Land Economics, 87(3): 412-432. 2011.

An Econometric Analysis of Land Development with Endogenous Zoning

Van Butsic
Department of Forest and Wildlife Ecology
University of Wisconsin
1630 Linden Dr.
Madison, WI 53706
butsic@wisc.edu
Ph: 608-262-9922

David J. Lewis
Department of Economics
University of Puget Sound
1500 N. Warner St., CMB 1057
Tacoma, WA 98416
djlewis@pugetsound.edu
Ph: 253-879-3553

Lindsay Ludwig
Industrial Economics, Inc.
2607 Massachusetts Ave.
Cambridge, MA 02140
lindsay.ludwig@gmail.com
Ph: 617-354-0074

Draft: April 20, 2010

Acknowledgements: We thank Bill Provencher, Dana Bauer, David Newburn, Andrew Plantinga, an anonymous reviewer, and seminar participants at the AERE sessions of the 2009 AAEA Annual Meeting and the 2009 Heartland Environmental and Resource Economics Workshop for helpful comments, Anna Haines and Dan McFarlane for spatial data assistance, and Randy Thompson for assistance with zoning data. We gratefully acknowledge support for this research by USDA McIntire-Stennis (# WIS01229). The GIS data construction was supported by the National Research Initiative of CSREES, USDA Grant # 2005-35401-15924.

Abstract

Zoning is a widely used tool to manage residential growth. Estimating the effect of zoning on development, however, is difficult because zoning can be endogenous in models of land conversion. We compare three econometric methods that account for selection bias in a model of land conversion - a jointly estimated probit-logit model, propensity score matching, and regression discontinuity. Our results suggest that not accounting for selection bias leads to erroneous estimates. After correcting for selection bias we find that zoning has no effect on a landowner's decision to subdivide in a rural Wisconsin county.

JELCode: R14; R52; Q24. **Keywords:** zoning, land use, sprawl, selection bias.

An Econometric Analysis of Land Development with Endogenous Zoning

One of the most widely discussed land management issues of recent years is urban sprawl – non-contiguous development on previously undeveloped agricultural and forested landscapes. Urban sprawl is criticized largely on the grounds that development consumes an excessive amount of land that would otherwise have provided market and non-market benefits associated with open space. A corollary of excess sprawl is the loss of farmland, since ex-urban growth often occurs in areas which are primarily agricultural. Local zoning ordinances remain probably the most widespread land use control influencing sprawl. In general, the effects of zoning on land development may vary across regions and are not well understood. Some argue that defining specific zones on the landscape for different types of development and open space can be viewed as a desirable feature of so-called "smart growth" policies (Danielson et al. 1999). In contrast, others argue that minimum lot zoning requirements can exacerbate sprawl by forcing consumption of larger lot sizes than the market would dictate in the absence of zoning (Fischel 2000). Empirical evidence regarding the effects of zoning on land development and sprawl is limited (McConnell et al. 2006), and requires an understanding of how individual landowners make decisions in response to local market conditions and zoning constraints.

Accounting for zoning policies in empirical land use models requires that researchers address the non-random application of zoning across a landscape. Including zoning in a model of land development can induce a form of selection bias in econometric estimation for at least two reasons. First, zoning policies may simply "follow the market" if local governments systematically consider the land market in the application of zoning and variance decisions (Wallace 1988). In particular, if a price differential exists between zones, then local governments will be pressured to expand the high-price zone, or to simply grant variances – an

often less costly alternative than re-writing the ordinance. To the extent that a researcher does not observe all factors which influence a parcel's development value, there is the strong possibility that the same unobservable factors that influence development will also influence zoning decisions, presenting a selection bias estimation problem commonly known as "selection on unobservables" (Cameron and Trivedi 2005; Ch. 25).

Second, zoning can induce selection bias in a land development model because parcels that are placed in a certain zone might have different distributions of the underlying covariates than parcels placed in an alternative zone. For example, parcels closer to busy roads may be less likely to be zoned restrictively due to the influence of road access on development potential. As such, zoning rules may be applied to a non-random sample (only the parcels with the unique attribute are zoned), and even if one can observe all characteristics which influence development decisions, parametric econometric methods can produce biased estimates due to differences in the distributions of the underlying covariates (Heckman et al. 1996). In this case it is difficult to separate the effect of zoning from the effect of the observed characteristic (e.g. proximity to busy roads), even though parcels are selected for specific zoning rules on "observable" characteristics.

The purpose of this paper is to conduct a parcel-level econometric analysis of the ability of local zoning (exclusive agriculture zoning (EAZ)) and statewide tax incentives (Wisconsin's Farmland Preservation Program (FPP)) to influence land use conversion in an exurban region outside of Madison, WI. Using a unique spatial panel dataset derived from five parcel level cross-sectional landscape observations between the years 1972 and 2005, we estimate the effect of EAZ and FPP on the likelihood of land development using multiple econometric techniques which correct for different forms of selection bias. While corrections for selection bias have been commonly applied to estimating the effects of zoning on property values in linear hedonic

models (Wallace 1988; McMillen and McDonald 1991; 2002), we extend the application of these techniques to non-linear models of the discrete decision of whether to subdivide and develop land – the key decision in analyses of urban sprawl.

We acknowledge the two types of selection bias discussed above and compare three econometric approaches to estimate the effects of endogenous land use policy on land development. First, we jointly estimate a parcel's selection into exclusive agricultural zoning (the zoning decision) with the decision to develop (subdivide) the parcel. The model specifies that these decisions are influenced by common observable characteristics (e.g. parcel size, distance from roads, etc.) and, importantly, common unobservable characteristics. The decisions are estimated within a joint discrete-choice framework that embeds correlated unobservables across the decisions (Greene 2006). Econometric estimation is performed with maximum simulated likelihood and allows for an empirical test of selection bias and unobserved heterogeneity with respect to inclusion of a parcel in an exclusive agricultural zone. Although the analysis extends the "selection on unobservables" approach for endogenous treatment effects to nonlinear models (e.g. Cameron and Trivedi 2005; Ch. 25), identification relies on potentially strong functional form assumptions.

Second, we perform propensity score matching to estimate the effects of zoning on the land-use decisions of those parcels that are "treated" with exclusive agricultural zoning (the average treatment effect on the treated). Matching methods exploit heterogeneity in the zoning status across parcels and provide potentially unbiased estimates of the treatment effect even if the zoning board selects parcels into exclusive agricultural zoning in a non-random fashion. In contrast to the joint discrete-choice estimation exercise, matching methods assume that selection into zoning is based only on characteristics *observable* to the researcher (e.g. prime farmland,

distance from service districts, etc.). Relative to joint estimation, the strength of matching is that it imposes minimal functional form restrictions in estimation, although the estimates will be biased if there are unobservable characteristics that influence both the zoning and development decisions.

Finally, we exploit a discontinuity in the application of FPP to examine the effects of income tax credits on the landowner's decision to subdivide. Wisconsin's Farmland Preservation Program provides income tax credits to landowners who maintain the agricultural status of EAZ parcels of at least 35 acres. Parcels less than 35 acres that are zoned EAZ are still subject to subdivision restrictions, but are not eligible for income tax credits. Thus, the discontinuity in eligibility for FPP at 35 acres allows for estimation of both semi-parametric and fully-parametric discrete-choice models over a sample where the application of FPP is quasi random.

Our estimation results yield the following basic conclusions. First, under the assumption that zoning is exogenous, exclusive agricultural zoning significantly *reduces* the probability of subdivision. In our application, this result leads to the *erroneous* conclusion that agricultural zoning significantly alters development patterns. Second, joint estimation of the zoning and development decisions provides strong evidence that the two decisions are influenced by correlated unobserved heterogeneity, contradicting the assumption of exogeneity in the first model. Joint estimation (waiving the assumption of exogeneity) indicates that zoning has *no effect* on the probability of subdivision. Third, results derived from matching methods largely confirm insights drawn from joint estimation – zoning has *no effect* on the probability of subdivision for the parcels that receive the treatment. Fourth, the discontinuity analysis shows

that eligibility for Wisconsin's Farmland Preservation Program has at most a weak effect on the probability of subdivision.

2. Empirical Analyses of Zoning

Previous economic analyses of zoning focused on the property price effects of various zoning restrictions. Relevant for our application, Henneberry and Barrows (1990) provide evidence that exclusive agricultural zoning (EAZ) *increases* farmland values in Wisconsin. However, the results from Henneberry and Barrows are contingent on an assumption that EAZ is exogenous in a model of land values. Wallace (1988) provides a widely-cited hedonic analysis of the effects of zoning on land values in King County, WA, concluding that zoning tends to "follow the market" – areas of high development value are more likely to be zoned to allow development. A series of papers by McMillen and McDonald (1989; 1991; 2002) provide further evidence on the effects of zoning on land values, concluding that zoning authorities systematically consider the local land market when selecting parcels for particular zoning rules (McMillen and McDonald 1989). A consistent estimation strategy in the hedonic literature on endogenous zoning is a two-stage estimation approach similar to Heckman's (1979) seminal two-stage sample selection model. In the first stage, the zoning decision is typically modeled as a discrete-choice decision process. In the second stage, results from the first stage are then used to correct for the endogeneity of zoning in a variant of a linear hedonic model of land values. The "selection on unobservables" approach used in this paper is motivated by the early hedonic research on two-stage models of zoning and land values, with the difference arising that our "second-stage" model is a non-linear model of the binary decision to develop land.

The recent economics literature on land-use change has focused on parcel-scale discrete-choice models of the land development decision. A variety of econometric approaches have been

used in prior work, including probit models of the binary development decision (Bockstael 1996; Carrion-Flores and Irwin 2004), conditional logit models of decisions involving agriculture, forest, and development (Newburn and Berck 2006; Lewis and Plantinga 2007), duration models of the time to conversion (Irwin and Bockstael 2002; Towe et al. 2008), and jointly estimated probit-Poisson models of the decision to develop and the decision of how many new lots to create (Lewis et al. 2009; Lewis 2009). In contrast to the hedonic literature cited above, most of the econometric land-use change literature treats zoning as exogenous in estimation (Irwin and Bockstael 2002; Newburn and Berck 2006; Towe et al. 2008), or ignores zoning altogether (Lubowski et al. 2006; Lewis and Plantinga 2007). While some analyses argue that zoning rules are exogenous in their application due to a natural experiment in policy design (McConnell et al. 2006; Towe et al. 2008; Lewis et al 2009), other analyses note the possibility that zoning is endogenous, but do not attempt to address the problem, often because zoning is not a central feature of the analysis.

Despite their common grounding in land values, it is evident that the discrete-choice land-use change literature has diverged substantially from much of the hedonic literature when it comes to handling potential selection bias associated with zoning.[1] One reason for the divergence is the fundamental difficulty associated with modeling selection bias in linear versus non-linear models. While linear models of land values can use widely-understood variants of Heckman's (1979) two-step empirical sample selection methodology, such methods are, in general, not appropriate for the type of non-linear models used in the land-use change literature (Greene 2006). However, recent advances in modeling selection problems with non-linear models (Greene 2006; Lewis et al. 2009), combined with widely-used quasi-experimental techniques such as matching methods and regression discontinuity, provide an opportunity to

reconsider how selection problems associated with zoning can be handled in discrete-choice models of land-use change.

3. Study Area, Relevant Land Use Policies, and Data

The study area for this analysis, Columbia County WI, is a fast growing county located just north of the Madison metropolitan area. While still considered rural in many areas, Columbia County has experienced significant growth in rural-urban fringe development from nearby Madison (McFarlane and Rice, 2007). Conflicts have arisen in Columbia County due to farm odors, slow machinery on roads, and the operation of machinery at late hours (Columbia County Planning and Zoning Department, 2007).

3.1 Agricultural Zoning and Farmland Preservation

In 1969, Columbia County began active attempts to slow the conversion of agricultural lands. EAZ was established in the county in 1973 in an attempt to limit rural subdivisions, and parcels zoned EAZ can subdivide under three conditions. First, EAZ parcels can create one new residence per 35 acres, as long as the residence is related to farm work. Second, landowners can ask the town board to re-zone their property to allow residential development. Third, landowners can request a variance from EAZ rules to develop their land. All three conditions appear to have been widely used since EAZ was originally established.

In 1977, the Farmland Preservation Program (FPP) was established by the State of Wisconsin to complement EAZ and preserve Wisconsin farmland through a system of tax credits and land-use restrictions. Owners of farmland can qualify for the tax credit if they sign a farmland preservation agreement restricting development of land for a specific amount of time or if their farmland is zoned for exclusive agricultural use (State of Wisconsin, 2007). Farmland owners who qualify for the tax credit may claim a sizable tax break each year; currently the

maximum an owner can claim is $4,200 a year, while the average payment in Columbia County is $641 per year. Generally, the tax credit increases as property taxes increase and household income decreases (State of Wisconsin, 2007). Given that data on whether land owners enroll in FPP is unavailable; our analysis assesses the impact of eligibility for this program.

A convenient feature of zoning in Columbia County is that areas not zoned EAZ have a uniform minimum lot size: 20,000 sq ft (15,000sq ft for panels prior to 1991). In our setting, we propose that minimum lot size is the most restrictive facet of zoning, as lot size restrictions will likely have a larger influence than other facets (such as minimum set backs or height restrictions) on the ability of a landowner to subdivide. Therefore, in this setting, the regulatory landscape can generally be described by two zones: EAZ and non-EAZ. This allows for estimation of zoning as a binary treatment variable.

3.2 Spatial-Temporal Data and Development Trends

We obtain spatial data on development decisions and parcel attributes over a number of years for two townships in Columbia County: Lodi and West Point, neighboring townships located in the southwest corner of the county bordering Lake Wisconsin and the Wisconsin River (Figure 1). The parcel level data was generated by the Center for Land Use Education at the University of Wisconsin – Stevens Point. Property boundaries were reconstructed over the study area for five points in time: 1972, 1983, 1991, 2000, and 2005. Using 2005 digital parcel data, historic property boundaries were recreated through a process of “reverse parcelization” that selects and merges parcels using historic tax records and plat maps (see McFarlane (2008) for a complete description of the data construction). Zoning data is constructed from the Columbia County Planning Department.

The full dataset has 21,798 individual parcel observations. Parcels that could not legally subdivide were dropped from this dataset; these include public lands and parcels too small to subdivide due to zoning restrictions. Additionally, all parcels adjacent to Lake Wisconsin were dropped from the analysis because waterfront parcels are arguably part of a different land market than non-waterfront property[2]. The final dataset used for estimation contains 5,764 observations. A host of variables are thought to influence the decision to enroll a parcel in EAZ and the decision to subdivide. A list of variables used in the econometric analysis is presented and summary statistics for the variables are presented in Table 1[3].

More than 30 years after EAZ and the FPP were established, Lodi and West Point townships are still experiencing a loss of agricultural lands. Out of 1,186 developable parcels in our data set in 1972, 328 (28%), subdivide by the year 2005. There are 539 parcels zoned EAZ that are eligible for FPP in 1972, and 132 (24%) of these parcels subdivide by 2005. There are 386 parcels in EAZ that are too small to qualify for FPP, and 77 (20%) of these subdivide by 2005. For the non-EAZ parcels 92 of the 228 (40%) parcels less than 35 acres subdivided by 2005, while 27 of the 33 (81%) parcels larger than 35 acres subdivided over our study period. Thus, summary statistics indicate that parcels zoned EAZ and those eligible for FPP payments are less likely to subdivide. However, summary statistics also indicate that development certainly happened on parcels with various combinations of EAZ and FPP, indicating the widespread application of re-zoning and variances in this region (see Ludwig 2008 for further information).

The data from Lodi and West Point townships admittedly represents a small geographic area compared to land use change models which use data from full counties (Lewis et al. 2009), multiple counties (Bockstael 1996., Lewis and Plantinga 2007), or nationally (Lubowski et al.

2008). In land use change models, a small geographic sample raises two concerns. First, if the small geographic location results in a small sample size, this can lead to type 1 errors. The panel nature of our data increases the sample size to 5,764 observations, large enough to assure statistical precision. Additionally, when we use econometric techniques that do not exploit the panel nature of the data (resulting in smaller sample sizes) our results remain relatively stable compared to models that use the full sample. Second, the transferability of these results to other settings may be hindered by the specialized sample. However, the townships examined here share multiple characteristics typical of ex-urban townships: proximity to urban areas, mixed agriculture and large-lot subdivision, and zoning boards comprised of local landowners.

Expanding our sample geographically is prohibitive for two reasons. First, historical reconstruction of parcel level land use change is labor intensive and expensive. Second, expanding the geographic area would hinder our identification strategy. The fact that zoning is binary in our sample (EAZ or non-EAZ) allows us to use econometric techniques appropriate for evaluating binary treatments. Using data from additional municipalities would introduce other land use policies, negating our ability to use these techniques. Overall then, while the sample comes from a small geographic area, the number of observations is large enough to ensure statistical precision, the townships are typical of exurban development, and the townships provide a unique mechanism for evaluating the effects of land use policy.

4. Estimating the Effects of Exclusive Agricultural Zoning on Development

The landowner's decision problem is cast as a problem of whether to subdivide and develop their land at time *t*. Much of the land-use literature is motivated by Capozza and Helsley's (1989) deterministic optimal stopping problem, whereby development takes place once development rents (assumed to be increasing over time) equal the rents from agriculture

(assumed to be constant over time). We cast the decision problem in terms of the reduced form net land value of subdividing at time *t*, where S_{nt}=1 if parcel *n* subdivides in time *t*, and S_{nt}=0 otherwise. Formally, the land value of subdivision is LV_{nt}, and subdivision occurs when:

$$LV_{nt} = V\left(w_{nt}, EAZ_{nt}\right) + \sigma\mu_n + v_{nt} > 0 \text{, (1)}$$

where w_{nt} is a set of observable parcel characteristics, EAZ_{nt} is a binary indicator of the zoning status of parcel n, and μ_n and v_{nt} denote parcel-specific characteristics observed by the parcel owner but not by the analyst. We model μ_n as an iid standard normal random effect to reflect the panel structure of our data – repeated parcel-level decisions are observed over time.

The zoning agency's decision problem is cast as a problem of whether to impose exclusive agricultural zoning status on parcel *n* (EAZ_{nt}=1) or not (EAZ_{nt}=0). As is typical in local governments throughout the United States, the landowner of parcel *n* can lobby the local government regarding the zoning decision. The net value to the zoning agency of imposing EAZ status on parcel n is defined as VZ_{nt}, and EAZ_{nt}=1 when:

$$VZ_{nt} = G(x_{nt}) + \varepsilon_{nt} > 0 \qquad (2)$$

where x_{nt} is a set of parcel characteristics observable to the researcher and the zoning agency, and ε_{nt} is a set of parcel characteristics observable to the zoning agency but not the researcher. In this setting, some of the same observable characteristics that influence zoning can also influence the net value of subdividing ($x_{nt} \in w_{nt}$), and, importantly, some of the same unobservable characteristics that influence zoning can be correlated with unobservable characteristics that influence the net value of subdividing. Such correlation implies that EAZ_{nt} is an endogenous variable when attempting to estimate the parameters in

4.1 FIML Estimation – Selection on Unobservables

One approach to obtaining a consistent estimate of the effects of EAZ_{nt} on development is to jointly estimate (1) and (2) with correlated unobservables across equations. Such a strategy can be implemented with a fully parametric approach, and we adopt such a framework in this section. In particular, we make the following assumption:

$$(\mu_n, \varepsilon_{nt}) \sim N[(0,0),(1,1,\rho)] \tag{3}$$

By further assuming that v_{nt} is logistically distributed, we follow Greene (2006) and model the two equations as a joint probit-logit model. In particular, by writing $V(w_{nt}, EAZ_{nt})$ as a linear function of parameters, the probability that farm *n* subdivides in time *t*, conditional on w_{nt}, μ_n, and EAZ_{nt} can be written:

$$P[S_{nt} = 1 \mid w_{nt}, \mu_n, EAZ_{nt}] = \frac{\exp[\beta w_{nt} + \lambda EAZ_{nt} + \sigma\mu_n]}{1 + \exp[\beta w_{nt} + \lambda EAZ_{nt} + \sigma\mu_n]} \tag{4}$$

Further, by writing $G(x_{nt})$ as a linear function, Greene (2006) shows that the probability of the observed EAZ behavior on farm *n* in time *t*, conditional on x_{nt} and μ_n, can be written as:

$$P[EAZ_{nt} \mid x_{nt}, \mu_n] = \Phi\left((2EAZ_{nt} - 1)[\alpha x_{nt} + \rho\mu_n] / \sqrt{1-\rho^2}\right) \tag{5}$$

where the term $2EAZ_{nt}$-1 is a computational and notational convenience that exploits the symmetry of the normal distribution. Conditional on w_{nt}, x_{nt}, and μ_n, the joint probability of the observed behavior on parcel n is:

$$\begin{aligned}&\Pr\left[EAZ_{nt}, S_{nt} \middle| x_{nt}, w_{nt}, \mu_n\right] = \\ &\Pr\left[EAZ_{nt} \middle| x_{nt}, \mu_n\right] \cdot \left\{\begin{array}{l}(1-EAZ_{nt}) \cdot \left(S_{nt} \cdot \Pr\left[S_{nt} = 1 \middle| w_{nt}, EAZ_{nt} = 0, \mu_n\right] + (1-S_{nt}) \cdot \Pr\left[S_{nt} = 0 \middle| w_{nt}, EAZ_{nt} = 0, \mu_n\right]\right) \\ +EAZ_{nt} \cdot \left(S_{nt} \cdot \Pr\left[S_{nt} = 1 \middle| w_{nt}, EAZ_{nt} = 1, \mu_n\right] + (1-S_{nt}) \cdot \Pr\left[S_{nt} = 0 \middle| w_{nt}, EAZ_{nt} = 1, \mu_n\right]\right)\end{array}\right\}\end{aligned} \tag{6}$$

The unconditional probability of the observed behavior is generally stated:

$$\Pr\left[EAZ_{nt}, S_{nt} \middle| x_{nt}, w_{nt}\right] = \int \Pr\left[EAZ_{nt}, S_{nt} \middle| x_{nt}, w_{nt}, \mu_n\right] \phi(\mu_n)\, d\mu_n \tag{7}$$

Equation (7) can be solved with maximum simulated likelihood by taking R draws from the normal distribution of μ_n. The log likelihood function to be maximized over *N* parcels is:

$$\sum_{n=1}^{N} \log\left[\frac{1}{R}\sum_{r}\prod_{t}\Pr\left[EAZ_{nt}, S_{nt} \middle| x_{nt}, w_{nt}, \mu_n\right]\right] \tag{8}$$

This function is maximized by choice of the parameter vector $(\beta, \lambda, \alpha, \sigma, \rho)$, and accounts for correlated unobservables across the decisions to zone and subdivide, and the panel structure of the data by modeling random parcel effects. The correlation coefficient ρ deserves special attention. In this model, ρ corrects and tests for unobserved selection bias between the decisions to zone and the decision to subdivide. The sign of ρ indicates the direction of correlation between the joint decisions, while its magnitude and standard error measure its significance. A negative statistically significant ρ indicates that parcels that are more likely to be zoned EAZ are less likely to subdivide.

4.2 Propensity Score Matching - Selection on Observables

An alternative to the FIML model is the use of propensity score matching. In this setting, EAZ is still modeled as endogenous to the decision to subdivide, but we assume that we can observe all important inputs to the decision to zone a parcel EAZ and the decision to subdivide. Additionally, we assume that the same characteristics that influence the decision to zone a parcel EAZ also influence the decision to subdivide. Matching works by comparing outcomes on parcels that were zoned EAZ and those that were not zoned EAZ but are similar in observed baseline covariates. The goal of matching is to make the covariate distributions of EAZ and non-EAZ parcels similar. In this way matching mimics a random sample. Following the notation used earlier, but with unscripted letters equaling population averages, the average treatment effect for the treated (ATT) is defined:

$$\tau_{ATT} = E(\tau \mid EAZ = 1) = E[S(EAZ = 1) \mid EAZ = 1] - E[S(EAZ = 0) \mid EAZ = 1) \quad (9)$$

The key is to find a proxy for the unobservable counter factual $E[S(EAZ = 0) \mid EAZ = 1)$. Under the assumption of common support and unconfondedness (Caliendo and Kopeinig 2008),

$$\tau_{ATT} = E_{C|Z=1}\{E[S_1 \mid EAZ = 1, C = c] - E[S_0 \mid EAZ = 0, C = c]\} \quad (10)$$

where C is a vector of characteristics that affect both the selection into EAZ and the likelihood of subdivision, and the subscript on S denotes the outcome (1 = subdivision; 0 = no subdivision). Matching on C implies controlling for a high dimensional vector. Thus we follow the insights of Rosenbaum and Rubin (1983a) and use the propensity score defined as $P(C) = prob(EAZ = 1 \mid C)$, which is the probability that a parcel is zoned EAZ given its set of covariates C. We can rewrite the estimate of ATT as:

$$\tau_{ATT}^{PSM} = E_{c|EAZ=1}\{E[S_1 \mid EAZ = 1, P(C)] - E[S_0 \mid EAZ = 0, P(C)]\} \quad (11)$$

In order to implement propensity score matching we must specify the zoning selection equation, which assigns a propensity score to each observation. The selection equation should only include variables that affect the participation decision (zoned EAZ or not) and the subdivision outcome (Heckman et al. 1998 and Dehejia and Wahba 1999). In our case, we use a probit specification similar to the "first stage" of the FIML model.

Formally, to derive equation (11), two conditions need to hold (Becker and Ichino 2002). First, the pretreatment variables must be balanced given the propensity score:

$$EAZ \perp C \mid P(C) \quad (12)$$

Second, the assignment to the treatment must be unconfounded given the propensity score:

$$S_0, S_1 \perp EAZ \mid P(C) \quad (13)$$

If equation (12) is satisfied, the distribution of the underlying covariates is the same regardless of treatment. That is, the treatment is randomly assigned. Therefore, treated and untreated parcels will be observationally identical on average. To validate these two requirements, we implement the propensity score matching algorithm derived by Becker and Ichino (2002) which assures that the propensity scores used for comparison are balanced in the underlying covariates.

A variety of matching estimators exist which have different trade-offs between variance and bias. The central questions when choosing a matching estimator are what constitutes a match and should one match with or without replacement? There is little theory to guide the choice of matching estimators –matching without replacement yields the most precise estimates – but only in relatively large datasets. We follow Caliendo and Kopening (2008) and test multiple matching estimators. We utilize radius matching, kernel matching and nearest neighbor matching without replacement to estimate the ATT of EAZ on parcels not eligible for FPP. Finally, we check for "hidden bias" that may occur if there is unobserved heterogeneity in our dataset using Rosenbaum bounds (Becker and Caliendo 2007).

We model each panel as an individual experiment where the treatment is applied at the beginning of each panel and the outcome is the state of the parcel at the beginning of the following panel. In total, we estimate 12 equations (3 matching estimators x 4 panels) to estimate the effect of EAZ on the likelihood of a parcel to subdivide. The effect of EAZ on development is identified separately from the effect of FPP by limiting our sample to those parcels less than 35 acres in size, and thus not eligible for FPP.

4.3 Regression Discontinuity (RD) –Effects of eligibility for the Farmland Preservation Program

Turning our attention to estimating the effect of FPP eligibility we return once again to the selection of a parcel into EAZ, equation (2). In our setting there is a sharp discontinuity

where parcels that receive the treatment in equation (2) are eligible for FPP only if they are larger than 35 acres. Thus we are faced with a second policy assignment:

$$FPP_n = \begin{cases} 1 \; if \;\; EAZ_n = 1 \; and \;\; acres \geq 35 \\ 0 \; if \;\; EAZ_n = 1 \; and \;\; acres < 35 \end{cases} \tag{14}$$

Where FPP_n represents the eligibility of an individual parcel for FPP, EAZ_n is the state of zoning and $acres$ is the size of the parcel. As acres is likely correlated with the decision to subdivide, the assignment mechanism is clearly not random and a comparison of outcomes between treated and non-treated parcels is likely to be biased. If, however, parcels close to 35 acres are similar in the baseline covariates, the policy design has some desirable experimental properties for parcels in the neighborhood of 35 acres.

Using the sharp regression discontinuity framework from Imbens and Lemieux (2008), we can estimate the average causal effect of eligibility for FPP by looking at the discontinuity in the conditional expectations of the outcome.

$$\lim_{acres \downarrow 35} E[S_n \mid Acres_n = acres] - \lim_{acres \uparrow 35} E[S_n \mid Acres_n = acres] \tag{15}$$

The average causal effect of eligibility for FPP at the discontinuity of 35 acres is:

$$\tau_{FPP} = E[S_n(FPP = 1) - S_n(FPP = 0) \mid acres = 35) \tag{16}$$

By assuming that the conditional regression functions describing the subdivision decision are continuous in acres at the discontinuity (Imbens and Lemieux 2008), we can rewrite the estimate of the treatment effect for being eligible for FPP as:

$$\tau_{fpp} = \lim_{acres \downarrow 35} [S \mid Acres = 35] - \lim_{acres \uparrow 35} [S \mid Acres = 35] \tag{17}$$

which is the difference of two regression functions at a point. Intuitively, by comparing parcels that are near the discontinuity that receive and do not receive the treatment, we can identify the

average treatment effect for parcels with values of *acres* at the point of discontinuity (Lee and Lemieux 2009).

We estimate this effect in two ways. First, we use a semi-parametric procedure developed by Nichols (2007) which uses local linear regression to estimate the average treatment effect for the treated around the point of the discontinuity. Second, we specify probit regressions with the discontinuity entering the estimation equation as a dummy variable (Imbens and Lemieux 2008). We specify these regressions over a number of distances away from the discontinuity. In both cases we present graphical evidence of the discontinuity. Finally, Lee and Lemiex (2009) show that in RD, panel datasets can be effectively analyzed as a single cross section. Thus, we estimate the probit models with clustered errors, but no random effects.

4.4 Summary of the models

The four models estimate different treatment effects and are based on different underlying functional form and selection bias assumptions. The FIML models from section 4.1 estimate the average treatment effect of both EAZ *and* FPP eligibility across all parcels. The matching estimator in section 4.2 estimates the average treatment effect for those parcels treated with EAZ, but not eligible for FPP. And the RD method in section 4.3 estimates the effect of FPP eligibility on parcels that are treated with EAZ. The FIML models are based on explicit assumptions regarding the underlying distributions of the unobservables, while the matching and RD estimators have much weaker functional form assumptions.

To demonstrate the importance of accounting for endogenous land use policy in models of land use conversion, we also estimate a binary logit model of the subdivision decision to quantify the effects of EAZ and FPP under the assumption that both policies are exogenously applied. In contrast, the FIML model assumes that parcels are selected into zoning based on

observable and unobservable factors that may also influence the development decision. Matching and RD estimators assume that zoning selection is based only on observable components, where identification is based off either the balancing of the propensity score, or manipulation of the sample, respectively. Table 2 presents a summary of the underlying assumptions concerning the endogeneity of EAZ and FPP eligibility in the analysis.

Finally, we note that the decision to subdivide may be different than the decision to develop. For instance, inherited farmland may be split between relatives, but the use of the land may remain agricultural. For policy purposes the change in ownership may be irrelevant unless land use changes in some way. To address this, we ran all the models on the same data but where subdivisions were only counted if a new structure was built by the year 2005 (the last year of our data). The results of these models mirror the results presented in the next section.

5. Results

5.1 *Regression techniques*

Estimated parameters for the FIML model and the independent probit and logit models of the zoning and subdivision decisions are presented in table 3 for the period 1972-2005.[4] We hypothesize that whether a parcel is zoned EAZ is a function of its size, land use, and location. The results of the first stage FIML probit regression bear this out: the size, land use, and location of the parcel all significantly influence the likelihood it is zoned EAZ. Of particular interest for this analysis is the estimate of ρ, the coefficient of correlation between the unobservables across the subdivision and zoning decisions, in the FIML estimator. Our estimate of ρ is -0.74, indicating that parcels with unobservables that make them *more* likely to subdivide have unobservables that make them *less* likely to be zoned exclusive agriculture. The estimate of ρ is

significantly different from zero at the 5% level and provides evidence that estimates of EAZ in the binary subdivision model suffer from selection bias.

The policy relevant variables in the logit model and the logit component of the jointly estimated FIML model, EAZ and FPP eligibility, are best interpreted through discrete change effects rather than parameter estimates. The discrete change effects of EAZ and FPP eligibility from the binary logit model are both negative and significantly different from zero (Figures 2 and 3), indicating that under the assumption that EAZ is exogenously imposed, parcels zoned EAZ are *less* likely to subdivide. However, when the assumption of exogeneity is relaxed in the FIML model, the results change substantially. The discrete change effects of EAZ and FPP eligibility estimated with the FIML model are not significantly different from zero at any reasonable confidence level[5], indicating that we cannot reject the null hypothesis that the zoning policies have *no effect* on the probability of subdivision when we allow correlated unobservables across the zoning and subdivision decisions.

5.2 Propensity score matching

The specification of the propensity score follows closely to the probit selection equation estimated using the regression techniques, with the addition of some higher order terms to assure proper balance between the covariates. Specifications of the selection equation vary slightly from panel to panel to assure that the balancing algorithm of Becker and Inchino (2002) is met for each specification[6]. Table 4 presents the results of the selection equation for 2001-2005, where EAZ is the dependent variable and a probit specification is used. Overall, the size of the parcel, distance to services, distance to Lodi, distance to water, and land use, significantly affect the likelihood of a parcel being in EAZ, the other panels mirror this result.

There is some variation between panels and between estimators in the magnitude and standard error of EAZ's average treatment effect on the treated (ATT). In all cases the ATT is negative – although for the panels 1972-1983 and 1983-1991 results are not significantly different from zero (Table 6). For the panel 1991-2000 the nearest neighbor algorithm estimates a statistically significant -7 percentage point change in the likelihood of subdivision, for 2000-2005 this estimate is statistically significant and -4 percentage points. All other estimates for 1991-2000 and 2000-2005 are not significantly different from zero.

When significant effects of EAZ were detected, we tested the sensitivity of these results to "hidden bias" using the basic formulation from Rosenbaum(1983b). We use the Mantel-Haenszel (Mantel and Haenszel 1959) test statistic to measure how strongly an unobserved variable would have to influence the selection process to undermine the implications of the matching analysis. The effect of an unobserved variable on the selection into EAZ, γ, is simulated over various values – where larger γ values simulate higher levels of hidden bias. For each value of γ, the Mantel-Haenszel statistic is calculated. As γ increases we can detect the point at which the implications of the matching estimator are no longer valid – the point at which Mantel-Haenszel statistic becomes statistically insignificant. For the 1991-2000 panel, we find the matching estimates are sensitive to unobserved bias which would increase the odds of being selected into EAZ by 40%. That is, the existence of an unobserved variable which would increase the odds of being zoned EAZ by 40%, makes our estimates of the treatment effect null. The 2000-2005 estimates are sensitive to bias that would increase the odds of being selected into EAZ by 20%.[7]

5.3 *Regression Discontinuity*

Graphical analysis plays an important role in RD and we present three graphs here (Figure 4). First, we note that there are many observations near the discontinuity of 35 acres. In our setting, 35% of all parcels in the data set are in EAZ and are between 25-45 acres in size, and 50% of all parcels in EAZ fall within this range. Figure 4 also presents the mean probability of subdivision for 5 acre bins along with the number of observations. Of particular note is the drop in the mean probability of subdivision between 25-35 acres and 35-45 acres (also note that the number of observations between 25-35 acres (n=259) is much smaller than between 35-45 acres (n=1727, which may increase the standard errors of our estimate). Finally we fit a kernel density function to this data and include a break at the discontinuity[8]. We note a large discontinuity at 35 acres, indicating that FPP eligibility may have an effect on the propensity to subdivide.

A semi-parametric methodology developed by Nichols (2007) is used to estimate the effect of FPP eligibility on the likelihood of a parcel to subdivide. In this method, local linear regressions are run on each side of the discontinuity to estimate the local Wald statistic which can be interpreted as the percentage point change induced by FPP eligibility in the area around the discontinuity.[9] The local linear regressions rely simply on the running variable –acres in this case – and the outcome variable –whether or not a subdivision happens, along with specifying the discontinuity. Estimates may be sensitive to bandwidth choice, which dictates how far observations are used from the discontinuity. McCrary (2007) suggest that visual inspection of the local linear regressions around the discontinuity is the most effective way to select a bandwidth. We do this and find an optimal bandwidth around 3. To check the sensitivity of our estimates we estimate the effect of FPP over multiple bandwidths.

An alternative RD method involves running a probit model over the sample data around the discontinuity (Greenstone and Gallagher 2008). In this case, the effect of FPP eligibility can

be estimated with a dummy variable[10]. Other variables that we assume affect the likelihood of subdivision are also included in the probit model such as acres, land use, and location of the parcel. The running variable – acres- enters the model linearly. Choosing which parcels are "near" the discontinuity (Imbens and Lemieux 2008) is admittedly at the discretion of the researcher, therefore we use multiple breakpoints to check for sensitivity in our analysis[11]. While there was some sensitivity in regards to standard errors, the main findings are consistent over the range of estimates. We present the full results of one probit model (all years, acres between 25-45) in Table 6.[12]

The RD results all find negative effects of FPP eligibility on the probability of subdivision, but only the semi-parametric design produces results that are statistically different from zero (Table 7). In general, these results suggest that the effect of FPP eligibility on the propensity of parcels which are zoned EAZ to subdivide may be negative around the discontinuity. Combined, the two RD methods provide some evidence in favor of an effect of FPP on subdivision, although the bulk of evidence indicates that this effect is weak.

5.4 Discrete change effects of EAZ and FPP eligibility

It is useful to scale the results such that they are easily comparable. Discrete change effects in this setting can be interpreted as the percentage point change in the probability of subdivision for the given treatment (either EAZ or FPP). Some care is still needed when interpreting the discrete change effects since the actual treatment effects vary between estimators. Overall, the majority of the estimates in Figure 2 suggest that we fail to reject a null hypothesis that EAZ has *no effect* on the propensity of landowners to subdivide. The binary logit models that assume no selection bias have discrete change effects around -5 percentage points. Given that correlated unobservables are found in the jointly estimated model, and the propensity

score estimates (nearest neighbor matching) are sensitive to unobserved "hidden bias", it is likely that these results are erroneous. The FIML estimates and the majority of the propensity score estimates find no effect of EAZ on the propensity to subdivide. The bulk of the evidence suggests that EAZ likely has no effect on the likelihood of a parcel to subdivide.

The story for FPP eligibility is less clear (Figure 3). Both the binary logit model (no assumed selection bias) and the semi-parametric RD model from 1983 produce statistically significant effects of FPP eligibility. As mentioned earlier, the binary logit model is likely affected by selection bias. The semi-parametric RD models, however do offer some evidence that FPP eligibility may affect the likelihood of a parcel to subdivide. In contrast, the FIML model and the probit discontinuity model find no evidence that FPP eligibility affects the likelihood of a parcel to subdivide. We conclude, therefore, that FPP eligibility likely has a weak effect (if any effect at all) on the likelihood that a landowner subdivides.

6. Discussion

We present multiple methods to estimate the effect of endogenous land use policy on the likelihood of rural landowners to subdivide. This exercise leads to two main results. First, we cannot reject a null hypothesis that Columbia County's exclusive agricultural zoning program (EAZ) has *no effect* on development decisions, while Wisconsin's Farmland Preservation Program (FPP) of tax credits has at most a *weak effect* on the development decisions of rural landowners in our study area. Second, we find evidence that including zoning as an exogenous explanatory variable in land development models can lead to selection bias resulting in erroneous inference regarding the effects of land-use policies on development decisions.

Our results show that consistent estimates of the effects of land use policy require the researcher to seriously consider the potential for selection bias in land conversion models. While

the hedonic literature on zoning has long accounted for endogenous policy application, less attention has been paid to this issue in the land conversion literature. In our setting, three very different econometric methods – FIML estimation, propensity score matching, and regression discontinuity – prove useful at addressing the endogeneity of zoning. Even though the propensity score matching estimator cannot account for unobserved selection bias, the examination of Rosenbaum bounds allows us to evaluate whether these estimates are sensitive to the presence of unobserved selection bias. The regression discontinuity analysis, in general, produces estimates of FPP eligibility that are consistent with results that correct for unobserved selection bias. Our favored estimates are from the FIML model of the jointly estimated zoning-subdivision decision, although we recognize the critique that this method relies extensively on functional form assumptions for identification. Nevertheless, joint estimation provides a plausible identification strategy and generates estimates that can be used in spatial landscape simulations where econometric estimates are linked with a GIS to examine how multiple individual decisions influence larger landscapes (Lewis and Plantinga 2007). Future research in land use conversion models would be well served by focusing more attention on methods to properly model selection bias arising from the non-random application of land use policy.

As a policy-relevant finding, we cannot reject the null hypothesis that EAZ has no effect on landowner development decisions, while FPP eligibility has at most a weak effect on these decisions. The fact that EAZ does not influence subdivision decisions hints that, in this application, zoning may simply "follow the market". That is, restrictive zoning – such as EAZ – is likely to be applied to parcels that are unlikely to subdivide whether they are zoned or not. This result is consistent with previous work done using hedonic analysis which finds that areas of high development potential are often zoned to allow development (Wallace 1988; McMillen

and McDonald 1989). The result that FPP at most weakly influences the landowner's decision to subdivide is not surprising given the small benefit to the landowner from FPP (on average $641 per year/ per farm) compared with the much larger gains possible from subdividing (upward of $7,000 per acre if left in agricultural use and possibly much higher in residential use (Anderson and Weinhold (2008)). This result indicates that, at least in the region of the state we analyze, the money Wisconsin spends on FPP annually has little effect on farmland preservation.

Our use of an admittedly small region – two townships in one exurban county near Madison, WI – leads to both strengths and weaknesses of our analysis. A clear strength of the small region is the reduction of zoning policy into a binary variable – exclusive agricultural zoning or not – amenable to contemporary treatment evaluation techniques. Analyzing significantly larger regions would provide far less policy clarity, given the fact that zoning rules typically exhibit significant variation across municipalities. However, while the small region of analysis provides empirical clarity, such clarity comes at the expense of generalizability of the results to other regions. Nevertheless, a primary purpose of our analysis is to demonstrate and examine multiple empirical *methods* to account for the endogeneity of zoning in land conversion models. To the extent that zoning rules in other exurban regions are set by democratically elected boards comprised of local residents and landowners – as occurs in our study region – then the methodology and empirical issue of endogenous zoning will likely be relevant issues for many other researchers.

The evidence presented here suggests that zoning does not alter land development. Corollaries of this result are troubling for other land conservation programs where landowners can influence whether or not they receive a conservation "treatment". For example, the purchase of development rights (PDR) by governments and non-profits are popular ways to preserve

farmland in perpetuity, and are often credited with preserving open space. However, it is easy to imagine a situation analogous to our findings concerning EAZ – those landowners who are least likely to subdivide in the absence of a conservation program (those who wish to continue farming) may be the most likely to sell their development rights. If this is the case, the amount of land "preserved" through PDR programs may be overstated – at least in the short run - due to the fact that some of the farmland likely would not develop even in the absence of the PDR payment. An analogous situation exists for conservation easements and nature reserves (Andam et al. 2008). More research investigating whether PDR programs and other conservation policies simply "follow the market" may be a valuable line of inquiry that would help policy makers better decide which lands to preserve and how to best go about preserving them.

References

Andam, K.S, P.J Ferraro, A. Pfaff, G.G. Sanches–Azofeifa, J.A., Robalino, 2008. “Measuring the effectiveness of protected area networks in reducing deforestation” *Proceedings of the National Academy of Sciences of the United States of America* vol.105 (42), 16089-16094.

Anderson, K. D. Weinhold. 2008. “Valuing Future Development Rights: The Cost of Conservation Eassements.” *Ecological Economics* 68(1) 437-436.

Becker,S., A. Ichino, 2002. "Estimation of Average Treatment Effects Based on Propensity Scores." *The Stata Journal* 2 (4):358-377.

Becker,S., M. Caliendo, 2007. “Sensitivity analysis for average treatment effects” *The Stata Journal* 7(1):71:83.

Bento, A., Towe, C., and J. Geoghegan. 2007. “The Effects of Moratoria on Residential Development: Evidence from a Matching Approach.” *American Journal of Agricultural Economics,* 89(5): 1211-1218.

Bockstael, N.E. 1996. “Modeling Economics and Ecology: The Importance of a Spatial Perspective.” *American Journal of Agricultural Economics* 78:1168-1180.

Caliendo, M., S. Kopeinig 2008. “Some Practical Guidance for the Implementation of Propensity Score Matching.” *Journal of Economic Surveys* 22(1):31-72

Cameron, A.C., and P.K. Trivedi. 2005. *Microeconometrics: Methods and Applications.* Cambridge University Press.

Capozza, D.R. and R.W. Helsley. 1989. “The Fundamentals of Land Prices and Urban Growth.” *Journal of Urban Economics,* 26: 295-306.

Carrion-Flores, C. and E.G. Irwin. 2004. "Determinants of Residential Land-Use Conversion and Sprawl at the Rural-Urban Fringe." *American Journal of Agricultural Economics* 86(4): 889-904.

Columbia County Planning and Zoning Department. 2007. "Columbia County Comprehensive Plan 2030." Adopted September 19, 2007.

Danielson, K.A., R.E. Lang, and W. Fulton. 1999. "Retracting Suburbia: Smart Growth and the Future of Housing." *Housing Policy Debate*, 10 (3): 513–553.

Dehejia, R.H., S. Wahba. 1999 "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs" *Journal of the American Statistical Association* 94(448):1053-1062

Fischel, W. 2000. "Zoning and Land Use Regulation." in: B. Boudewijn and G. De Geest (Eds.), *Encyclopedia of Law and Economics, Volume II: Civil Law and Economics*, Edward Elgar, Cheltenham: 403–423.

Greene, W. 2006. "A General Approach for Incorporating Selectivity in a Model." Working Paper, Department of Economics, New York University.

Greenstone, M., J. Gallagher. 2008. "Does Hazardous Waste Matter? Evidence form the Housing Market and the Superfund Program" *Quarterly Journal of Economics* 123(3):951-1003

Heckman, J.J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica*, 47(1): 153-161.

Heckman, J.J., H.Ichimura, J.Smith, and P. Todd. 1996. "Sources of Selection Bias in Evaluating Social Programs." *Proceedings of the National Academy of Sciences*, 93(23): 13416-13420.

Heckman, J.J., H. Ichimura, P. Todd 1998. “Matching as an Economic Evaluation Estimator” *Review of Economic Studies* 65(2):261-294

Henneberry, D.M., and R.L. Barrows. 1990. “Capitalization of Exclusive Agricultural Zoning into Farmland Prices.” *Land Economics*, 66(3): 249-258.

Imbens,G.W., T. Lemieux. 2008. “Regression Discontinuity Designs: A Guide to Practice” *Journal of Econometrics* 142(2):615-635.

Irwin, E. and N.E. Bockstael. 2002. “Interacting Agents, Spatial Externalities and the Evolution of Residential Land Use Patterns.” *Journal of Economic Geography* 2: 331-54.

Irwin, E.G., and N.E. Bockstael. 2004. “Land Use Externalities, Open Space Preservation, and Urban Sprawl.” *Regional Science and Urban Economics* 34: 705-725.

Lee D.S, and T. Lemieux 2009. “Regression Discontinuity Designs in Economics” National Bureau of Economic Research. Working Paper 14723

Lewis, D.J., and A.J. Plantinga. 2007. “Policies for Habitat Fragmentation: Combining Econometrics with GIS-Based Landscape Simulations.” *Land Economics*, 83(2): 109-127.

Lewis, D.J. 2010. “An Economic Framework for Forecasting Land-Use and Ecosystem Change.” *Resource and Energy Economics*, 32: 98-116.

Lewis, D.J., B. Provencher, and V. Butsic. 2009. “The Dynamic Effects of Open-Space Conservation Policies on Residential Development Density.” *Journal of Environmental Economics and Management*, 57: 239-252.

Lubowski, R.N., A.J. Plantinga, and R.N. Stavins. 2006. “Land-Use Change and Carbon Sinks: Econometric Estimation of the Carbon Sequestration Supply Function.” *Journal of Environmental Economics and Management*, 51(2): 135-152.

Lubowski, R.N., A.J. Plantinga, and R.N. Stavins. 2009. "What drives land use change in the United States? A national analysis of landowner decisions." *Land Economics* 84(4): 529-550.

Ludwig, L. 2008. "The Effect of Exclusive Agricultural Zoning in Columbia County: Incorporating Zoning into Land Use Models." Unpublished M.S. thesis, Department of Agricultural and Applied Economics, University of Wisconsin Madison.

Haenszel. 1959. "Statistical aspects of the analysis of data from retrospective studies." *Journal of the National Cancer Institute* 22: 719–748.

McConnell, V., M.Walls, and E. Kopits. 2006. "Zoning, TDRs, and the Density of Development." *Journal of Urban Economics*, 59: 440-457.

McCrary, Justin. 2007. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test" NBER Technical Working Paper No. 334

McFarlane,D. 2008. "Identifying parcelization patterns in three rural Wisconsin towns – Columbia County project summary" available from the University of Wisconsin Stevens Point –Center for Land Use Education http://www.uwsp.edu/cnr/landcenter/parcelizationstudy/columbia_final_report.pdf

McFarlane, D. and K. Rice. "Integrating landscape amenities with historic parcelization trends to better target landscape conservation efforts." Presented at the 48th annual conference of the American Collegiate Schools of Planning. October 18-21, 2007.

McMillen, D.P., and J.F. McDonald. 1989. "Selectivity Bias in Urban Land Value Functions." *Land Economics*, 65: 341-352.

McMillen, D.P. and J.F. McDonald. 1991. "A Simultaneous Equations Model of Zoning and Land Values." *Regional Science and Urban Economics*, 21: 55-72.

McMillen, D.P., and J.F. McDonald. 2002. “Land Values in a Newly Zoned City.” *Review of Economics and Statistics*, 84: 62-72.

Newburn, D.A., and P. Berck. 2006. “Modeling Suburban and Rural-Residential Development Beyond the Urban Fringe.” *Land Economics*, 82(4): 481-499.

Nichols, A. 2007. “rd: Stata module for regression discontinuity estimation.” http://ideas.repec.org/c/boc/bocode/s456888.html

Rosenbaum, P. R., and D. B. Rubin,, 1983a, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika* 70: 41–55.

Rosenbaum, P.R., D.B.Rubin 1983b. “Assessing the Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome.” *Journal of the Royal Statistical Society Series B(Methodological)* 45(2):212:218

State of Wisconsin. 2007. “Wisconsin farmland preservation credit: For use in preparing 2007 claims.” Publication 503.

Towe, C.A., C.J. Nickerson, and N. Bockstael. 2008. “An Empirical Examination of the Timing of Land Conversions in the Presence of Farmland Preservation Programs.” *American Journal of Agricultural Economics*, 90(3): 613-626.

Wallace, N.E. 1988. “The Market Effects of Zoning Undeveloped Land: Does Zoning Follow the Market?” *Journal of Urban Economics*, 23: 3

Table1 Description of variables and Summary statistics by policy

		EAZ < 35 acres n=1923		Non-EAZ <35 acres n = 1411		EAZ>35 acres n= 2047		Non-EAZ > 35 acres n=110	
Variable	Description	Mean	Std. Dev.	Mean	Std. Dev	Mean	Std. Dev.	Mean	Std. Dev.
Acres	GIS calculated size of parcel (hundred acres)	0.14	.977	0.58	.728	0.45	1.334	0.46	1.341
Slope	Average parcel slope (percent*100)	8.45	6.89	8.84	7.11	8.23	5.31	8.07	5.40
% Crop	Percentage of parcel cropped or tilled	46.61	40.68	22.30	35.36	63.13	34.53	50.70	32.53
% Past	Percentage of parcel in pasture	12.95	24.27	21.82	34.22	9.20	16.45	16.06	17.46
% Forest	Percentage of parcel in forest	30.69	36.67	27.74	36.66	26.01	31.79	29.93	33.21
% Water	Percentage of parcel in water	0.66	5.35	0.01	0.30	0.24	1.95	0.18	1.06
Services	Parcel is within public service district (0 - no, 1 - yes)	0.04	0.19	0.01	0.34	0.00	0.18	0.01	0.31
Servdist	Distance from parcel edge to service district boundary (ten miles)	1.49	1.17	1.36	1.51	1.61	1.21	0.10	0.12
Lodidist	Distance to the town of Lodi (ten miles)	0.22	0.10	0.22	0.14	0.23	0.11	0.17	0.13
Waterdist	Distance from parcel edge to water (ten miles)	7.01	5.25	5.14	6.45	7.13	5.51	9.01	9.19
Road	Parcel adjacent to state/federal highway (0 - no, 1 - yes)	0.07	0.47	0.09	0.35	0.06	0.49	0.07	0.44
Schools	Travel time to nearest school (ten minutes)	0.72	2.36	0.63	3.04	0.68	2.38	0.52	2.57
EAZ	Parcel zoned Exclusive Agriculture	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Split	Parcel subdivides (1,0)	0.04	0.20	0.07	0.25	0.06	0.25	0.25	0.43
Large	Parcel > 35 acres	0	0	0	0	1	0	1	0
FPP	Eligible for farmland preservation program	0	0	0	0	1	0	0	0
Developed	Parcel has structure (1,0)	0.28	0.45	0.47	0.49	0.26	0.44	0.53	0.50

Table 2. Comparison of econometric methods

Model	Treatment effect	Selection Bias Assumptions	Selection Bias Correction	Other assumptions/concerns
Binary logit model of the subdivision decision	ATE of EAZ and FPP eligibility on likelihood of parcel to subdivide over entire dataset	no unobserved correlation between zoning and subdivision	None	Functional form assumptions and model specification may strongly influence the coefficient estimates along with st. errors (LaLonde, 1986).
FIML model of the joint zoning and subdivision decisions ; probit – logit specification	ATE of EAZ and FPP eligibility on likelihood of a parcel to subdivide over entire dataset	unobserved correlation between zoning and subdivision decision	Correlated unobservables across the probit and logit models. The parameter ρ is the correlation coefficient.	Functional form assumptions and model specification may strongly influence the coefficient estimates along with st. errors (LaLonde, 1986).
Propensity score matching of the subdivision decision.	ATT on EAZ parcels that are not eligible for FPP	correlation between zoning and subdivision decisions due to observable factors only	Matching on the underlying covariates such that: $EAZ \perp C \mid p(C)$ and $S_0, S_1 \perp EAZ \mid p(C)$	Only corrects for selection bias resulting from differences in observable covariate distributions. Cannot correct for "hidden bias"
Semi-parametric regression discontinuity of the subdivision decision	ATT at the discontinuity, given the parcel is zoned EAZ	Eligibility for FPP is non-random and is based on the size of the parcel	In the neighborhood around 35 acres the assignment of FPP eligibility is quasi-random The conditional regression functions describing the subdivision decision are continuous in acres at the discontinuity	Does not correct for selection into EAZ Estimate is only valid in the area around 35 acres unless one assumes a homogenous treatment effect Local linear regression is used on a discrete dependent variable Estimate sensitive to bandwidth
Fully parametric regression discontinuity of the subdivision decision	ATT at the discontinuity, given the parcel is zoned EAZ	Eligibility for FPP is non-random and is based on the size of the parcel	In the neighborhood around 35 acres the assignment of FPP eligibility is quasi-random The conditional regression functions describing the subdivision decision are continuous in acres at the discontinuity	Does not correct for selection into EAZ Dummy variable may simply be picking up some non-linearity in acres Sensitive to what range around 35 is included in estimation

Table 3. FIML, probit, and logit results for data from 1972-2005. Dependent variable in the probit model is selection into EAZ. Dependent variable in the logit model is whether a subdivision happens.

FIML Probit	Coef	Std. Err.	t-value	Binary Probit	Coef	Std. Err.	t-value
Intercept	-32.82*	5.60	-5.86		-2.48*	0.20	-12.36
Slope	-5.88	8.57	-0.69		0.01	0.01	1.54
% Crop	13.53*	2.32	5.83		0.02*	0.002	9.63
% Past	7.39*	1.80	4.10		0.01*	0.002	5.05
% Water	131.69*	55.03	2.39		0.08*	0.02	3.55
% Forest	7.82*	1.97	3.97		8.21E-03*	0.002	2.79
Watdist	16.55*	3.42	4.84		7.62E-02*	2.15E-02	5.70
Schools	29.53*	5.44	5.43		0.15*	0.02	8.89
Road	-4.43*	1.78	-2.48		-0.22	0.18	-1.27
d72	1.40*	0.39	3.60		-0.02	0.05	-0.43
d83	0.84*	0.34	2.46		-0.07*	0.04	-1.95
d91	0.54**	0.31	1.74		0.04	0.03	1.38
Acres	46.98*	8.16	5.75		0.03*	0.003	10.09
FIML Logit				Binary Logit			
Intercept	-3.21*	0.54	-5.94		-2.89*	0.45	-6.48
Slope	-1.32	1.25	-1.06		-0.01	0.01	-1.07
% Crop	-0.09	0.38	-0.24		-1.22E-04	2.58E-03	-0.05
% Past	0.30	0.38	0.79		4.38E-03	2.53E-03	-1.73
% Water	1.60	1.49	1.07		0.02*	0.01	2.09
% Forest	0.55	0.39	1.43		-3.42E-05	2.55E-03	-.01
Watdist	-2.45*	0.98	-2.50		-8.61E-05*	3.13E-05	-2.75
Watdist^2	1.08	1.21	0.89		2.19E-09*	1.43E-09	1.54
Schools	-1.30	1.63	-0.80		0.02	0.12	0.20
Schools^2	0.84	1.38	0.61		2.58E-03	7.99E-03	0.32
Road	0.36	0.25	1.42		0.23	0.23	0.99
Dummy72	0.22	0.18	1.23		0.32*	0.16	2.06
Dummy83	-0.30	0.19	-1.63		-0.26	0.17	-1.57
Dummy 91	0.12	0.17	0.72		0.18	0.16	1.19
Acres	1.80*	0.51	3.57		0.03*	0.005	7.54
EAZ	0.96**	0.55	1.76		-0.82*	0.18	-4.48
Large	0.01	0.35	0.03		0.16	0.30	0.53
FPP	-0.59**	0.33	-1.81		-0.64*	0.29	-2.23
$\rho/\sqrt{1-\rho^2}$	-1.0875*	0.3326	-3.2697				
σ	19.54*	3.28	5.95				

n=5764 * denotes significance at 5% level ** denotes significance at the 10%level

Note: Coefficients in the FIML probit model are normalized by $\sqrt{1-\rho^2}$.

Table 4. Results from EAZ selection equation for panel data from 2001-2005. Dependent variable is EAZ

Probit model	Coef.	Std. Err.	z
Intercept	-6.884*	0.447	-15.390
developed	-0.199	0.129	-1.540
acres	0.098*	0.022	4.420
acres2	-0.002*	0.001	-2.590
% Crop	0.009*	0.003	3.520
% Past	0.007*	0.003	2.410
% Forest	0.003	0.003	1.110
% Water	0.029	0.025	1.150
Servdist	-1.10E-04*	2.87E-05	-3.820
Servdist^2	9.55E-09*	1.58E-09	6.040
Lodidist	0.001*	5.390E-05	10.410
Lodidist^2	-1.55E-08*	1.78E-09	-8.690
Watdist	4.15E-04*	3.38E-05	12.290
Watdist^2	-1.35E-08*	1.38E-09	-9.730

n=1109 * denotes significance at 5% level ** denotes significance at the 10% level

Pseudo R^2=.4984

Table 5. Propensity score matching results; the effect of EAZ on parcels zoned EAZ but not eligible for FPP.

Year	Matching Method	Coefficient	Std. Err.	t-stat
1972 n=614	Radius	-0.05	0.06	-0.88
	Kernel	-0.03	0.06	-0.55
	Nearest Neighbor	-0.02	0.05	-0.34
1983 n=834	Radius	-0.04	0.03	-1.18
	Kernel	-0.07	0.05	-1.24
	Nearest Neighbor	-0.02	0.03	-0.55
1991 n=845	Radius	-0.01	0.02	-0.63
	Kernel	0.00	0.01	-0.16
	Nearest Neighbor	-0.07*	0.03	-2.41
2001 n=1041	Radius	-0.03	0.02	-1.62
	Kernel	-0.03	0.02	-1.10
	Nearest Neighbor	-0.04**	0.02	-1.86

* denotes significance at the 5% level ** denotes significance at the 10% level

Table 6. Full estimation results from probit discontinuity model for parcels between 25-45 acres and zoned EAZ. Marginal effects are reported, discrete change effects are reported for binary variables.

variable	dy/dx	Std. Err.	z
Slope	0.0007	0.0011	0.61
% Crop	-0.0051*	0.0015	-3.42
% Past	-0.0050*	0.0014	-3.39
% Forest	-0.0061*	0.0022	-2.72
% Water	-0.0046	0.0014	-3.09
Watdist	-1.36E-06	0	-0.44
Watdist^2	-4.14E-11	0	-0.28
Schools	0.0126	0.0143	0.88
Schools^2	-0.0004	0.001	-0.42
Road	0.0277	0.0301	0.92
Dummy 72	0.0248	0.0165	1.51
Dummy 83	-0.0067	0.0146	-0.46
Dummy 91	0.0001	0.0154	0.06
Acres	0.0005	0.0033	0.16
>35	-0.0593	0.0520	-1.14

n=1986 * denotes significance at 5% level ** denotes significance at the 10% level

Table 7. Estimated Regression Discontinuity Results (marginal effects reported for probit model)

Years	Estimator	Bandwidth	Coefficient	Std. Err	Z
1972-2005	Local-linear regression	3.37	-0.12**	0.07	-1.78
	Local-linear regression	5.60	-0.08	0.06	-1.41
	Probit Model n=1986	Parcels from 25-45 acres	-0.06	0.05	-1.11
	Probit Model n= 901	Parcels from 30-40 acres	-0.05	0.054	-0.83
1983-2005	Local-linear regression	3.37	-0.18**	0.09	-2.03
	Local-linear regression	5.59	-0.14*	0.07	-2.16
	Probit Model n=1472	Parcels from 25-45 acres	-0.09	0.06	-1.29
	Probit Model n= 679	Parcels from 30-40 acres	-0.12	0.095	-1.34

* denotes significance at 5% level ** denotes significance at the 10%level

Figure Captions.

Figure 1. Lodi and Westport townships in Columbia County, WI.

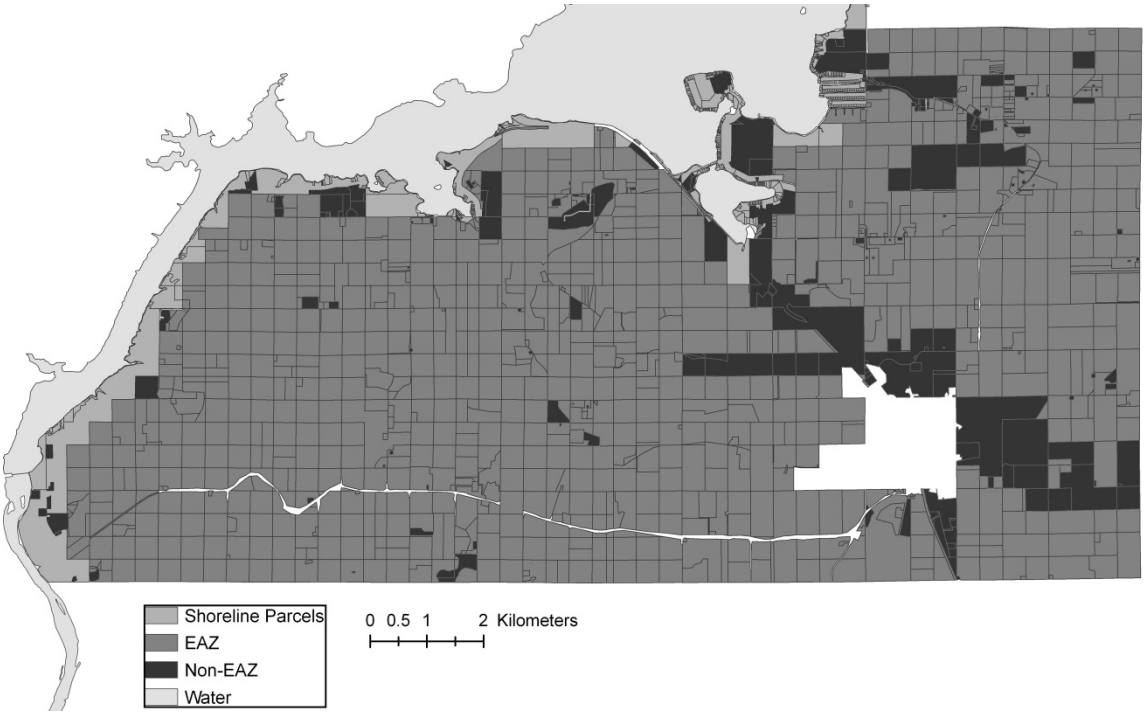


Figure 2. Discrete change effects of EAZ estimates (bands indicate confidence intervals).

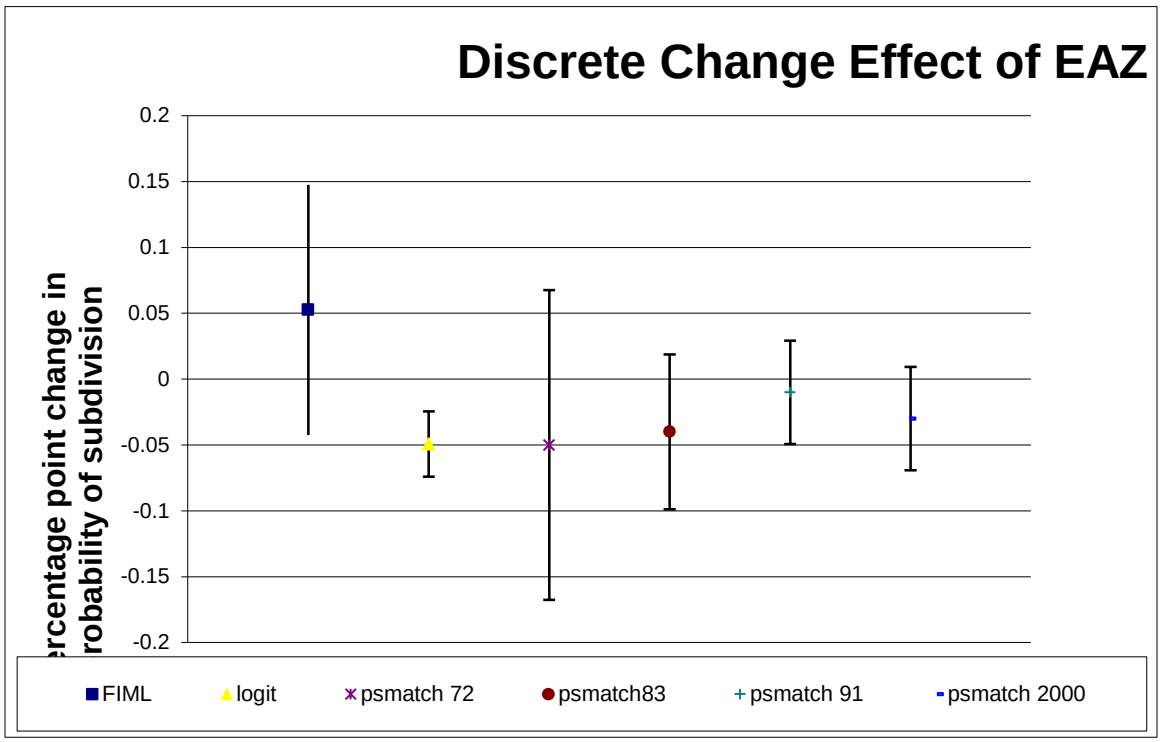


Figure 3. Discrete change effects of FPP eligibility estimates(bands indicate confidence intervals).

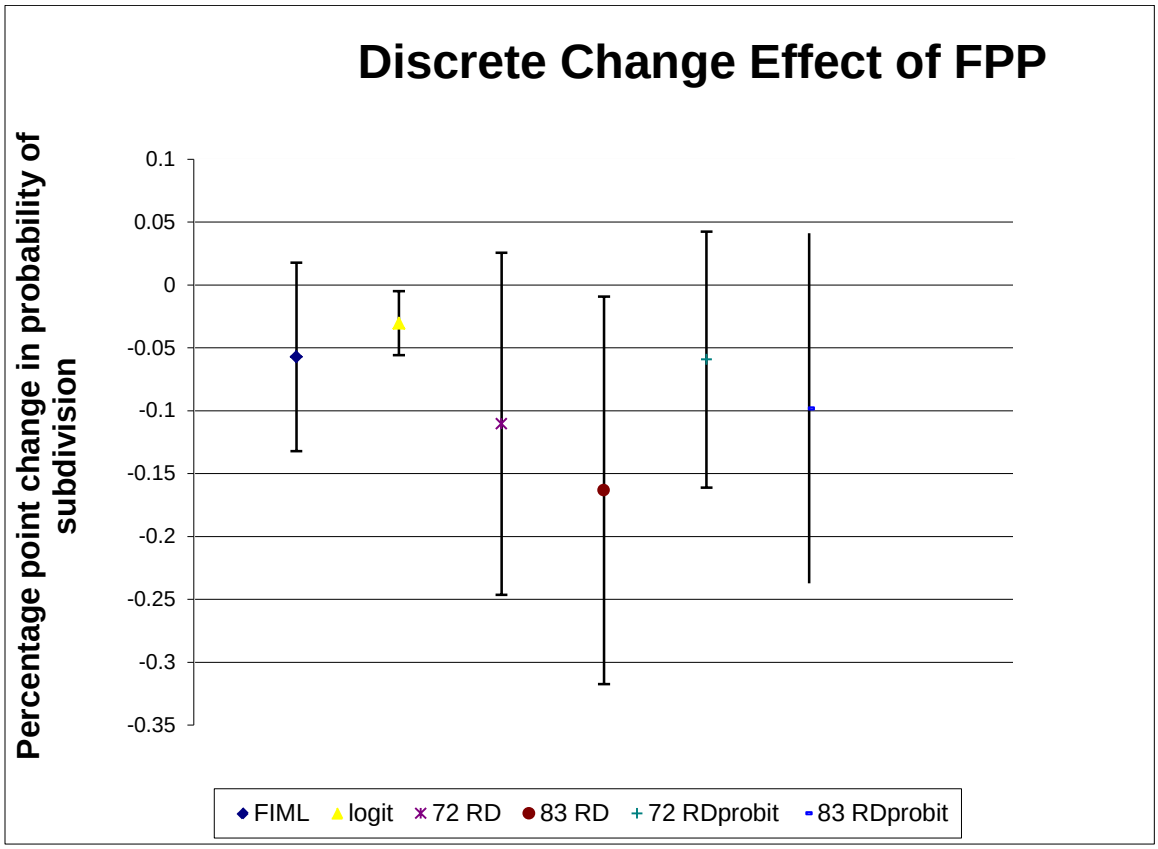
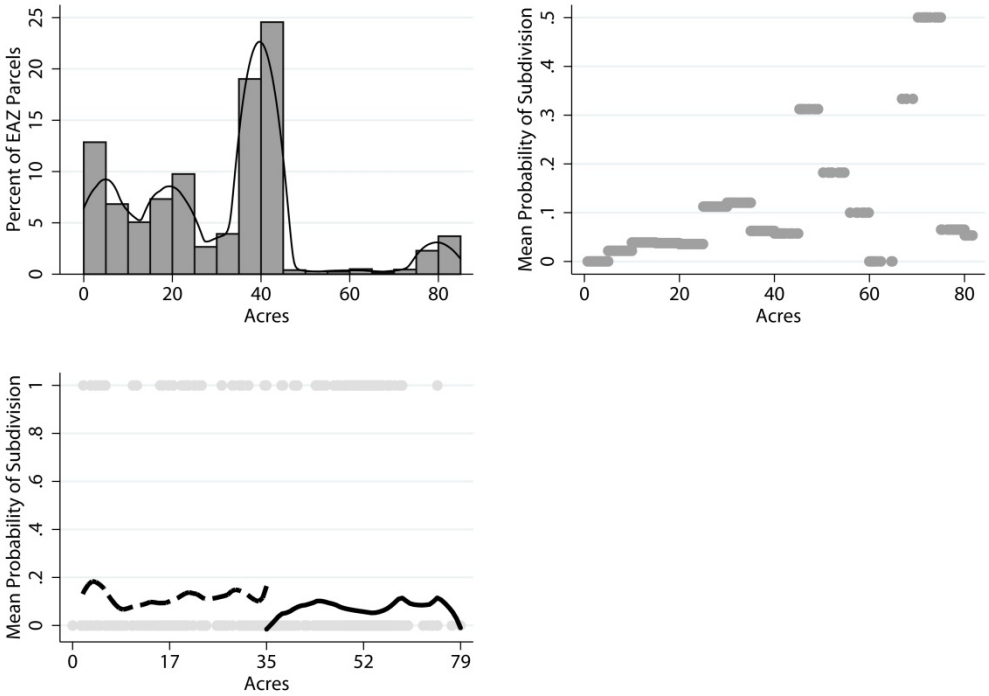


Figure 4. Regression discontinuity summary graphs. From top left to right 1. Number of parcels in each 5 acre bin. 2. Mean probability of subdivision within each bin and number of observations. 3. Kernel estimation of mean probability of subdivision on each side of the discontinuity (bandwidth = 3.37).



[1] One exception in the land-use change literature is the analysis of Bento et al. (2007), who use matching methods to estimate the effects of development moratoria on land-use change in a selection on observables analysis.
[2]Eliminating waterfront parcels reduces the dataset by 315 observations. All econometric models were run with the full dataset and with the restricted dataset – the results are not sensitive to the exclusion of waterfront parcels. Results from the full dataset are available upon request.
[3] In addition to the variables used in the estimation, many other geographic variables – distance to Madison, a township dummy, distance to public open space, alternative road measures – among others were created but found to not influence to the likelihood of zoning or subdivision and were thus left out the final estimated equations.
[4] Since EAZ when into effect in 1973 and FPP in 1977, we also estimate the models with just the 1983-2005 period. Parameter estimates using just the 1983-2005 period are available from the authors upon request, and the relevant estimates are not affected by which time period is used.
[5] Standard errors for the discrete change effects for the FIML model are estimated using the Krinksky-Robb method.
[6] Results of the selection equations are available upon request from the authors.

[7] Full Mantel-Haenszel statistics are available from the authors upon request.
[8] We employ the method suggested by Lee and Lemieux (2009) to choose a bandwidth of .75 acres.
[9] The use of a linear probability model when faced with discrete data is less than ideal. To check the robustness of using this model, we compare the discrete change effects of a probit model with those from a linear model. The marginal effects are nearly identical between the two models, hinting that in this case the use of the linear probability model on a discrete dependent variable is not problematic.
[10] We also run the regression with an interaction term FPP*acres, as suggested by Lee and Lemieux(2009) the results do not change qualitatively.
[11] A probit model with panel-robust standard errors was run over the range of data from 34-36 acres, 33-37 acres, 32-38 acres, 31-39 acres along with 30-40 acres and 25-45 acres.
[12] As an additional robustness check, we tests for a discontinuity at 25 acres. The results of the local linear and probit models are both null.