange County is the wealthiest of the counties, followed by Los Angeles, Ventura, Riverside and San Bernardino. As noted earlier the exceptionally low value for San Bernardino County is due to the inclusion of military bases.

The procedure for imputing the value of land for single-family residential parcels is the same as that employed for the other land uses, except of course that the dummy variables on plan land use is SFR land uses.

We have mentioned already that we have evidence, which will be presented in section 6, that vacant land is systematically underassessed. Because of our imputation procedure, this underassessment would result in underestimation of the aggregate value of land in developed uses.

	Number of developed parcels	Aggregate land area of developed parcels (10^9 ft^2)	Market value of land per ft^2	Aggregate market value of land (\$$10^9$)
Los Angeles	1268399	11.18	9.86	110.21
Orange	507290	4.27	15.63	66.76
Riverside	478150	5.52	4.99	27.53
San Bernardino	467472	6.15	5.14	31.63
Ventura	174575	1.91	11.53	22.06
Total	2895886	29.03	8.89	258.19

Table 13: Imputed aggregate land values for single-family residential (SFR) properties.

The relative magnitudes, across counties, of the market value of SFR land per ft^2 look reasonable. The absolute values look reasonable to us too, even though they are probably downward biased too through their imputation from vacant parcels.

It is reassuring that the ratio of aggregate land values for NSFR to SFR properties is similar across the counties, ranging from a high of 0.62 for Los Angeles County to a low of 0.50 for Riverside County. Furthermore, it is highest for the least residential of the counties and lowest for the most residential. We were surprised that the proportion of developed land that is SFR is not higher. The reason appears to be that much of the NSFR developed land is in "public" and "transportation, communication and utilities" use, much of which, even though classified as developed, is in fact vacant.

	Number of developed parcels	Aggregate land area of developed parcels (10^9 ft^2)	Market value of land per ft^2	Aggregate market value of land ($\$10^9$)
Los Angeles	1893741	38.63	4.62	178.65
Orange	680062	13.35	7.61	101.62
Riverside	632804	28.23	1.46	41.25
San Bernardino	645595	186.68	0.26	48.43
Ventura	246680	8.47	4.05	34.29
Total	4098882	275.36	1.47	404.24

Table 14: Imputed aggregate land values for developed properties.

It is of interest to compare the results for the aggregate market value of land for developable vacant properties (from Table 9) to aggregate market value of land for developed properties (Table 14). In Los Angeles County, for example, the estimated aggregate market value of developable vacant land was $14.72 billion, while that of developed land is $178.65 billion, a ratio of 0.082. As would be expected, the ratio is higher for the less developed counties (Riverside, Ventura, and San Bernardino) than for the more developed counties (Los Angeles and Orange).

5. *The Aggregate Value of Vacant and Developed Land*

This section collects results from the previous two sections. Table 16 repeats the estimates presented previously of aggregate land values for: i) single-family residential parcels; (ii) developed, non-SFR parcels; (iii) developable vacant land; (iii) undevelopable vacant land; and (iv) all land.

	ALV_{SFR} (\$$10^9$)	ALV_{NSFR} (\$$10^9$)	ALV_{DV} (\$$10^9$)	ALV_{NDV} (\$$10^9$)	Total ALV (\$$10^9$)
Los Angeles	110.21	68.43	14.72	18.08	211.44
Orange	66.76	34.86	3.97	47.34	152.93
Riverside	27.53	13.72	13.51	6.60	61.36
San Bernardino	31.63	16.80	5.52	4.81	58.76
Ventura	22.06	12.23	5.26	5.60	45.15
Total	258.19	146.05	42.98	82.44	529.66

Table 15: Aggregate land values imputed per regressions (1) and (2).
Notes: ALV denotes aggregate land value, SFR single family residential, NSFR developed, non-SFR, DV developable vacant, and NDV undevelopable vacant.

All the land values discussed to this point in the paper are rather peculiar. A parcel's land value imputed from (1) estimates "what the parcel's 2007 assessed value would have been had the parcel last been sold in 2000." We now convert this peculiar concept to one that is more familiar. First, we make an adjustment to estimate instead "what the parcel's assessed value immediately after sale would have been had the parcel last sold in 2000." Since, under Proposition 13, in principle at least, the State of California automatically reassesses all parcels, whether vacant or developed, at their sales prices immediately after sale, this is equivalent to estimating the parcel's market value in 2000.

Thus, we need to ascertain "what the parcel's assessed value immediately after sale would have been had the parcel last sold in 2000" from "what the parcel's 2007 assessed value would have been had the parcel last sold in 2000". As explained in detail in footnote 6, under California's Proposition 13 a parcel's assessed value in year t is determined by a formula that depends on the year of last sale and the sales price at that time. While the formula is quite complicated, over the period of interest it simplifies so that a parcel's estimated market value in 2000 equals "what the parcel's 2007 assessed value would have been had the parcel last sold in 2000" divided by $(1.02)^{2007\text{-}2000} = 1.149$. The adjustment therefore entails only application of this scaling factor. Applying this scaling factor, we obtain that estimated aggregate land value in 2000 as $(529.66 \text{ X } \$10^9)/1.02^7 = 461.10 \text{ X } \$10^9$.

Is \$461 billion dollars large or small? We find such large numbers easier to absorb when we choose an alternative metric. All previous studies, using other methods that were discussed in Section 2.2, found aggregate land value at the national level to be similar in magnitude to GNP. This is an intuitive metric one can relate to. Table 16 gives aggregate land value (taken from Table 15), aggregate 2000 land value, aggregate income (measured on a residency basis, as county income in the 2000 Census), and the ratio of aggregate 2000 land value to aggregate income, for the Greater Los Angeles Area, as well as for its constituent counties.

The ratio of aggregate land value to aggregate income is 1.114.

	ALV ($\$10^9$)	2000 ALV ($\$10^9$)	2000 Aggregate Income ($\$10^9$)	Ratio
Los Angeles	211.44	184.07	260.30	0.707
Orange	152.93	133.13	73.51	1.811
Riverside	61.36	53.42	28.89	1.849
San Bernardino	58.76	51.15	28.81	1.776
Ventura	45.15	39.31	22.32	1.761
Total	529.66	461.10	413.83	1.114

Table 16: The ratio of 2000 aggregate land value to 2000 aggregate income.
Note: Recall that aggregate income is measured on a residency basis.

It is interesting to speculate on why the ratio differs in the way it does across counties. One hypothesis is that the ratio of aggregate land value to aggregate income is highest for the more agricultural counties. Another hypothesis rests on a result from microeconomic theory, that, in the production of a particular good, the land share of income depends on the elasticity of substitution between land and other factors in production. If the elasticity of substitution is greater than one, then in locations where the factor rent on land is relatively high, the land value share is relatively low. Yet another hypothesis is that measuring income on a residency rather than on a source basis biases the ratio downwards for the bedroom counties. None of these hypotheses, however, explains why the ratio of aggregate land value to aggregate income is so much lower for Los Angeles County than that for the other four counties. Our conjecture is that the downward bias from estimating developed land values from vacant land values, controlling for observables, is greatest for Los Angeles County[23].

The ratio of aggregate 2000 urban land value (aggregate 2000 land value minus the aggregate 2000 land value of undevelopable parcels) to aggregate income equals the ratio of aggregate 2000 land value to aggregate income, 1.114, times the ratio of aggregate 2000 urban land value to aggregate 2000 land value, 0.844, yielding 0.940.

---

[23] A promising way to investigate this hypothesis is the "tear down" method. Since the value of a developed property, whose structures are about to be torn down with new structures built on site, is almost entirely in its land, compare the sales prices of developed properties that have been redeveloped to its imputed land values per our procedure

From the time dummy variables in regressions 1A and 1B, one can compute a vacant land value index[24]. Over the period from 1985 to 2007, this index was at its minimum in 2000 (see Figure A1), suggesting that the ratio of aggregate land value to aggregate income was considerably lower than average in 2000. Since we have consistently made assumptions that lead to more conservative estimates of aggregate land values, we are confident that 1.114 is a lower bound estimate of the average ratio of aggregate land value to aggregate income over time in the Greater Los Angeles, at least until the property market crash in 2008.

Another way of expressing aggregate numbers intuitively is in per capita terms. Table 17 gives population, per capita income and land value per capita for each of the five counties and then for the Greater Los Angeles Region in 2000. In 2000 land value per capita for the Region was $28,164, was lowest in Los Angeles County at $19,337, and was highest in Ventura County at $52,199.

	Population (10^6)	Per capita 2000 income, $	Per capita 2000 land value, $
Los Angeles	9.519	27,350	19,337
Orange	2.846	25,830	46,780
Riverside	1.545	18,700	34,574
San Bernardino	1.709	16,860	29,932
Ventura	0.753	29,640	52,199
Total	16.372	25,280	28,164

Table 17: Population, per capita income, and land value per capita in 2000

Appendix D presents our estimates of aggregate land values by major land use categories. Perhaps the most interesting results relate to single-family residential. According to our calculations, single-family residential constitutes 63.9% of the total developed land value, 10.5% of total developed land area, 48.75% of total land value, developed and vacant, and 2.81% of total land area, developed and vacant. Furthermore, the average land value per ft^2 is higher for single-family residential than for any other major land use category, including retail and office. Though this jars with our intuition, we have no plausible explanation for why our estimate for single-family residential land value should be biased upward relative to that for other developed land use categories.

---

[24] This is not a conventional land value index. See Appendix A for a discussion of how to interpret this index, and how to convert it to a conventional land value index.

6. Some Robustness Checks

There are many, many robustness checks that we could have undertaken, most of which relate to the poor quality and incompleteness of the SCAG parcel database. We could, for instance, have collected information, from other data sources, on characteristics of vacant land not included in the SCAG parcel database, such as steepness of the terrain and access to water, both of which are important determinants of vacant land values in southern California, and on characteristics of developed land not included in the SCAG parcel database, such as standard explanatory variables used in hedonic valuation regressions for single-family housing. We attempted neither simply because of the huge size of the database compared to our modest budget. We could have undertaken such exercises for a sample of parcels, such as for a single city, but this would have raised issues concerning the representativeness of our sample. We could also have undertaken more extensive ground-truth tracking to check the accuracy of data obtained from the county assessment offices. We could also have pushed harder in attempting to uncover actual assessment practice, in contrast to the rules laid out in the State's instructions on assessment. But again we were constrained by our limited resources.

We decided instead to focus on the robustness of our estimates to, first, spatial autocorrelation, and, second, to the use of sales prices rather than assessed values.

6.1 Spatial autocorrelation

If the object of our analysis were the accurate valuation of individual properties, spatial autocorrelation would have been of central importance. But the object of our analysis has instead been to estimate the *aggregate* value of land. It is our intuition that accounting for spatial autocorrelation would amplify the bumps on the land value surface without substantially affecting the volume under the surface.

As far as we are aware, the techniques for generating out-of-sample predictions are not well developed in the spatial econometrics literature, the focus of which has been primarily to obtain unbiased estimates of regression coefficients. However, making out-of-sample prediction is the key to imputing aggregate land value in this paper. To impute land value for vacant parcels for which data on either assessed value or sale year is missing is to make out-of-sample predictions. To impute land value for developed parcels from vacant land sales and assessments is also to make out-of-sample predictions too.

It is also computationally challenging, if not infeasible, to run standard spatial regressions in this paper. In the OLS regressions for equation (1) on vacant parcels, there are 99169 developable vacant parcels in the sample of regression (1A) and 94368 undevelopable vacant parcels in the sample of regression (1B). In each of the two regressions, there are more than two hundred explanatory variables of which the coefficients need to be estimated, including 173 city dummies, 29 sale year dummies, 21 (1A)/14 (1B) current land use dummies, and 38 planned land use dummies. The usual sample size of spatial regressions in the literature ranges from hundreds to several thousands, and with many fewer explanatory variables. Regression on a more disaggregated level such that spatial regression is computationally feasible is unreliable, since the large number of explanatory variables cannot be reliably estimated without a large sample.

Ultimately we decided not to generate estimates of aggregate land value using the spatial econometrics techniques that we are familiar with. Perhaps the computational burden could be made manageable by the use of more sophisticated and in principle computationally efficient techniques, such as those introduced in Lesage & Pace (2004) and Bourassa *et al* (2010), or by the use of geo-statistical software like GeoDa and ArcGIS. But since neither of us has the programming skills to implement these techniques, we chose to leave the estimation of aggregate land values using spatial econometric techniques to experts in spatial econometrics.

6.2 Using sales prices rather than assessed values throughout

We explained earlier why we decided to use assessed values rather than sales prices. In the SCAG parcel database, reliable data on sales prices are available for only Orange and Riverside Counties, and for these two Counties, data on assessed values are more complete than data on sales prices. Also, when we made this decision, we believed, on the basis of the State's documentation of its assessment practices, that all classes of property are assessed at market value. However, when we obtained results on vacant land values, we began to question this assumption.

Most other states provide preferential property tax treatment to agricultural land with the aim of preserving family farms. They do this by taxing agricultural land on the basis of agricultural income; in particular, the property tax assessed on a parcel in agricultural use is based on its "agricultural value" -- "what the market price of the land would be if it were held in agriculture forever" -- rather than on its market value. Might not the State of California do this too, claiming that it this is consistent with the income method of market value assessment?

We took Riverside County as an example to see how the aggregate land value of vacant parcels imputed using sales price differs from that imputed using assessed value.

The general procedure we employ in imputing the 2000 aggregate vacant land value in Riverside County using sales price data is as follows:

- For all developable vacant parcels in Riverside County that contain data on sales price and assessed value and the year of most recent sale (and if it is no earlier than 1980)[25], regress the natural log of the parcel's sales price per equation (3).

[25] In the rest of this subsection, when we say that a vacant parcel has "complete" data on both assessed value and sales price, we shall mean that it has data on assessed value *and* sales price *and* on year of most recent sale *and* the year of most recent sales was later than 1980.

Equation (3) is the same as equation (1) except that the dependent variable in equation (3) is natural log of sales price and the sample is different.
- For the same set of parcels, use the coefficients on time dummies in that equation to estimate each parcel's 2000 market value.
- For the complementary set of developable vacant parcels in Riverside County, use the regression equation to impute each parcel's 2000 market value.
- Repeat the procedure for undevelopable vacant parcels in Riverside County.

Analogously, the general procedure we employ in imputing the 2000 aggregate vacant land value in Riverside County using assessed value data is as follows:

- For all developable vacant parcels in Riverside County that contain data on sales price and assessed value and the year of most recent sale (and if it is no earlier than 1980), regress the natural log of 2007 assessed parcel value per equation (4). Equation (4) is exactly the same as equation (1), except the sample employed.
- For the same set of parcels, use the coefficients on time dummies in that equation to estimate each parcel's 2007 assessed value had it been sold on 2000.
- For the complementary set of developable vacant parcels in Riverside County, use the regression equation to impute each parcel's 2007 assessed value had it been sold in 2000.
- For all developable vacant parcels in Riverside County, compute each parcel's 2000 market value by dividing the 2007 assessed value had it been sold in 2000 by the scaling factor of 1.02^7.
- Repeat the procedure for undevelopable vacant parcels in Riverside County.

The details of the above two procedures are analogous to those explained in detail in section 3. Full regression results of (3) and (4) are reported in appendix B.

We shall now compare, by developability, for the subset of vacant parcels in Riverside County that have complete data on both sales price and assessed value, the aggregate vacant land value imputed using sales price data to that imputed using assessed value data. This subset of vacant parcels corresponds exactly to the regression samples of (3) and (4). The estimated 2000 sales price of a parcel in this subsample is obtained simply by inflating or deflating the available sales price data using coefficients of sale year dummies in equation (3). The estimated 2000 assessed value of a parcel in this subsample is obtained by inflating or deflating the available assessed value data using coefficients of sale year dummies in equation (4), and then dividing by the scale factor of 1.02^7.

Table 18 records the relevant results. Column 1 gives the number of vacant parcels in this subsample, and column 2 gives the aggregate land area of vacant parcels in this subsample. Column 3 gives the 2000 aggregate land value imputed from sales price (ALV_{SP}) in this subsample. Column 4 gives the 2000 aggregate land value imputed from assessed value (ALV_{AV}) in the subsample. Column 5 gives the ratio of column 3 to column 4.

	Number of parcels	Aggregate area of parcels (10^9 sq. ft)	2000 aggregate land value estimated from sales price (\$$10^9$)	2000 aggregate land value estimated from assessed value (\$$10^9$)	ALV_{SP} / ALV_{AV}
Developable Vacant	27239	4.29	24.95	7.07	3.53
Undevelopable Vacant	14811	7.72	12.49	2.54	4.91

Table 18: The comparison of 2000 aggregate vacant land value estimated from sales price to that estimated from assessed value, for the subset of vacant parcels in Riverside County that have complete data on both sales price and assessed value.

We shall now compare, by developability, for the complementary set of vacant parcels in Riverside County that do not have complete data on both sales price and assessed value, the aggregate vacant land value imputed using sales price data to that imputed using assessed value data. Recall that this subset of vacant parcels corresponds exactly to the parcels that are not included in the regression samples of (3) and (4). The estimated 2000 sales price of a parcel in this subsample is imputed using regression (3A) and (3B) and setting its sale year to 2000. The estimated 2000 assessed value of a parcel in this subsample is imputed using regressions (4A) and (4B), by setting the sale year to 2000 and dividing the results by the scale factor of 1.02^7.

Table 19 corresponds to Table 18 exactly, but applies to the set of vacant parcels in Riverside County that do not have complete data on both sales price and assessed value.

	Number of parcels	Aggregate area of parcels (10^9 sq. ft)	2000 aggregate land value imputed from sales price (\$$10^9$)	2000 aggregate land value imputed from assessed value (\$$10^9$)	ALV_{SP} / ALV_{AV}
Developable Vacant	61069	14.76	29.13	11.34	2.57
Undevelopable Vacant	57227	148.92	42.08	23.59	1.78

Table 19: The comparison of 2000 aggregate vacant land value imputed from sales price to that imputed from assessed value, for the subset of vacant parcels in Riverside County that do not have complete data on both sales price and assessed value.

Table 20 combines results from Table 18 and Table 19, and reports on the comparison of aggregate land value imputed from sales price data with that imputed from assessed value data for all vacant parcels in Riverside County, using regressions (3) and (4).

	Number of parcels	Aggregate area of parcels (10^9 sq. ft)	2000 aggregate land value imputed from sales price ($\$10^9$)	2000 aggregate land value imputed from assessed value ($\$10^9$)	ALV_{SP} / ALV_{AV}
Developable Vacant	88308	19.05	54.08	18.41	2.94
Undevelopable Vacant	72038	156.64	54.57	26.13	2.09

Table 20: The comparison of 2000 aggregate vacant land value imputed from sales price to that imputed from assessed value, for all vacant parcels in Riverside, using regressions (3) and (4).

Recall that earlier in sections 3 and 5 we reported the estimated aggregate values of vacant parcels in 2000 in Riverside County using assessed value data too. For Riverside County, the 2000 aggregate land value of developable vacant parcels imputed using assessed value was at $13.51/1.02^7 = 11.76$ billion dollars (Table 9), and the 2000 aggregate land value of undevelopable vacant parcels imputed from assessed value was $6.60/1.02^7 = 5.74$ billion dollars (Table 10). The aggregate vacant land values imputed from assessed value reported in Table 20 differ from those reported in those earlier sections because they are imputed from different regressions. Those values in sections 3 and 5 were imputed from regressions (1), while the corresponding values in Table 20 were obtained from regressions (4). Regressions (4) are exactly the same as regressions (1) except for the regression sample employed. The regression samples of (1) include all vacant parcels in the five counties with complete data on assessed value, while the regression samples of (4) is smaller, including only the vacant parcels in Riverside County that have complete data on both sales price and assessed value.

The aggregate value of developable vacant parcels in Riverside County imputed from (4A) is 18.41 billion dollars, which is 1.57 times that obtained from (1A), while the aggregate value of undevelopable vacant parcels in Riverside County imputed from (4B) is 26.13 billion dollars, which is 4.55 times that imputed from (1B). This indicates that, controlling for observables, in Riverside County at least, vacant parcels with complete data on both assessed value and sales price are of higher quality on average than vacant parcels with complete data on only assessed value. Whether this is due to a particular assessor employing different practice for the two classes of parcels, or due to assessors working in different parts of the County employing different practice, is impossible to ascertain, though intuitively the latter seems more likely. Whichever is the case, our finding casts doubt on the consistency of assessment practices in the County.

On the basis of these findings, it is reasonable to conjecture that vacant parcels for which neither assessed value nor sale prices is recorded are of lower quality than vacant parcels for which both sales price and assessed value are recorded, and for vacant parcels for which only assessed value is recorded. If this conjecture is correct, then our imputation procedures overestimate both the assessed values and sales prices of parcels for which neither assessed value nor sales price is recorded.

The degree of underassessment of the value of vacant parcels also differs across land use types. Appendix B provides more information on that.

7. Conclusion

Knowing the aggregate value of land is important in a variety of contexts. The first is macroeconomic. There are important interactions between the land market and the rest of the economy. Shocks from outside the real property market can affect the asset value of real property, generating wealth effects, which is a potentially important channel through which the shocks can affect aggregate economic performance. Shocks from inside the real property market can affect aggregate economic performance too, as is evidenced by the recent Great Recession. These effects are quantitatively more important the larger is the share of the nation's wealth that is held in the form of real property, including vacant land.

The second context in which knowing the aggregate value of land is important is fiscal. Most economists favor increased use of land taxation in principle. But practically they have had two main objections, the difficulty in assessing land value and the small size of the land tax base. As we have seen, the first objection remains, though with increased use of GIS and of parcel databases, it is considerably less severe than it used to be. The second objection would be largely neutralized if it were demonstrated that the land tax base is considerably larger than is conventionally believed.

The third is theoretical. The larger is the aggregate value of land, the more remiss is mainstream economic theory in ignoring land and space. The fourth is empirical. Getting the numbers right is important, especially when evaluating the effects of economic policies.

The original aim of this paper was to estimate the aggregate value of all land, and also of all urban land, in the Greater Los Angeles Region in 2000 from a comprehensive land parcel database assembled by the South California Association of Governments (SCAG). This is the first paper, to our knowledge, to attempt to estimate the aggregate value of urban land from a land parcel database. We conjectured that other methods that have been employed to estimate the aggregate value of land, which rely on national accounts data, seriously underestimate the value. Parcel databases hold great promise and have been underexploited by researchers. The paper has been less successful than we had hoped in meeting this original aim because of the many problems we encountered with the assessment data in the SCAG parcel database. But there are offsetting benefits. Knowledge of the pitfalls we have uncovered in the use of parcel databases should prove useful to the increasing number of researchers who are using them[26]. This knowledge should also convey a strong *caveat emptor* -- distrust the empirical results of all studies that use parcel databases unless the researchers have not only acknowledged the problems with the data but also have also devoted considerable time and effort to attempting to deal

[26] Private companies such as DataQuick are now selling parcel data. Whether their quality is higher or lower than the quality of those assembled by regional planning agencies is unknown. The same caveat applies. Distrust the validity of all data unless it is well documented and unless strong evidence is provided that its accuracy has been checked.

with them. Knowledge of the problems with parcel databases will hopefully induce researchers, policy analysts, and policy makers to put pressure on the regional planning agencies that assemble these databases to put more effort into ensuring their accuracy, on assessment offices to fulfill their legal obligation to provide complete and accurate data, as well as full documentation of their assessment practices, and on the relevant governmental bodies to provide the increased funding needed to improve the data quality in parcel databases.

Knowing that the SCAG parcel database has fields for the sales price and date of the most recent sale, as well as on current and plan land use, originally we had intended to estimate a land value surface, differentiated by plan land use on the basis of sales of vacant land, and then to impute land values to developed parcels from the estimated land value surface (with the current land use replacing plan land use). Time dummies for the date of land sale would have allowed us to bring all land values to a common base year, 2000. We recognized that this would result in underestimation since, first, holding constant all available descriptive characteristics of a land parcel, higher quality parcels tend to be developed earlier (so that the land value of a developed parcel with a given set of available descriptive characteristics will tend to be higher than that of a vacant parcel with the same available characteristics), and second, 2000 was a low point in the land value cycle. Unfortunately, we found that vacant land sales price data are available for only two of the counties in the five counties of the LA Metro Area, and in those counties, the data were seriously incomplete. Data were, however, available on assessed value for all five counties in the LA Metro Area, and, though incomplete, they were not as incomplete as the data on the most recent sale. Since the State of California in principle uses market value assessment, and since the formula that the State employs (under Proposition 13) to update assessed value from one year to the next is known, we applied the procedure we had originally intended to apply to sales price data, to assessed value data instead. The bulk of the paper reports on the details of this procedure, and on the results of applying it. Subsequently, we checked on the soundness of our assumption that the assessed values of vacant land conform well to their sales prices, and found that, the principle of market value assessment notwithstanding, vacant land is systematically underassessed, especially for some land uses such as farming.

Fortunately, the three major sources of bias (imputing land values for vacant properties to developed properties, using the year 2000, and using assessed value rather than sales price) that we identified all result in underestimation of the aggregate value of land. Thus, we are confident that our central estimates of the aggregate value of land and of urban land provide a lower bound on the true value. These values are 1.114 and 0.940 times as large as aggregate regional income in 2000, which are higher than previous estimates from national accounting data. Thus, we are confident in asserting that the aggregate value of vacant land is higher than previous estimates. We are not however confident in estimating the degree of downward bias in our estimates for an average year. Our conjecture is that the degree of downward bias is substantial, and therefore that the true aggregate value of land is several times the previous estimates of the aggregate value of land, based on national accounts data. If our conjecture is correct, then land plays a quantitatively substantially more important role in the national economy than currently believed, and also the taxable capacity of land is considerably larger than previously believed.

We hope that our paper will stimulate other researchers to take over from where we left off, using other regional planning authorities' land parcel databases, as well as land parcel databases assembled and marketed by private data companies. Such parcel databases hold great promise to provide researchers, planners, and policy makers with more spatially detailed data on regional land markets, which should prove very valuable in the planning process. But for such databases to realize their potential, considerably more effort needs to be expended in ensuring their accuracy and completeness, which will require more transparency in assessment practice.

References

Angel, S., Parent, J., Civco, D., & Blei, A. (2012). *Atlas of Urban Expansion.* Cambridge, MA: Lincoln Institute of Land Policy.

Arnott, R. (1979). Optimal city size in a spatial economy. *Journal of Urban Economics, 6*(1), 65-89.

Arnott, R. (2004). Does the Henry George theorem provide a practical guide to optimal city size? *American Journal of Economics and Sociology, 63*(5), 1057-1090.

Arnott, R. (2005). Neutral property taxation. *Journal of Public Economic Theory, 7*, 27-50.

Arnott, R. & Guo, Y. (2012). Developable Vacant Land in the LA Metro Area - An Economic Approach. *MPRI working paper 2012-8*.

Arnott, R., & Lewis, F. (1979). The transition of land to urban use. *Journal of Political Economy* 87: 161-169.

Arnott, R. J., & Stiglitz, J. E. (1979). Aggregate land rents, expenditure on public goods, and optimal city size. *The Quarterly Journal of Economics, 93*(4), 471-500.

Bourassa, S. C., Cantoni, E., & Hoesli, M. (2010). Predicting house prices with spatial dependence: a comparison of alternative methods. *Journal of Real Estate Research*, *32*(2), 139-159.

California, State Board of Equalization. 2012. California Property Tax: An Overview. Publication 29.

Davis, M. A., & Heathcote, J. (2007). The price and quantity of residential land in the United States. *Journal of Monetary Economics, 54*(8), 2595-2620.

Davis, M. A., & Palumbo, M. G. (2008). The price of residential land in large US cities. *Journal of Urban Economics, 63*(1), 352-384.

Flatters, F., Henderson, V., & Mieszkowski, P. (1974). Public goods, efficiency, and regional fiscal equalization. *Journal of Public Economics, 3*(2), 99-112.

Kanemoto, Y., Ohkawara, T., & Suzuki, T. (1996). Agglomeration economies and a test for optimal city sizes in Japan. *Journal of the Japanese and International Economies, 10*(4), 379-398.

LeSage, J. P., & Pace, R. K. (2004). Models for spatially dependent missing data. *The Journal of Real Estate Finance and Economics*, *29*(2), 233-254.

Nichols, J. B., Oliner, S., & Mulhall, M. R. (2010). *Commercial and residential land prices across the United States*. Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board.

Nichols, J. B., Oliner, S. D., & Mulhall, M. R. (2013). Swings in commercial and residential land prices in the United States. *Journal of Urban Economics*, *73*(1), 57-76.

Saiz, A. (2010). The geographic determinants of housing supply. *Quarterly Journal of Economics* 125: 1253-1296.

Serck-Hanssen, J. (1969) The optimal number of factories in a spatial market. In H.C. Bos, ed. *Towards Balanced International Growth*. Amsterdam: North-Holland, 269-281.

Shoup, D. C. (1970). The optimal timing of urban land development. *Papers in Regional Science*. 25: 33-44.

Starrett, D. A. (1974). Principles of optimal location in a large homogeneous area. *Journal of Economic Theory, 9*(4), 418-448.

Vickrey, W. (1977). The city as a firm. *The economics of public services*, 334-343.

World Bank. (2006). *The Changing Wealth of Nations: Measuring Sustainable Development in the New Millenium.* Washington, DC: World Bank Press.

Zhang, H. & Arnott, R. (2011). Value and Rent per ft2 of Non-residential and Non-vacant Parcels at Model Zone Level. *MPRI working paper 2011-7*.

Zhang, H. & Arnott, R. (2012). Value and Rent per ft2 of Residential Parcels at Model Zone Level. *MPRI working paper 2012-2*.

Zhang, H. & Arnott, R. (2013). The Aggregate Value of Urban Land in the Greater Loss Angeles Region. Earlier draft of current paper.

Appendix A
Regression Results

To impute the land value of vacant parcels, we ran two regressions in the form of equation (1), one for developable vacant parcels and the other for undevelopable vacant parcels. Table A1 presents the full regression coefficients for each of developable vacant parcels and undevelopable vacant parcels. Note that a dummy variable has a blank regression coefficient when the corresponding regression sample does not contain data points for which the dummy variable equals one. All distance measures are in miles.

Vacant land value indexes by developability can be constructed from the two regressions in Table A1. The coefficient on dummy sale year t relates the 2007 assessed land value of a vacant parcel sold in year t to the 2007 assessed land value for the same vacant parcel sold in the base year 2000. In particular, taking 2000 as the reference year, the developable land value index for year t is calculated as the exponential of the coefficient on the sale year dummies in the regressions for developable vacant parcels. The undevelopable land value index is calculated in the same way from the regression for undevelopable vacant parcels. Figure A1 shows the two land value indexes.

The index takes into account that, under Proposition 13, a parcel is reassessed at its sale price when sold, with its assessed value between sales being determined by a formula that depends on the number of years since the parcel's last sale and its sale price at that time (recall footnote 8). Thus, the index is not a standard price or value index. The index measures "what the 2007 assessed value would have been if the land had been sold in year t rather than the year 2000". A standard value index would measure "what the assessed value immediately after sale would have been immediately after had the land been sold in year t rather than in the year 2000." Thus, to convert the derived index to a standard value index, one needs to multiply the derived index by "what the assessed value immediately after sale would have been immediately after sale" divided by the assessed value in 2007. Applying the Proposition 13 formula for the rate of increase in assessed value between sales, for a property that most recently sold in year t the adjustment factor is 1.02^{t-2007}.

Readers should be careful that the land value index is meaningful only for comparing land values of the same developability across years. Comparing the value of two indexes at a given year across developability is inappropriate.

Table A1: Full regression results on assessed value of developable and undevelopable vacant parcels per regressions (1A) and (1B)

	Dependent variable: ln (Assessed value)	
	Developable vacant parcels	Undevelopable vacant parcels
ln (land area)	0.639***	0.583***
Distance to the nearest sub-center	-0.037***	0.007***
Distance to the CBD	-0.016***	0.006***
Squared distance to the CBD	4.16E-5***	-1.01E-4***
Distance to the nearest freeway or highway	0.017***	-0.010***
Distance to the nearest coast	-0.019***	-0.051***
Squared distance to the nearest coast	9.02E-5***	1.90E-4***
city: Adelanto (base category)		
city: Agoura Hills	-0.279	-1.037***
city: Alhambra	1.123	7.877***
city: Aliso Viejo		0.511
city: Anaheim		3.485*
city: Apple Valley	0.683*	0.071
city: Artesia		4.585***
city: Avalon		0.542*
city: Azusa	-3.858***	-3.356***
city: Baldwin Park	1.149	6.178***
city: Banning	0.591	2.966***
city: Barstow	1.903***	
city: Beaumont	-0.815	-1.626***
city: Bell Gardens	1.068*	2.765***
city: Bellflower		3.013**
city: Beverly Hills	1.676***	
city: Big Bear Lake	2.391***	6.285***
city: Blythe	0.019	5.590***
city: Bradbury	0.095	2.064**
city: Brea	0.044	-0.346
city: Burbank		1.094***
city: Calabasas	-4.811***	-0.578***
city: Calimesa	0.79	-1.308***
city: Camarillo	0.68	
city: Canyon Lake	2.645	0.067
city: Carson	0.437	-1.451***
city: Cathedral City	0.079	2.156***

city: Cerritos	-0.152	-3.717***
city: Chino	1.508***	0.064
city: Chino Hills	-0.482	-0.151
city: Claremont	0.477	0.573*
city: Coachella	0.694**	2.982***
city: Colton	-1.827***	3.651***
city: Commerce	1.647**	5.517***
city: Compton	-0.984**	
city: Corona	-0.017	-1.992***
city: Costa Mesa	0.709	2.742***
city: Covina	1.410**	4.795***
city: Cudahy	-0.018	6.678***
city: Culver City	1.025*	6.429***
city: Cypress	0.39	
city: Dana Point	2.292***	4.915***
city: Desert Hot Springs	0.38	-1.141***
city: Diamond Bar	1.133***	5.958***
city: Downey	0.476	
city: Duarte	1.686*	-0.326
city: El Monte	0.133	-2.900*
city: El Segundo	0.717	5.991***
city: Fillmore	0.598	-2.045***
city: Fontana	0.759**	0.600**
city: Fullerton	1.182	
city: Garden Grove	1.007	
city: Gardena	-1.353**	1.508
city: Glendale	-0.518	-0.186
city: Glendora	-0.877**	0.027
city: Grand Terrace	-0.167	4.512**
city: Hawthorne	-1.436*	-0.026
city: Hemet	-0.064	-1.023***
city: Hermosa Beach		3.774***
city: Hesperia	-0.513	4.428***
city: Hidden Hills	-0.985	-2.647***
city: Highland	0.109	4.792***
city: Huntington Beach	0.442	
city: Huntington Park		3.242***
city: Indian Wells	0.898**	4.456***
city: Indio	0.116	6.188***
city: Industry	0.488	
city: Inglewood	-0.195	0.572
city: Irvine	-0.039	3.718***
city: Irwindale	1.57	3.171***

city: La Canada Flintridge	-2.153**	-0.043
city: La Habra	-0.357	2.716**
city: La Habra Heights	-0.102	-0.139
city: La Mirada	-1.455*	-3.305***
city: La Puente	0.418	-5.491***
city: La Quinta	-0.111	0.545*
city: La Verne	-1.491***	-0.490*
city: Laguna Beach	1.432***	1.376*
city: Laguna Hills	1.958	1.886***
city: Laguna Woods	1.259	4.881***
city: Lake Elsinore	-0.823**	2.060***
city: Lake Forest	1.146	1.586***
city: Lakewood	0.439	
city: Lancaster	1.357***	0.02
city: Loma Linda	1.147***	3.631*
city: Lomita	0.699	2.473
city: Long Beach	0.705	4.783***
city: Los Alamitos		0.975
city: Los Angeles	-0.293	-0.172*
city: Lynwood	0.454	
city: Malibu	1.389***	4.442***
city: Manhattan Beach	1.498	-0.184
city: Maywood	2.509	-0.18
city: Menifee	0.109	-1.562***
city: Mission Viejo	0.755	1.511***
city: Monrovia	0.501	3.933***
city: Montclair	0.721	
city: Montebello	1.012**	3.647*
city: Monterey Park	1.403***	3.860**
city: Moorpark	1.348**	3.086***
city: Moreno Valley	-0.172	-0.808***
city: Murrieta	0.715**	2.105***
city: Needles		6.958***
city: Newport Beach	0.918	1.011
city: Norco	0.138	-1.847***
city: Norwalk	0.927	
city: Ojai	1.507**	4.434***
city: Ontario	-0.101	
city: Orange	0.053	-0.655
city: Oxnard	1.383***	1.767
city: Palm Desert	1.086***	3.714***
city: Palm Springs	0.15	-0.453**
city: Palmdale	0.837**	-0.087

city: Palos Verdes Estates		5.811***
city: Paramount	-0.148	1.674***
city: Pasadena	-0.007	1.315**
city: Perris	0.636*	2.920***
city: Pico Rivera	0.102	-8.236***
city: Pomona	-0.134	4.101***
city: Port Hueneme	-0.252	
city: Rancho Cucamonga	0.935***	5.937***
city: Rancho Mirage	0.809**	3.571***
city: Rancho Palos Verdes	-1.224***	-4.247**
city: Rancho Santa Margarita	0.403	-0.49
city: Redlands	0.859**	1.547
city: Redondo Beach	1.385	1.973**
city: Rialto	1.393***	2.256***
city: Riverside	-0.341	-1.548***
city: Rolling Hills	0.005	
city: Rolling Hills Estates	0.461	1.458
city: Rosemead	1.15	1.457
city: San Bernardino	0.489	0.491
city: San Buenaventura	0.960**	3.710***
city: San Clemente	1.633	2.965**
city: San Dimas	-0.387	-3.997***
city: San Fernando	1.738	
city: San Gabriel	0.681	0.077
city: San Jacinto	0.346	-0.338*
city: San Juan Capistrano	1.766***	-0.099
city: Santa Ana		0.088
city: Santa Clarita	-0.336	4.270***
city: Santa Fe Springs	0.295	
city: Santa Monica	0.86	2.366***
city: Santa Paula	0.483	
city: Seal Beach	1.012**	1.579
city: Sierra Madre		0.031
city: Signal Hill	-1.418***	2.103**
city: Simi Valley	0.063	-2.132***
city: South El Monte	0.121	3.995***
city: South Gate	0.740*	
city: South Pasadena	-0.674	4.296**
city: Stanton	1.301	
city: Temecula	0.221	-0.915***
city: Thousand Oaks	0.426	-0.706
city: Torrance	-0.079	4.094***
city: Tustin	0.908**	6.195***

city: Twentynine Palms	0.908**	3.585***
city: Upland	0.777**	3.208***
city: Victorville	1.318***	4.937***
city: Walnut	-0.142	1.325*
city: West Covina	-0.072	1.166**
city: West Hollywood	2.686***	2.479*
city: Westlake Village	-1.516*	0.131
city: Westminster	-0.106	
city: Whittier	1.060**	-0.501**
city: Wildomar	0.359	-1.396***
city: Yorba Linda	-1.191	
city: Yucaipa	0.838**	1.765***
city: Yucca Valley	1.149**	3.621***
city: unincorporated_la	0.301	3.607***
city: unincorporated_or	0.033	-0.788***
city: unincorporated_rv	0.187	2.688***
city: unincorporated_sb	0.259	3.883***
city: unincorporated_vn	0.394	-1.068***
sale year: 1980	-0.261**	0.093
sale year: 1981	0.114	0.196***
sale year: 1982	0.156*	0.183***
sale year: 1983	-0.027	0.104*
sale year: 1984	0.042	-0.065
sale year: 1985	0.089	0.157***
sale year: 1986	0.069	0.291***
sale year: 1987	0.133	0.362***
sale year: 1988	0.301***	0.562***
sale year: 1989	0.615***	0.917***
sale year: 1990	0.627***	0.923***
sale year: 1991	0.400***	0.621***
sale year: 1992	0.136	0.592***
sale year: 1993	0.048	0.284***
sale year: 1994	0.072	0.450***
sale year: 1995	0.157	0.306***
sale year: 1996	0.142	0.152**
sale year: 1997	0.098	0.262***
sale year: 1998	0.014	0.162***
sale year: 1999	-0.099	0.105*
sale year: 2000 (base category)		
sale year: 2001	0.071	0.083*
sale year: 2002	0.213***	0.150***
sale year: 2003	0.279***	0.222***
sale year: 2004	0.269***	0.370***

sale year: 2005	1.217***	0.832***
sale year: 2006	1.068***	0.899***
sale year: 2007	0.932***	0.744***
sale year: 2008	0.307*	0.344***
land use: Former Military Base - Built-Up Area	0.394	
land use: Former Military Vacant Area	0.229	
land use: Under construction	1.292***	
land use: open space and recreation	-0.645	-3.150***
land use: Golf courses	0.230***	
land use: Local parks and recreation		-0.627***
land use: Developed local parks and recreation		-4.163***
land use: Regional parks and recreation		-0.142
land use: Developed regional parks and recreation		-2.960***
land use: Undeveloped regional parks and recreation		-3.274***
land use: Cemeteries	0.151	
land use: Wildlife preserves and sanctuaries		0.001
land use: Specimen gardens and arboreta		2.108*
land use: Beach parks		-1.945***
land use: Other open space and recreation		-2.352***
land use: Agricultural	-0.034	
land use: Cropland and improved pasture land	0.685	
land use: Irrigated Cropland and Improved Pasture Land	-0.380***	
land use: Non-Irrigated Cropland and Improved Pasture Land	-0.294***	
land use: Orchards and Vineyards	-0.122*	
land use: Nurseries	0.159	
land use: Dairy, Intensive Livestock, and Associated Facilities	-0.074	
land use: Poultry Operations	0.629*	
land use: Other Agriculture	0.470***	
land use: Horse Ranches	0.480***	
land use: Vacant (base category)		
land use: Vacant Undifferentiated	-0.439***	-4.287***
land use: Abandoned Orchards and Vineyards	-0.171	
land use: Vacant With Limited Improvements	0.121*	
land use: Beaches Vacant	0.747***	
planned land use: missing (base category)		
planned land use: Residential	0.624	0.811*
planned land use: Single Family residential	0.173***	0.192***
planned land use: Multi-family residential	-0.160***	0.343***
planned land use: Mobile homes and Trailer parks	-0.340***	-0.382***
planned land use: Commercial and services	0.502***	1.040***

planned land use: Office	0.559***	0.890***
planned land use: Retail stores and commercial stores	0.304***	0.963***
planned land use: Other commercial	0.519***	0.983***
planned land use: Hotels and motels	-0.131	0.828*
planned land use: Public facilities	-0.948*	7.103***
planned land use: Special use facilities	0.027	-0.960***
planned land use: Educational facilities	-0.660***	-0.391***
planned land use: Military installations	-0.517	-1.799***
planned land use: Former Military Vacant Area	-0.149	
planned land use: Former Military Vacant Area	-1.602	
planned land use: Industrial	0.185*	0.017
planned land use: Light industrial	0.026	0.007
planned land use: Heavy industrial	0.12	-0.292***
planned land use: Extraction	-0.257	1.293*
planned land use: Wholesaling and warehousing	-0.345*	-0.439***
planned land use: Wholesaling and warehousing	1.975	
planned land use: Transportation, Communication, and Utilities	-1.686	
planned land use: Transportation	0.737**	0.466*
planned land use: Communication facilities	-1.507***	-0.354***
planned land use: Utility facilities	-0.891	-2.365*
planned land use: Mixed commercial and industrial	0.076	-0.585***
planned land use: Mixed urban	0.238***	0.587***
planned land use: open space and recreation	-1.404***	-0.890***
planned land use: Golf courses	0.081	-1.570***
planned land use: Local parks and recreation	-1.192***	-0.673***
planned land use: Regional parks and recreation	-2.008***	-0.524
planned land use: Cemeteries	1.129	1.357*
planned land use: Wildlife preserves and sanctuaries	-1.265***	-0.958***
planned land use: Beach parks		-1.239**
planned land use: Other open space and recreation	-2.182***	-1.236***
planned land use: Agricultural	-0.400***	-0.402***
planned land use: Vacant	-0.971	
planned land use: Water	-2.207***	-1.625***
Constant	7.243***	7.232***
Adjusted R-squared	0.264	0.405
Number of observations	99169	94368

* p<0.05, ** p<0.01, *** p<0.001

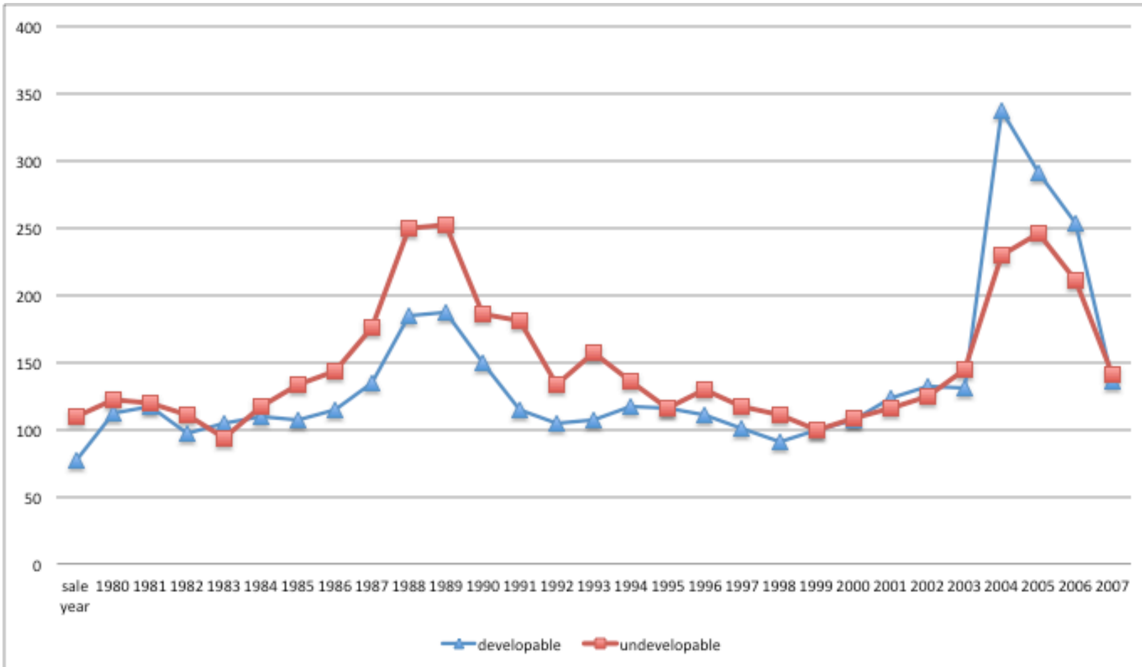

Figure A1: Land value index for developable and undevelopable land (nominal $).Year 2000 is the reference year with both developable and undevelopable land value indexes being 100.

Appendix B
Results with Sales Price compared to Assessed Value

To compare the 2000 aggregate vacant land value imputed from assessed value to that imputed from sales price in Riverside County, we ran two sets of regressions. In the first set, the dependent variable is the natural logarithm of sales price, and in the second set the dependent variable is the natural logarithm of 2007 assessed value. Within each set, we ran two regressions, one for developable vacant parcels and the other for undevelopable vacant parcels.

The full regression coefficients for developable vacant parcels are presented in table B1, and the full results for undevelopable vacant parcels are presented in table B2. In each table, column 1 reports on the regression coefficients from sales price data, and column 2 reports on the regression coefficients from assessed value data. The sample in each regression was all parcels, either developable or undevelopable as the case may be, with complete data on both sales price and assessed value. Thus, for example, when the dependent variable is the sales price of developable vacant parcels, the sample is all developable vacant parcels with complete data on both sales price and assessed value.

The degree of underassessment (measured as aggregate 2000 sales price divided by aggregate 2000 assessed values) varies across detailed land use types. Table B3 reports the comparison of 2000 aggregate vacant land value estimated from sales price with that estimated from assessed value, for each developable vacant land use type, for the subset of developable vacant parcels that have complete data on both sales price and assessed value. Note that the last row of Table B3 corresponds exactly to the first row of Table 18.

The ratios of 2000 aggregate land value estimated from sales price to that estimated from assessed value are higher than 1.0 for all detailed developable vacant land uses, except for "Cemeteries" and "Cropland and improved pasture land"[27]. The ratio is highest for "Poultry Operations" at 8.6, and second highest for "Irrigated Cropland and Improved Pasture Land" at 8.18.

---

[27] SCAG parcel database uses a four-digit number to represent a land use. The first two digits represent the major land use, and the third digit represents a subcategory of the major land use represented by the first two digits, and the fourth digit represents a sub-subcategory of the land use represented by the first three digits.
For example, land use code 21** ("Cropland and improved pasture land") includes two sub-categories: land use code 2110 ("Irrigated cropland and improved pasture land") and land use code 2120 ("Non-Irrigated cropland and improved pasture land"). A parcel with land use code 2100 means that SCAG does not know whether the parcel belongs to subcategory 2110 or subcategory 2120. Similarly, land use code 2000 ("Agricultural") means that SCAG does not which subcategory the parcel belongs to among all 2*** land uses. Land use code 3000 ("Vacant") means that SCAG does not which subcategory the parcel belongs to among all 3*** land uses.

Table B1: Full regression results for *developable* vacant parcels, sales price versus assessed value, Riverside County

	Dependent Variable: ln (sales price) ($)	Dependent Variable: ln (2007 assessed value) ($)
ln (land area)	0.264***	0.629***
Distance to the nearest sub-center	-0.042***	-0.061***
Distance to the CBD	-0.021***	-0.005**
Squared distance to the CBD	3.17E-5***	-6.36E-06
Distance to the nearest freeway or highway	0.053***	-0.017
Distance to the nearest coast	-0.009***	-0.005**
Squared distance to the nearest coast	6.41E-5***	-1.16E-05
city: Banning (base category)		
city: Beaumont	-0.485	-1.079**
city: Blythe	0.623**	0.374
city: Calimesa	0.088	0.301
city: Canyon Lake	-0.977	2.639*
city: Cathedral City	1.289***	0.13
city: Coachella	1.121***	0.136
city: Corona	-0.361	0.215
city: Desert Hot Springs	-0.182	0.305
city: Hemet	0.597**	-0.555**
city: Indian Wells	0.937***	1.516***
city: Indio	0.837***	0.205
city: La Quinta	0.913***	1.070***
city: Lake Elsinore	0.629**	0.388*
city: Menifee	0.224	0.078
city: Moreno Valley	0.635**	-0.348
city: Murrieta	0.132	0.299
city: Norco	-0.462*	-0.121
city: Palm Desert	1.135***	0.822***
city: Palm Springs	1.803***	-0.233
city: Perris	0.27	0.555**
city: Rancho Mirage	0.667**	0.634**
city: Riverside	-0.897***	-0.194
city: San Jacinto	0.009	-0.324
city: Temecula	1.202***	0.053
city: Wildomar	0.09	0.019

city: unincorporated_rv	-0.132	-0.028
current land use =1274 (base category)		
current land use =1275	-0.396	-1.299
current land use =1700	-0.965	0.475
current land use =1810	-0.335	-0.317
current land use =1840	-0.37	0.17
current land use =2000	0.165	-1.041
current land use =2100	-0.95	1.8
current land use =2110	0.333	-1.295
current land use =2120	-0.105	-1.165
current land use =2200	0.035	-1.335
current land use =2300	-0.546	-0.694
current land use =2400	-0.559	-0.372
current land use =2500	-0.218	-1.34
current land use =2600	-0.784	-0.641
current land use =2700	-0.449	-0.606
current land use =3000	-0.791	-0.7
current land use =3100	-1.189	-1.125
current land use =3200	-0.444	-1.324
current land use =3300	-1.502*	-0.547
planned land use =0 (base category)		
planned land use =1100	0.855	0.684
planned land use =1110	0.068	0.068
planned land use =1120	0.246***	-0.1
planned land use =1130	-0.811***	-0.189***
planned land use =1200	-0.156	-0.063
planned land use =1210	-0.166	-0.013
planned land use =1220	-0.066	0.097
planned land use =1230	-0.191**	0.325***
planned land use =1233	-0.513***	-0.577***
planned land use =1240	-1.273***	0.383
planned land use =1250	0.156	-0.108
planned land use =1260	-0.379	-0.215
planned land use =1270	-5.554***	-4.031***
planned land use =1300	-0.099	-0.507***
planned land use =1310	-0.716***	-0.291***
planned land use =1320	0.135	0.231**
planned land use =1340	0.044	0.151
planned land use =1410	0.886	-0.234
planned land use =1420	-2.822**	-3.465***
planned land use =1430	1.602	-5.223***
planned land use =1500	-0.149	0.276***
planned land use =1600	-0.226***	0.144*

planned land use =1800	-0.546**	-1.545***
planned land use =1810	0.54	-2.483***
planned land use =1820	1.260***	-1.096***
planned land use =1830	0.398	-1.302***
planned land use =1850	0.329**	-0.738***
planned land use =1880	-0.742***	-3.619***
planned land use =2000	0.716***	-0.175*
planned land use =4000	0.023	-1.360***
sale year =1980	-1.375*	0.095
sale year =1981	-0.51	0.154
sale year =1982	-1.507***	-0.251
sale year =1983	-1.478***	-0.141
sale year =1988	-2.181***	-0.448
sale year =1989	-0.315	-0.822
sale year =1990	-1.037	0.591
sale year =1995	0.473	-1.674
sale year =1998	-0.759	-1.409*
sale year =1999	-1.14	-0.922
sale year =2000 (base category)		
sale year =2001	2.220***	0.285
sale year =2002	-0.383	-0.254
sale year =2003	0.248	-0.181
sale year =2004	0.810*	-0.085
sale year =2005	1.171**	0.819*
sale year =2006	1.244**	0.923*
sale year =2007	1.193**	0.697
Constant	12.960***	8.415***
R-squared	0.357	0.384
N	27239	27239

* p<0.05, ** p<0.01, *** p<0.001

Table B2: Full regression results for *undevelopable* vacant parcels, sales price versus assessed value, Riverside County

	Dependent Variable: ln (sales price) ($)	Dependent Variable: ln (2007 assessed value) ($)
ln (land area)	0.494***	0.608***
Distance to the nearest sub-center	-0.060***	-0.035***
Distance to the CBD	-0.059***	-0.004*
Squared distance to the CBD	1.72E-4***	-5.99E-4***
Distance to the nearest freeway or highway	-0.053	0.117*
Distance to the nearest coast	0.119***	-0.117***
Squared distance to the nearest coast	-8.05E-4***	8.80E-4***
city: Banning (base category)		
city: Beaumont	1.731***	-5.464***
city: Blythe	1.131	1.249
city: Calimesa	1.890***	-6.092***
city: Canyon Lake	-1.644**	-4.712***
city: Cathedral City	-2.305***	-1.428***
city: Coachella	1.670***	0.116
city: Corona	-2.939***	-5.814***
city: Desert Hot Springs	-1.815***	-6.089***
city: Hemet	-1.463**	-6.432***
city: Indian Wells	-0.295	-0.512
city: Indio	-1.826***	1.779***
city: La Quinta	-1.542**	-5.566***
city: Lake Elsinore	-2.053***	-1.152***
city: Menifee	-3.258***	-6.271***
city: Moreno Valley	-0.951*	-5.981***
city: Murrieta	-0.883***	0.601**
city: Norco	-1.161*	-5.449***
city: Palm Desert	-0.486	-1.056
city: Palm Springs	-1.425***	-5.950***
city: Perris	-0.700***	-0.367*
city: Rancho Mirage	0.756	1.305
city: Riverside	-2.622***	-5.672***
city: San Jacinto	1.403***	-5.998***
city: Temecula	1.406***	-5.025***
city: Wildomar	-2.453***	-6.298***
city: unincorporated_rv	-1.617***	-0.656***

current land use =1800 (base category)		
current land use =1820	1.569*	3.059***
current land use =1821	3.098***	-1.539*
current land use =1830	1.54	5.385***
current land use =1832	-0.824	-5.022***
current land use =1850	1.296	3.161***
current land use =1880	2.005**	-1.879***
current land use =3000	5.539***	0.443
current land use =3100	4.638***	-5.339***
planned land use =0 (base category)		
planned land use =1110	-0.993***	-0.530***
planned land use =1120	-0.678***	0.134
planned land use =1130	-1.280***	-1.012***
planned land use =1200	0.134	0.093
planned land use =1210	-0.805***	-0.007
planned land use =1220	-0.701**	-0.154
planned land use =1230	-1.073***	0.336*
planned land use =1240	2.203	7.853***
planned land use =1250	-0.85	-0.981
planned land use =1260	-0.488	-1.236***
planned land use =1300	-1.680***	-0.735***
planned land use =1310	0.29	-0.193
planned land use =1320	-0.3	-0.172
planned land use =1340	-0.504	-1.588***
planned land use =1410	-0.950**	-0.03
planned land use =1420	1.419*	1.306**
planned land use =1500	-1.502***	0.106
planned land use =1600	-0.174	0.177
planned land use =1800	-1.616***	-1.427***
planned land use =1820	-0.363	-1.471***
planned land use =1830	1.123	-5.023***
planned land use =1850	1.103***	-1.621***
planned land use =1880	-1.859***	-1.548***
planned land use =2000	-0.627***	-1.272***
planned land use =4000	-0.915**	-2.568***
sale year =1980	-2.267	0.737
sale year =1981	-0.429	-0.158
sale year =1982	-1.045	-0.24
sale year =1983	-1.770**	-0.333
sale year =1984	-0.805	0.691
sale year =1985	-1.391	-1.812
sale year =1986	-1.121	-0.658
sale year =1987	-0.605	0.785

sale year =1989	0.152	-0.378
sale year =1990	0.191	-1.524
sale year =1992	1.694	0.254
sale year =1996	0.475	0.762
sale year =1998	-0.473	-1.086
sale year =1999	-0.605	-0.728
sale year =2000 (base category)		
sale year =2001	3.916***	0.327
sale year =2002	-0.132	-0.318
sale year =2003	0.244	-0.204
sale year =2004	0.596	-0.11
sale year =2005	1.254*	0.392
sale year =2006	1.185*	0.592
sale year =2007	0.954	0.337
Constant	9.163***	13.622***
R-squared	0.549	0.513
N	14811	14811

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Land use code	Land use description	Number of parcels	Aggregate area of parcels (10^9 sq. ft)	2000 aggregate land value imputed from sales price ($\$10^9$)	2000 aggregate land value imputed from assessed value ($\$10^9$)	ALV_{SP} / ALV_{AV}
1274	Former Military Base - Built-Up Area	3	1.168	0.004	0.003	1.25
1275	Former Military Vacant Area	5	0.004	0.005	0.005	1.01
1700	Under construction	13883	0.206	7.349	3.054	2.41
1810	Golf courses	366	0.090	0.566	0.131	4.31
1840	Cemeteries	5	0.002	0.004	0.011	0.33
2000	Agricultural	1128	0.297	1.772	0.410	4.32
2100	Cropland and improved pasture land	1	1.43E-05	1.20E-04	2.60E-04	0.46
2110	Irrigated Cropland and Improved Pasture Land	1163	1.340	3.240	0.396	8.18
2120	Non-Irrigated Cropland and Improved Pasture Land	680	0.370	1.704	0.326	5.22
2200	Orchards and Vineyards	897	0.808	1.880	0.346	5.43
2300	Nurseries	83	0.029	0.065	0.030	2.19
2400	Dairy, Intensive Livestock, and Associated Facilities	12	0.012	0.020	0.020	1.00
2500	Poultry Operations	10	0.006	0.025	0.003	8.60
2600	Other Agriculture	149	0.074	0.093	0.048	1.95
2700	Horse Ranches	327	0.104	0.247	0.153	1.62
3000	Vacant	3469	0.280	5.035	1.010	4.99
3100	Vacant Undifferentiated	3875	0.592	2.227	0.969	2.30
3200	Abandoned Orchards and Vineyards	15	0.009	0.030	0.005	5.90
3300	Vacant With Limited Improvements	1168	0.063	0.684	0.147	4.64
Total		27239	4.290	24.951	7.067	3.53

Table B3: The comparison, by detailed land use code, of 2000 aggregate land value estimated from sales price to that estimated from assessed value, for the subset of "developable" vacant parcels in Riverside County that have complete data on both sales price and assessed value.

Table B4 has the same structure as Table B3, and reports on the comparisons for each undevelopable vacant land use type, for the subset of developable vacant parcels that have complete data on both sales price and assessed value. Note that the last row of Table B4 corresponds exactly to the second row of Table 18.

The ratios of 2000 aggregate land value estimated from sales price to that estimated from assessed value are higher than 1.0 for all detailed undevelopable vacant land use types, except for "Regional parks and recreation". The ratio is highest for "Local parks and recreation" at 12.14, second highest for "Vacant" at 8.53.

Land use code	Land use description	Number of parcels	Aggregate area of parcels (10^9 sq. ft)	2000 aggregate land value imputed from sales price ($\$10^9$)	2000 aggregate land value imputed from assessed value ($\$10^9$)	ALV_{SP} / ALV_{AV}
1800	Open space and recreation	7	0.003	0.008	0.007	1.19
1820	Local parks and recreation	31	0.005	0.129	0.011	12.14
1821	Developed local parks and recreation	30	0.007	0.017	0.011	1.61
1830	Regional parks and recreation	1	1.39E-05	1.98E-04	3.99E-04	0.50
1832	Undeveloped regional parks and recreation	71	0.175	0.004	0.002	1.87
1850	Wildlife preserves and sanctuaries	37	0.045	0.161	0.021	7.55
1880	Other open space and recreation	584	0.014	0.231	0.094	2.46
3000	Vacant	2238	0.572	4.972	0.583	8.53
3100	Vacant Undifferentiated	11812	6.898	6.964	1.813	3.84
Total		14811	7.719	12.487	2.542	4.91

Table B4: The comparison, by detailed land use code, of 2000 aggregate land value estimated from sales price to that estimated from assessed value, for the subset of "undevelopable" vacant parcels in Riverside County that have complete data on both sales price and assessed value.

Appendix C
Classification of Vacant Land
into
Developable and Undevelopable

This appendix summarizes Technical Report 2012-8. "Developable Vacant Land in the LA Metro Area -- An Economic Approach", Yuntao Guo and Richard Arnott, which is available on the project website at **la-plan.org**. We should note that the appropriateness of the approach and the accuracy of the results reported in the Technical Report have not yet been approved via the editorial process. Thus, the reader is justified in being skeptical of the accuracy of our decomposition of aggregate vacant land value into aggregate developable vacant land value and aggregate undevelopable land value. It should also be noted, however, that the decomposition should have only minor impact on our estimates of aggregate vacant land values.

Previous papers in the literature that have aimed to identify "undevelopable" land have identified such land based on topographical characteristics such as elevation and steepness of the terrain (e.g. Angel *et al.*, 2012; Saiz, 2010, and perhaps also on access to water. The approach adopted in Arnott and Guo (2012) is fundamentally different, basing the distinction between developable and developable on the value of land per unit area. Thus the approach follows the hypothesis that "the market" knows best what land is developable and what is undevelopable. This hypothesis in turn is based on the well-known efficient markets hypothesis, that, under ideal conditions, markets efficiently aggregate information. The idea is that if someone has private information indicating that an asset is undervalued, he can profit from his private information by buying the asset (or if his private information is that the asset is overvalued, by selling the asset short) but in doing so conveys his private information to the market. Conditions are not of course ideal. Land is a heterogeneous commodity; individual parcels are sold infrequently; and certain landowners hold and exercise market power. Despite these qualifications, in most contexts the price of an asset does reflect its profitability well. A more specific result, due to Arnott and Lewis (1979), is that in steady state, land will be developed when its value per unit area has reached a threshold level that depends on the interest rate, the growth rate of the underlying earnings stream, and the costs of developing land.

The empirical question is then how to determine the threshold land value per unit area such that land is judged developable.

The Arnott-Guo procedure proceeded as follows. The developability status of most vacant parcels was assigned on the basis of a parcel's current land use, as described in Table 1 of the paper. There were only three land use categories, 1275 ("Former Base Vacant Area" in the SCAG land use classification), 3000 ("Vacant" in the SCAG land use classification) and 3100 ("Vacant Undifferentiated" in the SCAG land use classification), whose land uses Arnott and Guo defined as "ambiguously developable." For each county, a sample of ambiguously developable parcels with assessed land value indicated was collected and subjectively characterized as being developable or undevelopable. For this sample, a linear logit regression was run of developability probability against a constant and assessed land value[28] per unit area. A threshold development probability was then calculated iteratively such that Type I and Type II errors were approximately equal (a Type I error occurs when a parcel that was subjectively judged to be undevelopable had assessed value per unit area above the threshold level, and a Type II error occurs when a parcel that was subjectively judged to be undevelopable had assessed value per unit area below the threshold level). Each parcel outside the sample was then imputed a tentative developability probability, based on its assessed value per unit area of land, equal to the fitted value from the logit regression. To reduce misclassification due to errors in individual parcels' assessed values, each parcel was then imputed a developability probability as the average of the tentative developability probabilities of the parcel and its neighboring parcels. Parcels were then classified as developable or undevelopable on the basis of their imputed development probabilities, with the qualification that parcels within a quarter mile of a major intersection were automatically categorized as developable After this procedure was applied, a detailed map was drawn of each county with developed parcels colored in yellow, undevelopable vacant land colored in red, developable vacant land colored in green, water colored in blue, and major roads marked in black. The map of Riverside County corresponded impressively well to Arnott's knowledge of the County. Nevertheless, some anomalies remained. Each anomaly was inspected by satellite imaging, with perhaps supplemental research. Almost all the anomalies had good explanations, corresponding for example to ghost towns and power stations in the desert.

Despite its apparent complexity, the procedure reduces to assigning developability status to a parcel on the basis its assessed value per square foot.

No doubt, the Arnott-Guo procedure could be improved upon through the use of more sophisticated spatial econometric techniques. Furthermore, hybrid procedures that exploit additional information, such as information on topography and hydrology, would result in even more accurate categorization. Nevertheless, the results persuaded us of the soundness of the principle underlying our procedure.

[28] Since the procedure uses assessed value rather than sales price, the hypothesis that "the market" knows best what land is developable and what land is not is replaced with the hypothesis that "the assessor" knows well what land is developable and what land is not.

Appendix D
Land Values
Disaggregation by Land Use Category

This appendix presents various statistics related to land values, disaggregated by major land use category. Of particular interest is the proportion of aggregate land value, as well as the proportion of aggregate land areas, associated with the twelve land use categories. Table D1 reports the numbers for *developed* land, Table D2 the corresponding numbers for all land.

	Land use	Number of Parcels	Aggregate Land Value ($\$10^9$)	Percentage of Total Developed Land Value	Aggregate Land Area (10^9 sq. ft)	Percentage of Total Developed Land Area
1	Single Family Residential	2895886	224.77	63.87%	29.03	10.54%
2	Multi Family Residential	290809	18.18	5.16%	3.32	1.21%
3	Mixed Residential	266383	14.20	4.03%	8.67	3.15%
4	Office	47510	11.27	3.20%	3.19	1.16%
5	Retail	249836	27.39	7.78%	8.19	2.97%
6	Other Commercial	13872	2.89	0.82%	1.48	0.54%
7	Public	88092	8.51	2.42%	185.11	67.22%
8	Warehousing	6450	0.91	0.26%	1.49	0.54%
9	Other Industrial	166780	25.93	7.37%	16.02	5.82%
10	Transportation, Communication, and Utilities	58012	16.54	4.70%	18.52	6.73%
11	Mixed	15252	1.31	0.37%	0.33	0.12%
	Total Developed	4098882	351.92	100.00%	275.36	100.00%

Table D1: Aggregate land value in 2000 of developed parcels by major land use categories.

About 73% of aggregate developed land value is associated with residential land use, 4% office and other commercial, 8% with warehousing and other industrial, 8% with retail, and somewhat over 7% with public and transportation, communication, and utility. In terms of total developed land area, the corresponding percentages are 15%, 2%, 6%, 3% and 75%. That 75% of the developed land area is either public, or transportation communication, and utilities is the result of much of it being classified as developed even though it is largely vacant. Also of interest is the average land value per sq. ft. for land in different uses: single-family residential ($7.74), multi-family residential ($5.48), mixed residential ($1.64), office ($3.53), retail ($3.34), other commercial ($1.95), public ($0.05), warehousing ($0.61), other industrial ($1.62), TCU ($0.89), and mixed ($3.98). These figures confound differences in the value of land at a particular location and different patterns of location for different land uses. Differences in the value of land across different land uses at a particular location, controlling for location and city, can be inferred from the magnitude of the regressions' land use dummy variables.

Table D2 is the same as Table D1 except that it includes vacant land. Vacant land constitutes 23.68% of aggregate land value and 73.33% of all land area. Developable vacant parcels have an average value per sq. ft of $0.65, and undevelopable vacant parcels an average value per sq. ft. of $0.10.

	Land use	Number of Parcels	Aggregate Land Value ($\$10^9$)	Percentage of Total Land Value	Aggregate Land Area (10^9 sq. ft)	Percentage of Total Land Area
1	Single Family Residential	2895886	224.77	48.75%	29.03	2.81%
2	Multi Family Residential	290809	18.18	3.94%	3.32	0.32%
3	Mixed Residential	266383	14.20	3.08%	8.67	0.84%
4	Office	47510	11.27	2.45%	3.19	0.31%
5	Retail	249836	27.39	5.94%	8.19	0.79%
6	Other Commercial	13872	2.89	0.63%	1.48	0.14%
7	Public	88092	8.51	1.85%	185.11	17.93%
8	Warehousing	6450	0.91	0.20%	1.49	0.14%
9	Other Industrial	166780	25.93	5.62%	16.02	1.55%
10	Transportation, Communication, and Utilities	58012	16.54	3.59%	18.52	1.79%
11	Mixed	15252	1.31	0.29%	0.33	0.03%
12	Developable Vacant	224976	37.42	8.11%	57.53	5.57%
13	Undevelopable Vacant	339824	71.77	15.57%	699.49	67.76%
	Total parcels	4663682	461.09	100.00%	1032.37	100.00%

Table D2: Aggregate land value in 2000 of all parcels, including both vacant parcels and developed parcels, by major land use categories.

Appendix E
Ground Truth Tracking/Trekking

Ground truth tracking or ground truth trekking refers to an ensemble of techniques that are employed to check the accuracy of spatial data. One is satellite photography; another is application of Google Street View, at locations where these views are available; yet another is a site visit.

In the larger project of which this essay is one of many products, ground truth tracking was used in a variety of different contexts. The relevant technical reports, which are unpublished and have not been through a refereeing process, are available on the project website **la-plan.org**.

The first context in which ground truth tracking was used was to ascertain the accuracy of parcel boundaries via satellite photography. (Technical Report 2010-5, "Systematic Validation of MRPI Databases -- Census Data (Block Group)", Michael Goodchild and Wenwen Li).

Ground truth tracking was also used was to check on estimates obtained from satellite imaging of the floor-area ratios of properties in Imperial County, for which parcel data on the floor area of structures were completely absent. (Technical Report 2010-7, "DEM-based Floor Area Estimation", Michael Goodchild and Wenwen Li). Since light bounces off structures while radar waves do not, one can in principle measure the average height of structures on a property by subtracting the average elevation of the ground estimated through radar from the average height of the property estimated through satellite photography -- the difference-in-elevation or DEM method. The accuracy of the method was checked through site visits. Unfortunately, the signals were too noisy for their difference to be statistically reliable at the level of the individual parcel. There were also problems with applying the technique to mountainous terrain.

At locations covered by Google Street View, the accuracy of the parcel data with respect to the floor area of structures can be estimated from a combination of Google Street View and satellite imaging[29]. In the first stage, satellite imaging was used to estimate the proportion of the property that is covered by structure. In the second stage, Google Street View is used to estimate to average height of the structures (and also to check on the identification of structures through satellite imaging) on the parcel. Manually applying these techniques entailed about one person-minute of time per parcel. In the third stage, site visits were used to determine the accuracy of estimates of the floor area of structures on a site obtained by multiplying the structure footprint estimated from satellite photography by the average building height obtained from Google Street View. (Technical Report 2011-6. "SCAG Parcel Database Validation Report on Accuracy of

[29] Here and elsewhere, account needs to be taken that the parcel data are for 2007, whereas the ground truth tracking is current.

Total Floor Space per Parcel by Ground Truth Trekking", Michael Goodchild, Wenwen Li, Alex Schild, and Nate Royal).

Satellite photography was used in several ways in the classification of vacant parcels into developable vacant land and undevelopable vacant land. The developability status of most vacant parcels was assigned on the basis of a parcel's current land use code. Parcels with land use code 3000 ("Vacant" in the SCAG land use classification) and 3100 ("Vacant Undifferentiated" in the SCAG land use classification) were classified as "ambiguously developable". For each county a sample of ambiguously developable parcels with assessed value indicated was collected and subjectively categorized as being developable or undevelopable on the basis of satellite images. For this sample, a logit regression was run of developability probability against assessed land value per unit area. A critical assessed land value per unit area was then iteratively calculated such that Type I and Type II errors were approximately equal. Parcels outside the sample were then assigned a tentative developability probability. To deal with data error, each parcel was then assigned a developability probability as the average of the tentative developability probability of the parcel and its neighboring parcels. Parcels were then classified as developable or undevelopable on the basis of their development probabilities. Satellite photography was again used to investigate "anomalies". Almost all anomalies had good explanations, such as ghost towns and power stations in the desert. (Technical Report 2012-8. "Developable Vacant Land in the LA Metro Area -- An Economic Approach", Yuntao Guo and Richard Arnott).

Satellite photography, followed up by site visits, was also used to ascertain the accuracy of SCAG's land use classification. For the Palm Springs model zone, it was found that land use classification is generally accurate[30], though reclassification of recently developed sites appears to be prone to error[31]. (Technical Report 2013-3. "Ground Truck Tracking of Land Use in SCAG's Parcel Database", Ramy Mami and Matthew Taylor).

Ground truth tracking holds considerable promise. The results from ground truth tracking can be statistically combined with parcel data to produce more accurate estimates. Ground truth tracking is however often expensive (in terms of the person-hours applied) and is typically done in an *ad hoc* fashion. We know of no statistical theory on the optimal design of ground truck tracking sample selection so as to achieve the most bang for the buck -- the maximum improvement in estimation/forecasting accuracy per dollar spent. If we had it all to do over again, we would have developed the theory and applied it to optimal sample selection.

---

[30] We cannot resist mentioning a golf course on a mountaintop because of the novelty of the concept.

[31] It is often difficult to determine on the basis of visual inspection whether a site with a house on it is under construction (and therefore classified as vacant, according to our definition) or developed, particularly since the real estate collapse in 2008 caused many developed homes to be abandoned and many homes under construction to be lived in.

Appendix F
Imputing the Land Value of Developed Parcels on the Basis of the Value of Vacant Land Imparts a Downward Bias

Casual experience suggests that, controlling for accessibility, if two parcels of equal accessibility are converted to urban use at different times, the one that is converted earlier is of higher quality. If this is correct, then the procedure we follow of imputing land value to developed properties on the basis of the value of vacant land, controlling for location, imparts a systematic downward bias to our estimates of aggregate land values. This appendix investigates whether casual experience is supported by theory.

The simplest model in which to address this question is the Arnott and Lewis (1979) model of the transition of land to urban use. The model assumes that a landowner/developer chooses when and at what density to convert his parcel of vacant land to urban use under perfect foresight so as to maximize discounted profit. Once developed, a parcel remains at the same density forever. Let R(t,A) be the rent per unit area of floor space at time t on a site has exogenous amenity level A, T be the time of development, K be the capital per unit area of land in developing the property in monetary units, $\mu(K)$ be the floor-area ratio with $\mu' > 0$ and $\mu'' < 0$, r be the interest rate, and V(A) be the market value of the vacant land today. The question of interest is whether $dT/dA < 0$, i.e. whether a site will a higher amenity level is developed earlier. The landowner/developer's maximization problem is

$$\max_{T,K} \int_{T}^{\infty} \mu(K)R(t,A)\, e^{-rt}\, dt - Ke^{-rT} - V(A). \tag{1}$$

The first-order profit-maximization conditions with respect to development timing and density are:

$$T: \quad [- \mu(K)R(T,A) + rK]e^{-rT} = 0 \tag{2}$$

$$K: \quad \int_{T}^{\infty} \mu'(K)R(t,A)\, e^{-r(t - T)}dt - 1]e^{-rT} = 0. \tag{3}$$

The timing condition states that land is developed at a time when the marginal benefit from postponing development, the cost of capital, rK, equals the marginal cost, the rent foregone. The density condition states that land is developed when the marginal benefit from spending an extra unit of capital at development time, the increase in the present value of rents from doing so, equals the unit cost. If the second-order conditions hold as strict inequalities, then $R_T(T,A) > 0$ and $\mu''(K) < 0$.

Total differentiation of the pair of equations with respect to A yields

$$- \mu(K)R_T(T,A)\, dT/dA + [- \mu'(K)R(T,A) + r]\, dK/dA = \mu(K)R_A(T,A) \tag{4}$$

$$[- \mu'(K)R(T,A) + r]\, dT/dA + [\int_{T}^{\infty} \mu''(K)R(t,A)\, e^{-r(t - T)}dt]\, dK/dA = \tag{5}$$

$$- \int_T^{\infty} \mu'(K)R_A(t,A)\, e^{-r(t - T)}dt$$

Thus,

$$[- \mu'(K)R(T,A) + r]\, dK/dA = \mu(K)R_A(T,A) + \mu(K)R_T(T,A)\, dT/dA \qquad (6)$$

Substituting this into the second equation gives

$$[- \mu'(K)R(T,A) + r]^2\, dT/dA = [- \int_T^{\infty} \mu'(K)R_A(t,A)\, e^{-r(t - T)}dt][- \mu'(K)R(T,A) + r]$$
$$- \{[\int_T^{\infty} \mu''(K)R(t,A)\, e^{-r(t - T)}dt](\mu(K)R_A(T,A) + \mu(K)R_T(T,A)\, dT/dA\} \quad (7)$$

or

$$\{[- \mu'(K)R(T,A) + r]^2 + [\int_T^{\infty} \mu''(K)R(t,A)\, e^{-r(t - T)}dt]\mu(K)R_T(T,A)\}dT/dA= \qquad (8)$$
$$[- \int_T^{\infty} \mu'(K)R_A(t,A)\, e^{-r(t - T)}dt][- \mu'(K)R(T,A) + r]$$
$$- [\int_T^{\infty} \mu''(K)R(t,A)\, e^{-r(t - T)}dt]\mu(K)R_A(T,A).$$

Now, the expression in curly brackets preceding dT/dA is negative via the second-order conditions for a maximum. Thus, the sign of dT/dA depends on the sign of the terms on the RHS; specifically

$$\text{sgn}\,(dT/dA) = \text{sgn}\,\{[\int_T^{\infty} \mu'(K)R_A(t,A)\, e^{-r(t - T)}dt][- \mu'(K)R(T,A) + r]$$
$$+ [\int_T^{\infty} \mu''(K)R(t,A)\, e^{-r(t - T)}dt]\mu(K)R_A(T,A)\}. \qquad (9)$$

Using (3), this can be simplified to

$$\text{sgn}\,(dT/dA) = \text{sgn}\,\{[\int_T^{\infty} \mu'(K)R_A(t,A)\, e^{-r(t - T)}dt][- \mu'(K)R(T,A) + r]$$
$$+ [\mu''(K)/\mu'(K)]\mu(K)R_A(T,A)\}. \qquad (10)$$

Thus far, we have placed no restrictions on how an increase in the amenity level affects the time path of rents. To simplify, where A_0 is an arbitrary amenity level, I shall assume that

A-1: $R(t, A) = AR(t, A_0)$ for all A and all t;

that is, a given change in the amenity level results in a proportional change in the rent function over time. Under **A-1**:

$$\text{sgn}\,(dT/dA) = \text{sgn}\,\{[\int_T^{\infty} \mu'(K)R_A(t,A)\, e^{-r(t - T)}dt][- \mu'(K)R(T,A) + r]$$
$$+[\mu''(K)/\mu'(K)]\mu(K)R_A(T,A)\} \qquad (11)$$

$$= \text{sgn}\,\{[\int_T^{\infty} \mu'(K)R(t,A)\, e^{-r(t - T)}dt][- \mu'(K)R(T,A) + r]$$
$$+ [\mu''(K)/\mu'(K)]\mu(K)R(T,A)\}.$$

Using (2) and (3), this reduces to

$$\text{sgn}\,(dT/dA) = \text{sgn}\,\{[-\,\mu'(K)R(T,A) + r] + [\mu''(K)/\mu'(K)]rK\}. \qquad (12)$$

Recall that the elasticity of substitution in intensive form under CRS is

$$\sigma = -\,\mu'(K)(\mu(K) - \mu'(K)K)/[\mu(K)\mu''(K)K]. \qquad (13)$$

Using $\mu(K)R(T,A) = rK$ from (2), $-\,\mu'(K)R(T,A) + r = r(-\,\mu'(K)K/\mu(K) + 1)$. Substituting this into (12) yields

$$\text{sgn}\,(dT/dA) = \text{sgn}\,\{[-\,\mu'(K)K + \mu(K)] + [\mu''(K)K\mu(K)/\mu'(K)]\}$$

$$= \text{sgn}\,\{(-\,\sigma + 1)[\mu(K)\mu''(K)K/\mu'(K)\} < 0 \text{ when } \sigma < 1. \qquad (14)$$

Consider the case where $\sigma = 0$. Then K is essentially constant. Thus, comparing two properties that differ in A, both have the same marginal benefit from postponing development, rK, but the property with the higher amenity level has the higher marginal cost, which is the rent foregone. When $\sigma \in (0,1)$, the landowner/developer responds to a higher amenity level by both bringing forward development and increasing development density. When $\sigma > 1$, the landlord-development will respond by constructing later at much higher density. Only under exceptional rental growth conditions is $\sigma > 1$ consistent with the second-order conditions, and all the empirical evidence indicate that σ is less than one.

Thus, casual experience is supported by theory. Controlling for accessibility, locations with higher amenities are developed earlier. Thus, controlling for accessibility but not for amenities, ascribing land values to parcels that have already been developed on the basis of the market value of vacant land imparts a downward bias to those estimates.