

Estimating Hedonic Models of Consumer Demand with an Application to Urban Sprawl.

Patrick Bajari and Matthew Kahn
University of Minnesota and NBER
University of California, Los Angeles

May 9, 2007

Abstract

In this paper, we describe a method for estimating hedonic models of consumer demand using a methods described in Bajari and Benkard (2005) and Bajari and Kahn (2005). We apply our methods to understanding suburbanization and the associated urban sprawl using a unique data set form Los Angeles in 2000-2003. Urban sprawl has both costs and benefits. One cost of sprawl is that commute times are increased since homes and places of work are more dispersed. A benefit is that sprawl allows consumers to purchase larger homes and lots. This paper uses a new data set on housing transactions in Los Angeles County to compare some costs (increased commuting) and benefits (larger homes) of sprawl. We use new methods in demand estimation to recover how heterogeneous home buyers tradeoff commuting versus larger homes at the margin. Finally, we evaluate the partial equilibrium welfare effects of two anti-sprawl policies.

1 Introduction.

Hedonic regressions are one of the most commonly used techniques in applied microeconomics for the study differentiated product markets. Hedonic regressions date back at least 80 years (see Waugh (1928)) and have been an active research area for several decades, with seminal contributions by Grilliches (1961), Grilliches (1971), Rosen (1974) and Epple (1987). In this paper, we describe a flexible, but computationally simple approach for estimating structural models of consumer demand using hedonics. The framework is an application of Bajari and Benkard (2005) and Bajari and Kahn (2005), which builds on the classic Rosen hedonic two-step (Rosen 1974, Epple 1987). In a first stage estimation, a flexible home price regression is estimated using local linear regression. Second, using the results from the local linear regression, we recover the implicit price faced by each household in our data set and the marginal utility of each household for every product characteristic. This allows us to generate a nonparametric distribution of random coefficients for the various product characteristics in our data set. Third, we regress the random coefficients on consumer demographics in order to learn about the joint distribution of tastes and demographic characteristics.

As an application of our methods, we study the incentives for suburbanization by households in a unique data set of housing purchases from Los Angeles during 2000-2003. In the year 2000, across all metropolitan areas, 54 % of employed heads of households lived in a single detached home and commuted to work by private vehicle. Flight from blight may explain some of the demand for suburbanization (Mieskowski and Mills 1993, Berry-Cullen and Levitt 1999) but suburbanization is ubiquitous across the entire United States (Margo 1992, Glaeser and Kahn 2004). Suburban housing products offer several benefits. The housing stock is newer. The home's interior space and lot size are larger. Suburban communities tend to self select richer, more highly educated households. Thus, on average, these communities offer greater local peer effects. In a monocentric city, a major cost of suburbanization is a longer commute. As employment has sprawled, this bundling of long commutes with new housing has been attenuated. Dispersed employment has shortened suburbanites' commute times (Gordon, Kumar and Richardson 1991).

In this paper, we use Los Angeles county home transaction data from 2000 to 2003 to estimate how heterogeneous home buyers value the various attributes of urban versus suburban housing. In particular, we estimate the valuation for square footage, lot size, community characteristics and commute times to employment centers. Los Angeles is the right place to study the benefits and costs of suburban sprawl. For households with incomes above \$53,000, 63% of employed Los Angeles heads of households live in single detached homes and commute to work by private vehicle.¹ This is a metropolitan area with dispersed employment centers (Giuliano and Small 1991). Only 6.5% of Los Angeles

¹ This fact is generated using the 1% IPUMS data from the 2000 Census of Population and Housing. The income cutoff of \$53,000 represents the median household income for the set of Los Angeles households where the head of household works. For all Los Angeles heads of households who work, 47% live in detached housing and commute by private vehicle.

workers commute by public transit.

Our housing demand model yields estimates of the willingness to pay for the physical characteristics of the home such as its age, the structure's square feet and the lot size. In addition, we estimate willingness to pay for community attributes such as access to high human capital neighbors and the demographic composition of the neighborhood. We also provide new estimates of the willingness to pay to avoid commuting.

After measuring the private benefits of sprawl, we then use our estimates to investigate the welfare effects of two anti-sprawl policies. There is an ongoing concern that there is excess urban expansion in major U.S cities (for an overview see Nechyba and Walsh 2004). "Cities, it is claimed, take up too much space, encroaching excessively on agricultural land. ... Excessive urban expansion also means overly long commutes, which generate traffic congestion while contributing to air pollution. Unfettered suburban growth is also thought to reduce the incentive for redevelopment of land closer to city centers, contributing to the decay of the downtown areas. Finally, by spreading people out, low-density suburban development may reduce social interaction, weakening the bonds that underpin a healthy society (Brueckner 2000)."

Using our housing demand model, we focus on the commute time consequences of low density living. In our first experiment, we shrink the lot size and square footage of all suburban homes. We then ask for each household, holding fixed each household's place of work, is the lower utility from smaller homes compensated for by the shorter commute? We also examine the incidence of this policy by demographic group to determine who are the winners and losers from such a policy. In our second policy experiment, we "monocentricize" Los Angeles by sending all employment back to the Central Business District. In this case suburban commutes become much longer. We estimate how much suburban home buyers would lose and how much urban home buyers would gain from such an employment sprawl reversal. Employment sprawl unbundles the commute time versus land consumption tradeoff. Our welfare analysis builds on recent work that has evaluated the intended and unintended consequences of different housing policies (see Arnott 1995, Glaeser and Luttmer 2003, Peng and Wheaton 1994, Phillips and Goodstein 2000, Thorsnes 2000).

2 The Data.

The data source is First American Real Estate Solutions. First American's MetroScan houses a comprehensive database of residential, commercial, industrial and vacant property obtained from county assessors and other agencies. The data in MetroScan initially comes from the county assessor's office. We focus solely on single detached homes sold in Los Angeles county over the years 2000 to 2003. For each home, we observe its sales price, year built, unit square footage, and lot square footage. This specific information on the unit's physical size and its lot size are crucial inputs

for measuring the demand for private space. Other data sources such as the Census of Population and Housing do not provide this information and the American Housing Survey surveys only a small number of homeowners in any metropolitan area.

Our data set provides each unit's street address. Thus, we know the unit's zip code, census tract and census bloc. We use these geographic identifiers to merge on data from the 2000 Census of Population and Housing. In particular, we merge on data regarding the socio-economic composition of the home owner's census tract and census block (namely the geographical area's educational attainment, racial and ethnic composition) and information on each census tract's population density and average one way commuting time in minutes. The 2000 Census data provides some tract level means by owner status. For example, we can construct the average household size and median household income for home owners in each census tract. We will use this information below when we discuss aggregation of consumer preferences.

The First American data provides us with the last name of each home purchaser. Based on this information, we assign each person a dummy variable that equals one if the last name is Hispanic and a second dummy variable that equals one if the last name is Asian.² In an immigrant city such as Los Angeles where Hispanics and Asians represent a large share of the population, it is important to recover housing preferences for such subgroups.

The data set includes over 173,000 transactions with a large enough number of transactions within zip codes to allow for within zip code hedonic estimation. We drop from the First American sample those observations that report a sales price of zero. We also trim the left and right tails of the pricing distribution dropping the bottom and top 1 percent of the pricing sample.

To appreciate the strengths of this data set it is useful to contrast it with more conventional micro data from the 2000 Census of Population and Housing. Home price data from the 2000 Census of Population and Housing is self reported and is top coded at \$875,000. In addition, home prices are partitioned into only 23 mutually exclusive categories. The last six categories measured in \$1,000s are 225, 275, 350, 450, 625, and 875. Clearly such crude categories could miss significant amounts of the "action" between communities. Census data only provides scanty information on the structure's quality such as the number of rooms and the unit's year built. Finally, the geography in the public use micro data is much more geographically aggregated. In the 5% sample of the 2000 IPUMS, PUMAs are identified. These are geographical units of roughly 100,000 people. In this paper, we seek to measure preferences for living in much smaller communities at the census bloc or census tract level.

While our data set has a number of advantages over traditional Census micro data for studying the demand for

² While we acknowledge that our assignment of people to ethnic categories based on last names may lead to some miscoding, this piece of household level information is useful for recovering preferences for specific subgroups of the population.

sprawled housing products, we must acknowledge the data set's limitations. As discussed above, we have very limited information on the demographics of each home buyer. In addition, unlike the Census, we do not have information on renters. Thus, the distributions of preferences we recover are for the select sample of households who chose to own and have purchased a home during the years 2000 to 2003 in Los Angeles county. In pooling repeat cross-sections over the year 2000 through 2003, we are assuming that the distribution of preferences is stable over this time period. A feature of the First American data base is that the data set only includes information on the most recent transaction. Thus, if a home sold in June 2000 and January 2002, we would not see the June 2000 transaction in our data set. To avoid attrition bias, we chose to include in our sample only homes that had transacted in the previous three years. We are confident for this short window that multiple sales are unlikely to be an important issue.

In Table 1, we report our sample's summary statistics. Home prices, structure quality and neighborhood attributes vary greatly across the county. Both average home prices and average commute times are much higher than the national average. The average sales price in our sample is \$329,000. Relative to new Hispanic owners, the average new Asian buyer is living at much lower density and has a slightly lower average commute time to work. Hispanics are choosing communities where 54 percent of the community is Hispanic. Asians are more likely to choose communities with more college graduates.

3 The Model.

In this section, we build a model of housing demand for purchasers of single detached housing in Los Angeles county during the years 2000 to 2003. A home $j = 1, \dots, J$ is modeled as a bundle of four types of characteristics. First are the structural characteristics of the home which will include attributes such as the structure's square footage, lot size and year built. The second set of characteristics are community attributes. In our application, this will include the average demographics of persons in your census block, such as the fraction of college educated households and the racial composition of households. Third, is the commuting time from home to work. Finally, there will be some characteristics of the home that are observed by households, but not the economist.

Households solve the following static utility maximization problem:

$$u_{ij} = \max_j u_i(x_j, t_{ij}, \xi_j, c) \text{ Subject to: } p_j + c \leq y_i \quad (1)$$

$$p_j = \mathbf{p}(x, \xi) \quad (2)$$

In equation (1), the term x_j represents the physical and the community attributes of home j , t_{ij} is the commute time

for the head of household i for living in home j , keeping i 's place of work fixed. The variable ξ_j is a product attribute observed by the consumers but not the economist.³ Prices are determined in equilibrium by the interaction of buyers and sellers. The function \mathbf{p} maps the characteristics (x, ξ) into their equilibrium prices. Note that households take prices as exogenous which is a plausible assumption if the housing market in Los Angeles is competitive. We are also assuming that recent home purchasers, who are migrants, believe that community attributes such as the percent of the community who are college graduates and the racial composition of the community are exogenous. We are only examining the demand for the migrant owners who recently bought into the community. Relative to the stock of owners and renters, this group is likely to be small and it reasonable to assume that they take the community attributes as exogenous.

Suppose that characteristic k is continuous and that j^* is household i 's optimal housing unit. Then it must be the case that:

$$\frac{\partial u_i(x_{j^*}, t_{ij^*}, \xi_{j^*}, y_i - p_{j^*})}{\partial x_{j,k}} - \frac{\partial u_i(x_{j^*}, t_{ij^*}, \xi_{j^*}, y_i - p_{j^*})}{\partial c} \frac{\partial \mathbf{p}(x_{j^*}, \xi_{j^*})}{\partial x_{j,k}} = 0 \quad (3)$$

$$\frac{\frac{\partial u_i(x_{j^*}, t_{ij^*}, \xi_{j^*}, y_i - p_{j^*})}{\partial x_{j,k}}}{\frac{\partial u_i(x_{j^*}, t_{ij^*}, \xi_{j^*}, y_i - p_{j^*})}{\partial c}} = \frac{\partial \mathbf{p}(x_{j^*}, \xi_{j^*})}{\partial x_{j,k}} \quad (4)$$

At the optimal bundle of product characteristics, equation (4) must hold. This implies that the marginal rate of substitution between product characteristic k and the composite commodity must be equal to the implicit price.

In the model above, we have assumed that consumers are static utility maximizers. However, equation (4) is generated by dynamic models of housing demand, such as Doherty and Van Order (1984). In these models, household's period utility depends on consumption of housing services and a composite commodity. At any time period, households can invest in housing stock, buy bonds or purchase the composite commodity. This first order conditions for this model imply an equation similar to (4), except that u corresponds to the period utility and \mathbf{p} corresponds to the user cost. We note that this first order condition only holds under stylized assumptions including no adjustment costs, time separable preferences and a competitive housing market. In our application, we will only apply the first order condition (4) to households that have recently moved. We recognize that this is a select sample of Los Angeles county residents. However, as we discussed above, we believe that this drawback is at least partially compensated for a superior data set on home prices than is available in the Census.

We will denote the characteristics used in application as follows:

³ Omitting ξ_j from the demand system will generate biased estimates of the willingness to pay for product attributes (see Berry, Levinsohn and Pakes (1995), Nevo (2001), Petrin (2001) and Bayer, McMillan and Rubin (2004)).

- $SQFT_j$ - The size of the home measured in square feet.
- $LOTSQFT_j$ - The size of the lot that the home is located on measured in square feet.
- AGE_j - The age of home j .
- $PRICE_j$ - The sale price of the home as recorded by First American.
- $RPRICE_j$ - The owner's equivalent rent for the home defined as 0.075 times the sale price deflated to a 2000 base year.
- $MBLACK_j$ - The percentage of people in the home's census block who are black.
- $MHISP_j$ - The percentage of people in the home's census block who are Hispanic.
- MBA_j - The percentage of people over the age of 25 in the home's census block who are college graduates.
- $MINC_j$ - The median income of households in the home's census block.
- ZIP_j - The zip code where the home is located.
- $COMMUTE_j$ - The one way average commute time measured in minutes for workers who live in the home's census tract.

From a single cross section, it is obviously not possible to recover a household's utility function globally. Following the literature on random coefficient discrete choice models, the utility specification we take to the data will be:

$$u_{ij} = \beta_{i,1} \log(SQFT_j) + \beta_{i,2} \log(LOTSQFT_j) + \beta_{i,3} \log(AGE_j) + \beta_{i,4} \log(MBLACK_j) + \beta_{i,5} \log(MHISP_j) + \beta_{i,6} \log(MBA_j) + \beta_{i,7} \log(COMMUTE_{i,j}) + \beta_{i,8} \log(\xi_j) + c \quad (5)$$

In equation (5), utility is a log-linear function of the product characteristics. The log specification allows product characteristics to have diminishing marginal utility. The terms $\beta_{i,k}$, $k = 1, \dots, 8$ are referred to as random coefficients. These terms allow the marginal valuation of the (log) characteristics to be person specific, since the terms, $\beta_{i,k}$ are person specific. In commonly used models, such as the logit or multinomial probit, the economist assumes that $\beta_{i,k} = \beta_k$ for all i . Therefore, our specification allows for a considerably richer specification of heterogeneity in tastes.

In most previous differentiated product studies, the $\beta_{i,k}$ are assumed to arise from a parametric distribution. Most commonly, they are assumed to be independently and normally distributed (see Berry, Levinsohn and Pakes (1995), Petrin (2001) and Nevo (2001)). In the context of our application, this is obviously a strong assumption. People with a higher valuation for big homes (i.e. high $\beta_{i,1}$) might be expected to value living in homes with larger lots (i.e. high $\beta_{i,2}$) and a higher proportion of college educated neighbors (i.e. higher $\beta_{i,6}$). This would not be consistent with the independence assumption. Furthermore, it is not clear, a priori why the random coefficients should have a normal

distribution. Therefore, in our application, we will not impose any parametric distribution on $\beta_i = (\beta_{i,1}, \dots, \beta_{i,8})$ and we will estimate the distribution of new home buyers' tastes nonparametrically.

It is worth noting that in (5), a random probit or logit error term is not included. As we shall discuss later, the model that we propose is just identified and therefore exhausts all of the degrees of freedom available in the data. Therefore, we can perfectly rationalize the observed data without using a random preference shock. We view this as a desirable feature of our model since when there are many choices available to consumers, random preference shocks may generate pathologies in the measurement of consumer welfare.

In our application, we will be interested in how the demographic characteristics of households in high sprawl areas differ from low sprawl areas. Therefore, we model the joint distribution of the random taste coefficients, β_i , and demographics. The demographic characteristics of the household we will consider are:

- $SIZE_i$ – The number of people in the household.
- INC_i - The household's annual income.
- $ASIAN_i$ – An indicator for whether the head of household is Asian.
- $HISP_i$ - An indicator for whether the head of household is Hispanic.

For product characteristics, $k = 1, \dots, 8$ we will append to the model (5) an additional equation of the form:

$$\beta_{i,k} = \alpha_{0,k} + \alpha_{1,k}SIZE_i + \alpha_{2,k}INC_i + \alpha_{3,k}ASIAN_i + \alpha_{4,k}HISP_i + \eta_{i,k} \quad (6)$$

$$E(\eta_{i,k} | SIZE_i, INC_i, ASIAN_i, HISP_i) = 0.$$

In equation (6), the random coefficient for product characteristic k is a function of household i 's demographics and an idiosyncratic household level preference shock, $\eta_{i,k}$. Appending an equation such as (6) to a random coefficient model is common practice in the literature. This equation allows us to learn about the distribution of tastes conditional on demographics. Also, as we shall demonstrate in our identification section this will allow us to learn about preferences under a more appealing set of assumptions than by just imposing equation (5) alone.

In equation (6), the relationship between tastes and demographics is assumed to be linear. In general, with microdata on household level characteristics, this assumption will not be required. However, in our application, we cannot match household level demographics to the transactions. We only have access to demographic characteristics aggregated at the level of the Census tract. As a result, the estimation approach that we propose will require us to aggregate in order to estimate $\alpha_k = (\alpha_{0,k}, \dots, \alpha_{4,k})$. Aggregation, in turn, requires an assumption of linearity.

It is worthwhile to contrast our estimation model with the classic hedonic two-step (see Cheshire 1998, Brown and Rosen 1982, Rosen 1974, Epple 1987, Eskelend, Heckman and Nesheim 2004). The framework that we apply

is derived from Bajari and Benkard (2005) and has the following features. First, we allow for consumers to be heterogeneous in their willingness to pay for product attributes. In a linear hedonic regression, the implicit price for a product attribute is commonly interpreted as the marginal willingness to pay and does not differ across consumers. We consider a nonparametric framework that allows the marginal willingness to pay to freely differ across consumers. Second, we derive consumer preferences for omitted product characteristics that are observed by consumers but not by the economist. Finally, we discuss a framework that allows the economist to recover preferences in a hedonic model with discrete attributes. The standard analysis of the hedonic model exploits a first order condition to uncover structural willingness to pay. We apply here a nonparametric hedonic model that allows for all three of the properties discussed above.

4 Estimation.

Our approach to estimation involves three steps. In the first step, we estimate the housing hedonic price function \mathbf{p} using flexible, non-parametric methods based on the techniques described in Fan and Gijbels (1996) and applied in Bajari and Kahn (2005). Second, we recover a vector of random coefficients for each household by applying first order conditions for optimality. Finally, we recover the joint distribution of demographics and household demographics. We only have access to demographics aggregated at the level of Census tracts. Therefore we propose techniques to estimate household level preferences with this aggregated data. The first two steps of our estimator are similar to those used in Bajari and Benkard (2005) and Bajari and Kahn (2005). The last step is novel to this paper.

4.1 First Step: Estimating the Hedonic Price Functional

In order to estimate the hedonic flexibly, we use methods from local linear methods discussed in Fan and Gijbels (1996). Fix a particular home j^* . In a local linear model, we assume that locally the hedonic price function \mathbf{p} from equation (1) satisfies:

$$p_j = \alpha_{0,j^*} + \sum_k \alpha_{k,j^*} (x_{j,k} - x_{j^*,k}) + \xi_j \quad (7)$$

In equation (7), we assume that in a neighborhood of (x_{j^*}, ξ_{j^*}) the hedonic is approximately linear. However, unlike a linear regression, where the relationship between the dependent and independent variables is globally linear, here the relationship is only locally linear. Thus, the coefficients have a subscript α_{\cdot,j^*} to emphasize that they will be specific to a particular bundle of characteristics (x_{j^*}, ξ_{j^*}) .

Following Fan and Gijbels (1996), for any j^* , $1 \leq j^* \leq J$, we use weighted least squares to estimate α_{j^*}

$$\alpha_{j^*} = \arg \min_{\alpha} (\vec{\mathbf{p}} - \mathbf{X}\alpha)' \mathbf{W}(\vec{\mathbf{p}} - \mathbf{X}\alpha) \quad (8)$$

$$\mathbf{p} = [RPRICE_j], \mathbf{X} = [x_j], \mathbf{W} = \text{diag}\{K_h(x_j - x_{j^*})\} \quad (9)$$

In equations (8) and (9), $\vec{\mathbf{p}}$ is the vector of the owner's equivalent rent for all products $j = 1, \dots, J$ in our cross section of homes, \mathbf{X} is a vector of regressors, which correspond to the observed product characteristics and \mathbf{W} is a matrix of kernel weights.

Note that the kernel weights \mathbf{W} are a function of the distance between product j^* and product j . Thus, the local linear regression assigns greater importance to observations near j^* . Local linear methods have the same asymptotic variance and a lower asymptotic bias than the Nadaraya-Watson estimator, whereas the Gasser-Mueller estimator has the same asymptotic bias and a higher asymptotic variance than local linear methods. In our estimates, normal kernels with a bandwidth set equal to 1.5 times the sample standard deviation were used to construct the weights.

Our estimates of equation (8) and (9) allow us to recover an estimate of the unobserved product characteristic

$$\xi_{j^*} = p_{j^*} - x_{j^*}\alpha_{j^*} \quad (10)$$

In equation (10), the unobserved product characteristic ξ_{j^*} is estimated as the residual to our hedonic regression. While there are certainly other interpretations of the residual in hedonics (e.g. measurement error in price), we believe that this interpretation is the most important in our data.

We use a reasonably large number of co-variates in many applications. Therefore we should not interpret these estimates as “nonparametric”. However, compared to other flexible functional forms, in our experience and judging from our previous experience in Bajari and Kahn (2005), local linear methods appear to give much more plausible estimates of the implicit prices for housing product characteristics.

Equations (8) and (9) use the standard hedonic assumption that unobserved product characteristics are mean independent of observed product characteristics. This is a strong assumption that would be objectionable in practice. For instance, we do not have school quality data available by school district. To control for local school quality, we include zip code fixed effects. We believe that a community's racial composition and percent of adults who are college graduates will be highly correlated with school quality. Bayer, McMillan and Rueben's (2004) study of San Francisco's housing market in 1990 provides direct evidence that controlling for community socio-economics, objective data on school quality adds little information.

In a first step, we run a linear regression of owner equivalent rent on the single detached home's square footage,

age and lot size and the census block’s percent college graduate, percent black and percent hispanic and zip code fixed effects. Zip code fixed effects absorb a number of important attributes such as distance from major employment centers, the beach, climate, air pollution, local property taxes, crime and local high school quality. We then subtract the zip code fixed effects from the owner’s equivalent rent. We then estimate the local linear regressions described in the previous section. This allows us to identify the implicit prices of the community characteristics using census block variation within a zip code. The average zip code has 14 census blocks within it. This emphasis of identifying willingness to pay for community attributes using within zip code community variation distinguishes our approach from previous hedonic studies such as DiPasquale and Kahn (1999). We use estimates from this linear hedonic regression to compute the implicit price of commuting.⁴ This method relies heavily on a linearity assumption, which is less general than the nonparametric approach that we discussed above. Census data provides average commute times for each census tract. In using these data, we are implicitly assuming that the average place of work for new home buyers in a given census tract is the same as the average place of work for all tract residents (i.e renters and owners, long time stayers and recent migrants). Since almost everyone drives in Los Angeles, we do not have to worry about transit mode differences generating differences in commute times. If a census tract’s average commute time is a noisy measure of the average home buyer’s true commute time, then this linear hedonic will under-estimate the marginal value of time.

4.2 Second Step: Recovering the Random Coefficients.

Suppose that household i chooses home j^* . Let $\hat{\alpha}_{j^*,k}$ be the coefficients of the local linear regression associated with x_{j^*} . These coefficients can be interpreted as the implicit prices faced by household i in the market. That is, we estimate the implicit price for characteristic k faced by household i , $\frac{\partial \mathbf{p}(x_{j^*})}{\partial x_{j^*,k}}$, as $\hat{\alpha}_{j^*,k}$. Given an estimate of the implicit prices, we can generate an estimate of $\hat{\beta}_{i,k}$ of household i ’s random coefficient for characteristic k as follows by applying equations (4) and (5):

$$\hat{\beta}_{i,k} = \hat{\alpha}_{j^*,k} x_{j^*,k}. \quad (11)$$

By applying equation (11) for every household in our data set, we nonparametrically estimate the population distribution of random coefficients. The joint cdf $F(\beta_i)$ can be estimated as using its empirical analogue:

$$\hat{F}(\bar{\beta}_i) = \frac{1}{I} \sum 1 \left\{ \hat{\beta}_{i,k} < \bar{\beta}_{i,k} \text{ for all } k \right\} \quad (12)$$

In equation (12), function $1\{\}$ denotes the indicator function.

⁴ When estimating this regression, we also include zip code fixed effects to control for unobserved community features at the level of the zip code.

In the cases where the characteristic k is discrete, we could follow Bajari and Benkard (2005) and Bajari and Kahn (2005) who propose simple a estimator for the preference parameters. In our application, all of the product characteristics are continuous and therefore this procedure is not required.

4.3 Third Step: Aggregation of Preferences.

We would like to describe how new home buyers' willingness to pay for housing attributes varies as a function of buyer demographics. We have individual level data on buyers' willingness to pay for each product attribute but we only have aggregate owner demographics by census tract for the set of all owners who live in that tract. This section proposes an aggregation strategy for matching our "micro" data to readily available "macro" data.

The idea behind our approach to aggregation is straightforward. Naively, we would want to aggregate equation (6) within a single Census tract. From step two above, we have the value of the dependent variable for each household and we can calculate tract specific means, for home owners, for the demographic characteristics in equation (6). We are concerned that a OLS estimate of census tract average marginal willingness to pay for a housing attribute regressed on census tract demographics will not yield consistent estimates for how marginal willingness to pay for attributes varies by demographic groups.

To recover these preferences, a slightly modified approach is required. We partition the $i = 1, \dots, I$ households into G groups each of size $n = \frac{G}{I}$. We perform this partition at random. We could do this for instance by sampling from the households without replacement and assign the first n households to the first group, the second n to the second group and so forth. Then, for each $g \in G$, form an aggregated version of (6) as follows:

$$\frac{1}{n} \sum_{i \in g} \widehat{\beta}_{i,k} = \frac{1}{n} \sum_{i \in g} \{ \alpha_{0,k} + \alpha_{1,k} SIZE_i + \alpha_{2,k} INC_i + \alpha_{3,k} HISP_i + \alpha_{4,k} ASIAN_i + \eta_{i,k} \} \quad (13)$$

$$\overline{\beta}_{g,k} = \alpha_{0,k} + \alpha_{1,k} \overline{SIZE}_g + \alpha_{2,k} \overline{INC}_g + \alpha_{3,k} \overline{HISP}_g + \alpha_{4,k} \overline{ASIAN}_g + \overline{\eta}_{g,k} \quad (14)$$

In equation (14), we define $\overline{\beta}_{g,k} = \sum_{i \in G} \widehat{\beta}_{i,k}$, $\overline{SIZE}_g = \frac{1}{n} \sum_{i \in G} SIZE_i$ and so forth. Because of the nature of our data, we do not observe these values directly. However, we can estimate them as follows. In our census tract data, we observe $t(i)$ the tract in which every household i lives and the average values of the demographic variables for home owners in tract t . Let $SIZE_t$, INC_t , $HISP_t$ and $ASIAN_t$ denote these values. Then

$$\overline{SIZE}_g = \lim_{t, n \rightarrow \infty} \frac{1}{n} \sum_{i \in g} \sum_t SIZE_t \cdot 1 \{t(i) = t\} \quad (15)$$

In equation (15), we estimate the average demographics of home buyers in group g by replacing it with the average of the tract that household i lives in. As the number of tracts and the size of the groups n become sufficiently large, we can consistently estimate \overline{SIZE}_g . Let $\widehat{SIZE}(g)$ be defined as $\widehat{SIZE}(g) = \sum_{i \in g} \sum_t SIZE_t \cdot 1\{t(i) = t\}$ and define $\widehat{INC}(g)$, $\widehat{HISP}(g)$ and $\widehat{ASIAN}(g)$ similarly. The regression equation we will estimate is:

$$\begin{aligned} \bar{\beta}_{g,k} &= \alpha_{0,k} + \alpha_{1,k} \widehat{SIZE}(g) + \alpha_{2,k} \widehat{INC}(g) + \alpha_{3,k} \widehat{HISP}(g) + \alpha_{4,k} \widehat{ASIAN}(g) + \bar{\eta}_{g,k} \\ \text{where } g &= 1, \dots, G \end{aligned} \quad (16)$$

Since we have constructed our groups at random, the expected value of $\bar{\eta}_{g,k}$ will be zero conditional on our covariates. We can therefore estimate (16) by regression. The coefficients α will be biased, but the bias will be small if the number of members n in each group g is sufficiently large.⁵

5 Results

We begin by comparing average housing attributes in high sprawl and low sprawl areas. For all census tracts in our First American Data Set, we partition them into those whose population density is less than or equal to the 25th percentile (the high sprawl tracts) and those census tracts whose density place them in the 75th or higher percentile (the low sprawl tracts). For these two separate groups of tracts, we use our First American Data to calculate means of housing structure and community attributes. These are reported in Table 2. The average home in the sprawl tracts is newer having been built 38 years ago while the average home in the low sprawl tracts was built 65 years ago. Relative to low sprawl homes, homes in the sprawl areas have higher square footage and larger lots. In terms of community attributes, the sprawl homes have much lower rates of blacks and Hispanics and much higher rates of college graduates. In the high sprawl areas, 35% of neighbors are college graduates while in the low sprawl areas 13% are college graduates. Ownership rates in high sprawl areas are much higher.

The one surprise that emerges in this comparison of high sprawl and low sprawl areas is the very small difference in average commuting times. Los Angeles is clearly not a monocentric city. Workers who live in the sprawl areas commute on average 32.5 minutes one way while workers in the low sprawl areas commute on average 31.3 minutes one way. In a monocentric city, this differential would be much larger. As pointed out by Gordon, Kumar and Richardson (1991), and Glaeser and Kahn (2001), suburbanized employment allows suburban workers to reverse

⁵ Due to collinearity issues we are not able to include a large number of explanatory variables in these second stage regressions. For example, we find that a census tract's percent Hispanic is highly negatively correlated with census tract average household income.

commute and drive at faster speeds at lower density.⁶

5.1 Hedonic Pricing Estimates.

The first step in implementing our three step estimation procedure is to estimate the hedonic pricing function. In Table 3, we display the distribution of implicit prices for the various product characteristics. Since we are using a nonparametric regression technique, each household faces a distribution of implicit prices. The average implicit prices all have the signs and magnitudes that are consistent with economic intuition. An extra square foot of interior space is priced at \$9.08 per year whereas an extra foot of lot size is priced at \$.16 per year. Of the community characteristics, we note that the percentage of one's neighbors that are college graduates is heavily capitalized into the prices. Increasing the fraction of college educated neighbors by 10 percentage points will, on average, cost \$1619.9 per year. The racial characteristics of neighbors have smaller implicit prices. For instance, increasing the proportion of hispanics in your block by 10 percentage points reduces real estate price by \$19 per year. Increasing the percentage of blacks in one's census block by 10 percentage points reduces annual home prices by \$1466.

Table 4 reports two hedonic regressions where we include each census tract's average commute time as an explanatory variable. The standard errors are clustered by census tract. We find that the implicit price of an additional minute in the one way commute is -218.00. Suppose that the head of household in this sample works 5 days per year for 48 weeks per year. Every additional minute of a one way commute leads to an additional 8 hours in the car (480 minutes=5 days×48 weeks×2). This implies an opportunity cost of time of \$27.30 per hour. The average household income for owners in the census tracts where the home sales we observe is \$56,300.00 per year. This leads to an average (pre-tax wage) of \$29.32 if the household works 48 weeks, 5 days per week and 8 hours per day.⁷ At the margin, this is consistent with the intuitive argument that the value of commuting is (approximately) equated to the household's hourly wage.⁸

We have documented that the Los Angeles housing stock is highly heterogeneous (see Tables 1 and 2) and that the hedonic pricing function of housing attributes is highly non-linear (see Table 3). Tables 1 through 4 provide an explanation for the wasteful commuting puzzle (see Hamilton 1982, Cropper and Gordon 1991, and Small and Song 1992). While households dislike commuting (see Table 4), each household recognizes that they face a bundling problem. Their dream house may not be near their place of work. While all else equal, a commuter would want a shorter commute, all else is not equal. The full set of housing products is not available within each community. Facing

⁶ The third to last row of Table 2 reports a commuting counterfactual that we will discuss in detail below.

⁷ This is the weighted average of census tract annual average household income for owners where the weights depend on what share of the home sales in our First American data base are from that census tract.

⁸ The transportation economics literature has tended to find smaller opportunity cost estimates for the value of time averaging around 50 percent of one's hourly wage (see Small 1992).

this spanning constraint, rational households tradeoff extra commuting time in return for a preferred housing structure and a preferred community. There would only be a wasteful commuting puzzle if all homes are identical. In this case, utility maximizers become commute minimizers who should Tiebout sort by place of work.

5.2 Preferences Estimates.

For each home buyer, we use our estimates of the implicit price he faces combined with our assumed functional form for the utility function, and his optimal consumption of this attribute, to recover his marginal valuation for various product characteristics. In order to present our estimates, it is most useful to change the units from the random coefficients in (5) to willingness to pay for a ten percent increase in consumption of the characteristic. This can be quickly computed for each household in our sample given our choice of the utility function. Suppose that household i 's current consumption of square footage is $SQFT_j$. Then willingness to pay for an extra square foot can be computed as:

$$WTPSQFT_i = \beta_{i,1} (\log(1.1 \cdot SQFT_j) - \log(SQFT_j)) \quad (17)$$

$$= \beta_{i,1} \log(1.1) \quad (18)$$

In Table 5, we present summary statistics for the empirical distribution of willingness to pay for housing attributes for all recent home buyers and for Hispanic recent home buyers. Recent home buyers care much more about structure size than lot size. The average recent home buyer is willing to pay \$1,430 for a 10 percent increase in structure size and only \$119 for a 10 percent increase in lot size. Hispanic owners have a lower demand for private space. Hispanics are willing to pay more for newer housing. In terms of community attributes, recent home buyers are willing to pay \$400 per year for a 10 percent increase in community residents who are college graduates. Hispanics reveal a positive but lesser demand for access to such peers. Community racial composition preferences appear small. All new home buyers are willing to pay \$88 per year to live in a community with 10 percent fewer black residents. Perhaps surprisingly, new Hispanic home buyers do not have a greater willingness to pay to live in a Hispanic community than the average home buyer.

In Table 6, we study the correlation between the willingness to pay for product characteristics. In most models with random coefficients, the correlation in tastes between different product characteristics is ignored. We find that the independence assumption commonly made in the literature appears to be strongly rejected by the data. Most of the signs are intuitively plausible with the exception of some of the race or age variables. We encourage the reader to

note, however, that Table 5 suggest that the willingness to pay for many of these variables is not particularly large. As shown in Table Six, unit square footage and community percent college graduates are complements. We are surprised that new housing and larger interior space are not complements.

Using the aggregation approach presented above, we now test hypotheses concerning how willingness to pay for housing attributes varies as a function of home owner demographics. Table 7 reports six OLS regressions examining willingness to pay for housing attributes and the consumption of housing attributes. We do not use the full set of demographic characteristics in equation (16) because of collinearity. In columns (1), (3), and (5), we report willingness to pay for interior unit square footage, the home's lot size, and the age of the housing unit as a function of a household head's age, household size, and median income. Richer people greatly value indoor space. An extra \$10,000 in income increases the willingness to pay for a 10 percent increase in indoor space by \$240. The income effect for lot size is much smaller. A \$10,000 increase in household income increases the willingness to pay for a 10 percent increase in lot size by \$9 per year. As shown in Table 7, controlling for household income, larger households are willing to pay less for more interior space and land. Our explanation for this counter-intuitive result is that household size proxies for per-capita income. For the same level of household income, larger families are poorer than smaller families and this income effect dominates any space demand effects. We are surprised by the negative income effect on the demand for new housing. In columns (2), (4) and (6), we report how actual consumption of unit square footage, the home's lot size and the age of the housing unit differ by demographic group. Richer people are buying newer, larger homes.

One of our methodological goals in this paper is to exploit both the "micro" information provided in the First American Data Base with the "macro" data from the Census. As discussed above, using information on each home buyer's last name we can identify home buyers who are Asian and Hispanic. In Table 8, we present another set of second stage preference regressions for housing structure attributes. For a household making an average of \$56,300.00 per year, Asian households value large homes more. The differences in willingness to pay for lot size and age, while statistically different, do not appear to be very large in magnitude.

In Table 9, we report our estimates of how home buyers differ with respect to their willingness to pay for community composition. Increasing household income by \$10,000 increases annual willingness to pay to live in a community with 10 percent more college graduates by \$95. Unlike the demand for high human capital peers, our estimated income effects for community racial composition are quite small. Increasing a household's income by \$10,000 increases annual willingness to pay to live in a community with 10 percent more black residents by \$19 per year.

We next consider two policy experiments where we first raise suburban density and second make the employment patterns monocentric. We conduct these policy experiments using our estimated utility parameters using a partial

equilibrium analysis. Ideally, we would allow for a general equilibrium analysis of these problems as in Bayer, McMillan and Rubin (2004) and Epple and Sieg (1999). Our policy analysis only studies the partial equilibrium effects, household by household. Therefore, the results from our policy experiments should be viewed as tentative. However, we believe that at a minimum, this analysis is a first step to understanding the incentives faced by consumers from alternative policies that could be used to discourage urban sprawl.

5.3 Policy Exercise #1: Raising Suburban Density

By estimating the housing product demand model for new home buyers, we have recovered the marginal willingness to pay for lot size and for avoiding a longer commute. Even if we have misspecified the utility function, this is sufficient information for conducting our first policy exercise.

Sprawl may impose an externality. When a household purchases a large home on a large lot, this increases his neighbor's distance to work. There is no market that prices the effect of large lots and large homes on the commute of other persons. While sprawl entails costs, it is clear that it also is associated with benefits, since larger homes and larger lots are valued. We therefore engage in the following thought experiment. Commutes could be shorter if each home's lot were "squished". There may be a co-ordination failure that given that the vast majority of housing has already been built, households cannot contract to "squish" their housing structures closer together to explicitly tradeoff less lot size for a shorter commute.

Suppose that we scrunched the city by ten percent. This would have the following effects. First, everyone would live ten percent closer to work and the commuting distance would be decreased by ten percent. Second, lot sizes and home sizes would be ten percent small. Third, population density would increase by $\frac{1}{.9} = \frac{1}{.81}$. Using data from the 1995 National Personal Transportation Survey on over 470 Los Angeles resident commuters, we estimate the following relationship between distance, population density and commuting time. The fitted equation that we use is:

$$\log(\text{commute}) = 1.13 + 0.54 * \log(\text{dist}) + 0.074 * \ln(\text{density})$$

That is, the log commute time to work is linear function of the log distance to the place of work and the log population density.

In Table 10, we estimate the effect of this change on consumer welfare in partial equilibrium, that is, holding fixed their location, but scrunching the city. On average, this would lower utility by about \$1,119 dollars per year. However, there is a wide dispersion around this number since there is heterogeneity in preferences for commuting time, home size and lot size.

We note that almost all of the negative utility from scrunching the city is because this policy makes home buyers live in smaller homes. The loss in welfare from forcing them to have smaller lots is much smaller. In fact, if we shrink the lot size, but not the home size, the results suggest that on net people would be better off by about \$2000 per year.

For each household we use our aggregation approach to calculate how the "squished city" affects different demographic groups based on their willingness to pay. The results in Table 11 show that richer households, and larger households would be willing to pay more to avoid this compression.

5.4 Policy Exercise #2: Monocentric Los Angeles

Our second policy exercise does not touch the existing residential housing stock. Instead, we propose to move all Los Angeles county employment back to the Central Business District. While nobody is proposing "monocentricizing" Los Angeles, it is of interest to calculate how home buyer well being has been affected by employment sprawl. As shown in Table 2, high sprawl residents have commute times that are only 1 minute longer than low sprawl residents. This indicates that most suburban residents must be working at suburban jobs. Employment sprawl has allowed them to avoid "monster" commutes by unbundling the commute versus land consumption tradeoff inherent in a monocentric city.

We now seek to measure this benefit of employment sprawl. To answer this question requires estimating what each household's commute time would be if they remained in their current home but worked downtown. For each zip code, we know its distance to the Los Angeles Central Business District as defined by the Census in 1982 (see Glaeser and Kahn 2001 for details).

The 1995 National Personal Transportation Survey provides micro data on household commuting patterns, distances traveled and transport modes. We use the Los Angeles subsample of 471 commuters in this data set to estimate the relationship between household's commute time, and distance to work. We use this data to estimate an OLS regression that yields:

$$\begin{aligned} \textit{commute} &= 9.2199 + 1.3538 * \textit{distance} \\ N &= 471, R^2 = .417 \end{aligned}$$

Employment sprawl has greatly benefited home buyers with a taste for larger newer housing structures. To demonstrate this, we use our First American data base's information on which zip code each household lives in to calculate its distance from the Central Business District. Plugging these values into the OLS regression equation, we predict what

each home buyer's commute would be if he had to commute to the CBD each day. These counter-factual commute times are reported in the third to last row of Table 2. Home buyers in sprawl areas commute 7.5 minute less one way than they would have if Los Angeles was monocentric (39.98-32.5). Home buyers in low sprawl areas would have much shorter commutes if Los Angeles was monocentric. This finding is quite intuitive. If all jobs were downtown, then people who live close to downtown in low sprawl areas would have shorter commutes. If willingness to pay to avoid commuting is \$29 per hour, then the average high sprawl home buyer would be willing to pay $$(29/60) \cdot 7.5 \cdot 2 \cdot 240 = \1740 per year for employment to stay sprawled.

In concluding this section, we must note that both of our policy experiments are partial equilibrium in nature. The Lucas Critique argues that substantive changes in government policy lead people to re-optimize. Yet in our experiments, we are assuming that home buyers choose the same housing product after the policy shock has taken place. For recent work that has attempted to build in such general equilibrium effects into residential sorting models see Sieg, Smith, Banzhaf and Walsh 2002, and Bayer, McMillan and Rueben 2004). In principal, it would be possible to compute such counterfactuals using our demand estimates. However, we were unable to merge data on individual specific place of work to our data. Thus, computing how households make their joint location/work decision was not possible and only a partial equilibrium analysis was possible.

Despite their limitations, we believe that these partial equilibrium policy experiments have useful lessons. Our first policy experiment suggests that there is a very large gap between the marginal valuation of land (as measured by the lot size) and the commuting externality it imposes. The current equilibrium is very far from efficient and could be improved, at least locally, by creating policies that encouraged homes to be built on smaller lots. The second policy experiment suggests that moving employment to a CBD would lead to large welfare losses. This suggests that the suburbanization of employment enhances the utility of many households.

6 Conclusion

Suburban home ownership embodies a tradeoff. Households face lower prices for newer, larger homes but are likely to spend more time commuting. Whether a household is willing to suburbanize hinges on its preferences over land consumption and avoiding commuting. Given that the educated tend to live in the suburbs, this will further encourage suburbanization for migrants who are willing to pay to live in high human capital communities.

Traditional census data cannot be used to measure this tradeoff. The housing data in the Census provides relatively little information on the types of housing households are purchasing. To provide new estimates of the private tradeoffs for new homeowners we have used recent transaction data in Los Angeles county. This investigation of housing

demand for recent home buyers in Los Angeles county reveals that the demand for sprawled housing products is fueled by the demand for interior space and access to high human capital communities. We find little evidence that the racial composition of communities sharply affects willingness to pay for housing.

One puzzle that emerges from our structural estimates is the low willingness to pay for a home's lot size. As shown in table 2, in sprawl areas the lots are twice as big as in low sprawl areas. This finding is based on within zip code hedonic regressions. Within these zip codes, observationally identical homes on larger lots do not sell for a much greater price. One explanation for large suburban lots supported by minimum lot zoning is that these lots help create a "moat" to keep poorer people from moving into communities. While our hedonics recover zip code specific fixed effects, we cannot estimate the counter-factual of what would be the value of these zip code specific hedonic fixed effects if communities did not have minimum lot zoning. In the absence of such zoning, more poor people could enter these suburban communities and this would raise local crime levels and reduce local school quality and this would be capitalized into home prices.

In addition to reporting new estimates of how home buyers tradeoff housing attributes, we have used these estimates to investigate the welfare effects of two specific "anti-sprawl" policies. Given our large estimates of the willingness to pay to avoid commuting and our small estimates of the willingness to pay for lot size, we conclude that the average home buyer would support "compacting" Los Angeles. We showed who would be the winners and losers from the radical anti-sprawl policy of moving all Los Angeles employment back to the Central Business District.

We have modeled Los Angeles county as a single housing market but in reality it is part of the greater Los Angeles region including other counties such as Orange, Ventura, San Bernardino and Riverside. The demand for sprawled suburban living continues to grow. Sixty miles east of Los Angeles county in Riverside county and San Bernardino county growth continues. Inland, in this warmer summer climate, average home prices are significantly lower. We expect that similar housing tradeoffs are taking place in such exurban fringe communities. This exurban growth should trigger employment to follow these suburbanites offering a test of Gordon, Kumar and Richardson's hypothesis that sprawl through encouraging the decentralization of employment helps to reduce commute times.

7 References.

- [1] Anas, Alex; Arnott, Richard; Small, Kenneth A. "Urban Spatial Structure," *Journal of Economic Literature*, September 1998, v. 36, iss. 3, pp. 1426-64
- [2] Anderson, S., DePalma, A., and J. Thisse (1995), *Discrete Choice Theory of Product Differentiation*, Cambridge, MIT Press.
- [3] Arnott, Richard. Time for Revisionism on Rent Control? *Journal of Economic Perspectives*, 9(1) 1995 99-120.
- [4] Bajari, Patrick and Benkard, C. Lanier, "Demand Estimation With Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach," *Journal of Political Economy*, 113(6), 1239-1276 (2005).
- [5] Bajari, Patrick and Matthew E. Kahn (2004), "Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities," *Journal of Business and Economic Statistics*, 23(1), pp. 20-33 (2005).
- [6] Bayer, Patrick, Robert McMillan and Kim Ruben, (2004) *The Causes and Consequences of Residential Segregation: An Equilibrium Analysis of Neighborhood Sorting*. Working Paper, Yale University.
- [7] Benkard, C. Lanier and Patrick Bajari, "Hedonic Price Indexes with Unobserved Product Characteristics," *Journal of Business and Economic Statistics*, 23(1), pp. 61-75 (2005).
- [8] Berry, Steven (1994), *Estimating Discrete Choice Models of Product Differentiation*, *Rand Journal of Economics*, 25(2), 242-62.
- [9] Berry, Steven; James Levinsohn, and Ariel Pakes (1995), *Automobile Prices in Market Equilibrium*, *Econometrica*, 63(4), 841-90.
- [10] Berry, Steven; James Levinsohn, and Ariel Pakes. "Voluntary Export Restraints on Automobiles: Evaluating a Strategic Trade Policy", *American Economic Review* 89 (3), 400-430.
- [11] Brown, James N and Harvey S. Rosen (1982), *On the Estimation of Structural Hedonic Price Models*. *Econometrica*, 50(3), 765-768
- [12] Brueckner, Jan (2000). *Urban Sprawl: Diagnosis and Remedies*. *International Regional Science Review*. 23(2) 160-171
- [13] Brueckner, Jan (2001). *Urban Sprawl: Lessons from Urban Economics*. *Brookings-Wharton Papers on Urban Affairs*. 65-97.
- [14] Calfee, John and Winston, Clifford "The Value of Automobile Travel Time: Implications for Congestion Policy," *Journal of Public Economics*, July 1998, v. 69, iss. 1, pp. 83-102.

- [15] Cheshire, Paul and Stephen Sheppard (1998), Estimating the Demand for Housing, Land and Neighbourhood Characteristics, *Oxford Bulletin of Economics and Statistics*, 60(3) 357-382.
- [16] Cropper, Maureen and Patrice Gordon. Wasteful Commuting: A Re-examination. *Journal of Urban Economics*. 1991, 29(1) 2-13.
- [17] Cullen, Julie Berry; Levitt, Steven D. "Crime, Urban Flight, and the Consequences for Cities," *Review of Economics and Statistics*, May 1999, v. 81, iss. 2, pp. 159-69
- [18] DiPasquale, Denise and Matthew E. Kahn (1999), Measuring Neighborhood Investments: An Examination of Community Choice, *Real Estate Economics* 27(3) 389-424.
- [19] Dougherty, A., and Van Order, R., 1982. "Inflation, Housing Costs, and the Consumer Price Index." *American Economic Review*. 154-164.
- [20] Epple, Dennis (1987), Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products, *Journal of Political Economy*, 95(1), 59-80.
- [21] Epple, Dennis and Holger Sieg (1999), Estimating Equilibrium Models of Local Jurisdictions, *Journal of Political Economy*, 107(4) 645-681.
- [22] Eskeland, Ivar, James Heckman, and Lars Nesheim (2004), Identification and Estimation of Hedonic Models. *Journal of Political Economy*, 112(1/2), S60-S109.
- [23] Fan, Jinqing and Irene Gijbels, *Local Polynomial Modelling and Its Applications (Monographs on Statistics and Applied Probability, (66), CRC Press.*
- [24] Giuliano, Genevieve and Kenneth Small (1991) Subcenters in the Los Angeles Region, *Regional Science and Urban Economics* 21(2) 163-182.
- [25] Glaeser, Edward L and Matthew E. Kahn (2001), Decentralized Employment and the Transformation of the American City *Brookings-Wharton Papers on Urban Affairs*, vol. 0, no. 0, 1-47.
- [26] Glaeser, Edward L.; Kahn, Matthew E. "Sprawl and Urban Growth," *National Bureau of Economic Research, Inc, NBER Working Papers: 9733 (2003).*
- [27] Glaeser, Edward L.; Luttmer, Erzo F. P "The Misallocation of Housing under Rent Control," *American Economic Review*, September 2003, v. 93, iss. 4, pp. 1027-46
- [28] Gordon, Peter; Kumar, Ajay; Richardson, Harry W "The Influence of Metropolitan Spatial Structure on Commuting Time," *Journal of Urban Economics*, September 1989, v. 26, iss. 2, pp. 138-51
- [29] Griliches, Zvi. "Hedonic Price Indexes for Automobiles: An Econometrics Analysis of Quality Change," in *The Price Statistics of the Federal Government*, General Series No. 73, New York: Columbia University Press for the

National Bureau of Economic Research, pp. 137-196.

- [30] Hamilton, Bruce W “Wasteful Commuting,” *Journal of Political Economy*, October 1982, v. 90, iss. 5, pp. 1035-51
- [31] Margo, Robert A “Explaining the Postwar Suburbanization of Population in the United States: The Role of Income,” *Journal of Urban Economics*, May 1992, v. 31, iss. 3, pp. 301-10
- [32] Mas-Colell (1977) The Recoverability of Consumer Preferences From Market Demand Behavior, *Econometrica*, 45(6), 1409-1430.
- [33] McCulloch, Robert E; Nicholas G Polson, and Peter E. Rossi (2000), A Bayesian Analysis of the Multinomial Probit Model with Fully Identified Parameters. *Journal of Econometrics*, 99(1), 173-193.
- [34] Mieszkowski, Peter; Mills, Edwin S. “The Causes of Metropolitan Suburbanization,” *Journal of Economic Perspectives*, Summer 1993, v. 7, iss. 3, pp. 135-47
- [35] Nechyba, Thomas and Randall Walsh. *Urban Sprawl*, *Journal of Economic Perspectives*, Forthcoming.
- [36] Nevo, Aviv “Measuring Market Power in the Ready-to-Eat Cereal Industry,” *Econometrica* , 69(2), 307-342, 2001
- [37] Painter, Gary, Stuart Gabriel and Dowell Myers. Race, Immigrant Status and Housing Tenure Choice, *Journal of Urban Economics* (2001) 49(1) 150-167.
- [38] Palmquist, Raymond (1984), Estimating the Demand for the Characteristics of Housing, *Review of Economics and Statistics*, 66(3), August, 394-404.
- [39] Peng, Ruijue and William Wheaton (1994). Effects of Restrictive Land Supply on Housing in Hong Kong: An Econometric Analysis. *Journal of Housing Research*, 5(2) 263-290.
- [40] Petrin,Amil (2002),Quantifying the Benefits of New Products: The Case of the Minivan, *Journal of Political Economy*, 110(4) 705-729.
- [41] Phillips, Justin and Eban Goodstein. (2000) Growth Management and Housing Prices: The Case of Portland, Oregon. *Contemporary Economic Policy* 18(3) 334-344.
- [42] Rauch, James (1993), Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities, *Journal of Urban Economics*, 34(3) 380-400.
- [43] Rosen, Sherwin (1974), Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, *Journal of Political Economy* 82 (January/February) 34-55.
- [44] Sieg, Holger, Kerry Smith and Spencer Banzhaf and Randy Walsh (2002), Interjurisdictional Housing Prices in Locational Equilibrium, *Journal of Urban Economics* 52(1) 131-153.

- [45] Small, Kenneth. Urban Transportation Economics. Fundamentals of Pure and Applied Economics, 51, Harwood Academic Press 1992.
- [46] Small, Kenneth and Shunfeng Song. "Wasteful" Commuting: A Resolution. Journal of Political Economy. 1992, 100(4) 888-898.
- [47] Smith, Kerry and Holger Sieg, and Spencer Banzhaf and Randy Walsh (2002), General Equilibrium Benefit Transfers for Spatial Externalities: Revisiting EPA's Prospective Analysis. RFF Working Paper 02-44 September 2002.
- [48] Thorsnes, Paul. Internalizing Neighborhood Externalities Journal of Urban Economics, 2000, v. 48, pp. 397-418.
- [49] Waugh, Frederick V. "Quality Factors Influencing Vegetable Prices," *Journal of Farm Economics*, 10:2, April, pp. 185-196.
- [50] Wheaton, William C "Land Use and Density in Cities with Congestion," Journal of Urban Economics, March 1998, v. 43, iss. 2, pp. 258-72

Table 1: Summary Statistics for Los Angeles Homes Sold Between 2000 and 2003

Variable	observations	mean	std. dev.	Asian Buyer Mean (N=14,274)	Hispanic Buyer Mean (N=50,766)
Price	173507	328695	1602895	363565	226008
Structure square feet	173507	1568.579	648.046	1793.183	1332.150
Lot Size in (sqft)	173507	7217.682	6005.796	7897.708	6703.124
Age of Structure	173507	48.105	20.940	43.171	51.420
Block Group % black	173507	0.084	0.145	0.039	0.095
Block Group % Hispanic	173507	0.349	0.265	0.262	0.538
Block group % college graduates	173507	0.265	0.188	0.333	0.152
Block group Median Income	173507	57781.090	25401.430	62567.010	46942.720
Population Density in 1000s per square mile	173507	3.216	2.388	2.911	3.906
One Way Commute time in Minutes	173505	31.282	4.970	31.392	31.903

Table 2: Housing Attributes in High Sprawl and Low Sprawl Areas in Los Angeles

Housing Attributes	Sprawl	Low Sprawl	All
Price	388597	231466	328695
Structure square feet	1847.696	1229.839	1568.579
Lot size in square feet	8715.238	5359.013	7217.682
Age of Structure	38.219	65.119	48.105
Tract % black	0.065	0.149	0.084
Tract % Hispanic	0.220	0.603	0.349
Tract % college graduates	0.345	0.130	0.265
Tract median income	71821.770	34474.620	57781.090
Tract % home owners	0.775	0.363	0.648
Tract % in poverty	0.079	0.252	0.118
Tract one way commute time In minutes	32.493	31.286	31.282
Tract one way commute time In minutes if all residents work In the Central Business District	39.983	22.813	32.706
Tract Population density in 1000s per square mile	1.308	7.981	3.216
Count of Home Sales	68,199	19,387	173,507

We sort census tracts by their population density. We assign those tracts whose density is equal to or greater than the 75th percentile to the “Low Sprawl” category while tracts whose density is less than or equal to the 25th percentile are assigned to the “High Sprawl” category. Cell specific averages are reported.

Table 3: Summary of Implicit Hedonic Prices

Variable	Mean	std. Dev.	25 %	50%	75 %
Square feet	9.078	1.718	7.848	8.810	9.953
Lot Size	0.159	0.083	0.126	0.167	0.207
Age of Structure	-9.463	33.946	-28.994	-19.556	2.386
Block % College Graduates	16199	3338.582	14381.480	15865.900	17233.410
Block % Black	-14656.590	3694.453	-16834.390	-14127.700	-12238.340
Block % Hispanic	-188.905	1560.991	-737.908	-295.219	88.052

Table 4: Linear Hedonic Price Regression to Recover Commuting Valuation

	Beta	S.e	Beta	S.e
Age of Structure	-0.473	7.017	73.331	13.782
Structure square feet	10.965	0.644	10.957	0.613
Lot size in square feet	0.123	0.039	0.051	0.043
Tract % Black	-6689.695	3178.315	-2237.928	1553.685
Tract % College Graduates	20574.020	2498.553	47787.220	3026.263
Tract % Hispanic	326.082	1444.595	4475.481	1393.349
One Way Commute in Minutes	-218.173	75.780	-488.231	93.102
Constant	8399.406	3237.439	4795.938	3612.349
Observations	173505		173505	
R ²	0.14		0.13	
Zip Code Fixed Effects	Yes		No	

The dependent variable in the regression is the average owner's equivalent rent. Standard errors are adjusted for tract level clustering.

Table 5: Differences in Consumer Willingness to Pay For Housing Attributes

Variable	All New Home Buyers			
	Mean WTP	25%	50%	75%
Square feet	1430.276	868.436	1180.252	1680.437
Lot Size	118.9519	66.630	101.345	141.048
Age of Structure	-15.48452	-120.937	-59.145	10.983
Block % College Graduates	399.6905	168.483	328.456	554.525
Block % Black	-87.9915	-115.014	-44.266	-9.116
Block % Hispanic	-6.185713	-17.178	-8.298	2.001
	Hispanic New Home Buyers			
Square feet	1058.366	713.311	923.517	1228.151
Lot Size	105.0876	69.930	98.900	126.671
Age of Structure	-94.20626	-148.276	-121.610	-61.611
Block % College Graduates	231.6113	92.974	173.721	307.285
Block % Black	-94.74129	-144.692	-48.272	-8.309
Block % Hispanic	-9.677789	-19.633	-11.100	-0.312

This table displays summary statistics for the population distribution of the willingness to pay for a 10 percent increase in each product characteristic. The units are in dollars per year.

Table 6: Correlation Between Willingness to Pay for Characteristics

	Square Feet	Lot Size	Age of Structure	% College Graduate	% Black	% Hispanic
Square Feet	1.000					
Lot Size	0.189	1.000				
Age of Structure	0.538	0.369	1.000			
Block % College Graduate	0.822	0.135	0.560	1.000		
Block % Black	0.156	0.245	0.115	0.213	1.000	
Block %Hispanic	0.464	0.056	0.004	0.341	0.315	1.000

Table 7: Willingness to Pay and Attribute Consumption
as a Function of Household Demographics

Column	Unit Square Footage		Lot Size		Age of Unit	
	(1) WTP	(2) Quantity	(3) WTP	(4) Quantity	(5) WTP	(6) Quantity
Household Size	-80.147 (34.932)	-21.445 (28.170)	-8.062 (6.391)	-324.408 (305.892)	-76.820 (6.846)	-3.673 (1.060)
Median Income in (\$1,000s)	23.971 (1.145)	14.997 (0.924)	0.916 (0.210)	14.661 (10.029)	3.213 (0.224)	-0.244 (0.035)
Constant	110.079 (165.514)	650.055 (133.474)	84.667 (30.282)	1449.377 (20.522)	20.522 (32.436)	75.985 (5.022)
Adjusted R ²	.408	.278	.033	.050	.411	.045

This table reports six separate OLS regression. There are 1000 “pseudo tracts” used in each regression. Each dependent variable represents the willingness to pay for a ten percent increase in the attribute. Willingness to pay’s units are dollars per year in owner’s equivalent rent. The quantity regressions are average consumption of the characteristic within the tract regressed on demographic characteristics. Standard errors are reported in parentheses.

Table 8: Willingness to Pay for Structure Attributes as a Function of Household Demographics and Ethnicity

Column	Unit Square Footage		Lot Size		Age of Unit	
	Hispanic (1)	Asian (2)	Hispanic (3)	Asian (4)	Hispanic (5)	Asian (6)
Household Size	-102.344 (21.159)	48.955 (53.108)	9.193 (3.603)	-11.789 (7.935)	-38.307 (3.657)	-82.441 (9.264)
Median Income in (\$1,000s)	19.041 (1.094)	28.238 (1.348)	1.589 (0.186)	1.219 (0.201)	3.074 (0.189)	3.777 (0.235)
Constant	389.439 (120.855)	-397.220 (217.490)	-15.784 (20.577)	86.275 (32.495)	-122.313 (20.887)	-0.120 (37.937)
Adjusted R ²	.369	.321	.071	.045	.439	.321

This table reports six separate OLS regression. There are 1000 “pseudo tracts” used in each regression. Each dependent variable represents the willingness to pay for a ten percent increase in the attribute. Willingness to pay’s units are dollars per year in owner’s equivalent rent. The quantity regressions are average consumption of the characteristic within the tract regressed on demographic characteristics. Standard errors are reported in parentheses.

Table 9: The Willingness to Pay for Community Attributes as a Function of Household Demographics

	Census Block % College Graduates s.e	Census Block % Black s.e	Census Block % Hispanic s.e
Household Size	-80.947 9.581	3.166 5.726	3.705 1.403
Median Income in (\$1,000s)	9.670 0.314	1.881 0.188	0.384 0.046
Constant	23.781 45.397	-222.030 27.129	-43.396 6.648
Adjusted R ²	.644	.115	.067

This table reports three separate OLS regression. There are 1000 “pseudo tracts” used in each regression. Each dependent variable represents the willingness to pay for a ten percent increase in the attribute. Willingness to pay’s units are dollars per year in owner’s equivalent rent.

Table 10: Welfare Effects of Compressing the City

Variable	MEAN	STD	25%	50%	75%
Loss from Smaller Unit Square Footage	3160	1988	1918	2607	3713
Loss from Smaller Lot Size	263	283	147	224	312
Gain from Change in Commute	2231	846	1711	2077	2507
Net Gain/Loss	-1192	2220	-1933	-715	98

Table 11: Distribution of Welfare Effects From Compressing the City

Variable	Coef.	S.E
Household Size	459.581	84.632
Median Household Income (\$1,000s)	-54.603	2.775
Constant	922.972	401.002
Observations	1000	
Adjusted R ²	.425	