Appendix A. Example derivations for the conceptual framework

General solutions of the conceptual framework are not quite intuitive. We look at some special cases here for illustrative purposes.

Case 1 – A Special Case

We consider an extrem case where farmers only derive utility from production and non-residents' production function is flat, which can be characterized by: $U_2^F = 0, \frac{\partial f_A^R}{\partial l_{Ai}} = 0$, and $\frac{\partial f_H^R}{\partial l_{Hi}} = 0$. Then, farmers' utility maximizing decision can be formulated as:

$$p_A \frac{\partial f_A}{\partial l_{Ai}} - C_A(S_i) + t_A = p_H \frac{\partial f_H}{\partial l_{Hi}} - C_H(S_i) + t_H; l_{Ai} + l_{Hi} = l_i. \quad \text{(A-1)}$$

This is essentially an equalization of marginal profits, suggesting that this decision is also maximize farming profit. If the production function is linear in acreage, the farmer will use all his disposable land for whatever crop type that generates the highest profit (which is usually not hay). Though less extreme cases would result in proportional land segmentation, the general story holds that: the optimal l_{Ai}^* and l_{Hi}^* for a farmer are decided subject to a set of exogenous variables: $\{p_A, p_H, t_A, t_H, C_A(S_i), C_H(S_i), S_i\}$. Considering that prices, costs, and tax credits may change with time (but the changes are similar for all farms in the study location), a proper analysis of proportional land use would have parcel specific attributes addressing the time-invariant part of the farmers' land cover management strategies and time fixed effects absorbing price shocks.

Non-resident owners' optimal decision can be formulated as:

$$D_i t_A - C_A(S_i) + \frac{U_2}{U_1} p_x \frac{\partial V}{\partial l_{Ai}} = D_i t_H - C_H(S_i) + \frac{U_2}{U_1} p_x \frac{\partial V}{\partial l_{Hi}}; l_{Ai} + l_{Hi} = l_i, \quad \text{(A-2)}$$

which is a problem of the trade-off between the benefits (tax credits plus aesthetic values) and costs of different land uses. Naturally, compared with other agricultural land uses, hay brings higher aesthetic value ($\frac{\partial V}{\partial l_{Hi}} > \frac{\partial V}{\partial l_{Ai}}$), and costs much less ($C_H(S_i) < C_A(S_i)$). The optimal l_{Ai}^{**} and l_{Hi}^{**} for a non-resident are decided subject to a set of exogenous factors $\{t_A, t_H, C_A(S_i), C_H(S_i), S_i\}$, the criteria for D_i, and the functional form of V. If the aesthetic value function V is linear in acreage, the corner solution for the maximization problem is reached and the non-residents will use all the disposable land for the dominant option (Hay). Though it is not definite for all individual parcels, it is most likely that $l_{Ai}^{**} \neq l_{Ai}^{*}$ and $l_{Hi}^{**} \neq l_{Hi}^{*}$.

Now we consider whether non-residents meet the requirement of agricultural tax assessment. Most legal criteria for agricultural tax assessment require a minimum level of agricultural revenue (R):

$$D_i = 1, if\ p_A f_A(l_{Ai}; S_i) + p_H f_H(l_{Hi}; S_i) \geq R. \qquad \text{(A-3)}$$

If the required revenue level is chosen carefully, an extreme situation can be reached, where R equals $\underline{R}_{it}$, the lowest possible revenue level of a certain parcel, and non-residents can always meet the requirement ($D_i \equiv 1$). But this situation is nearly impossible since it requires a heterogeneous regulation (essentially, the regulation achieving this goal would require a specific revenue level for every operation).

Under a uniform revenue requirement R, some non-residents will not be able to meet the criteria. However, the requirement can be reached by *renting* the farmland to an operating farmer.

The lessee farmer just needs to certify that the rental land is an active part of his agricultural enterprise, and the lessee farm has over a minimum threshold of revenue. Given the non-farming-profit nature of non-resident owners, the rental will be minimal in exchange for a preferred land cover plan identified by the following equation:

$$t_A + \frac{U_2}{U_1} p_x \frac{\partial V}{\partial l_{Ai}} = t_H + \frac{U_2}{U_1} p_x \frac{\partial V}{\partial l_{Hi}}; \; l_{Ai} + l_{Hi} = l_i. \tag{A-4}$$

Given the complexity of the negotiation process, some non-residents may not follow this rental plan and fail to meet the requirement. Their land-use decision will be:

$$-C_A(S_i) + \frac{U_2}{U_1} p_x \frac{\partial V}{\partial l_{Ai}} = -C_H(S_i) + \frac{U_2}{U_1} p_x \frac{\partial V}{\partial l_{Hi}}; \; l_{Ai} + l_{Hi} = l_i. \tag{A-5}$$

The average real-world situation should be a mixture of the two cases identified by Eq. (A-4) and Eq. (A-5). The optimal land use for non-residents l_i^{**} would almost surely depart from the farmer's optimal land use l_i^*, where hay is likely to be increasingly operated to improve the aesthetic value and reduce the costs to meet the agricultural assessment requirement.

Case 2 – A More General Case

We consider a more general situation, where farmers also derive utility from amenities and non-residents' production function is not flat. Specifically, the case can be characterized as: $V(l_{Ai}, l_{Hi}) = \beta l_{Ai} l_{Hi}$, $f^F(l_{Ai}, l_{Ai}) = \alpha^F l_{Ai} l_{Hi}$, $f^R(l_{Ai}, l_{Ai}) = \alpha^R l_{Ai} l_{Hi}$, and $\alpha^F > \alpha^R$. Note that the constraint on α reflects the difference in the production function of farmers and non-residents: nonresidents having relatively limited resource (i.e., time, equipment, or skills) to increase crop productivity, and even if the land is leased to a farmer, the operation of the leasing farmer is constrained by special conditions of the contract, where the nonresidents are likely to retain a certain degree of amenities and prohibit certain forms of operation.

Farmers' utility maximization can then be characterized by

$$p_A\alpha^F l_{Hi} - C_A(S_i) + D_i t_A + \frac{U_2^F}{U_1^F} p_x \beta l_{Hi} = p_H \alpha^F l_{Ai} - C_H(S_i) + D_i t_H + \frac{U_2^F}{U_1^F} p_x \beta l_{Ai},$$

$$l_{Ai} + l_{Hi} = l_i, \tag{A-6}$$

and non-residents' utility maximization can be characterized by

$$p_A\alpha^R l_{Hi} - C_A(S_i) + D_i t_A + \frac{U_2^R}{U_1^R} p_x \beta l_{Hi} = p_H \alpha^R l_{Ai} - C_H(S_i) + D_i t_H + \frac{U_2^R}{U_1^R} p_x \beta l_{Ai},$$

$$l_{Ai} + l_{Hi} = l_i. \tag{A-7}$$

And thus, for a farmer, the optimal solution is

$$l_{Hi}^* = \frac{\left(p_H\alpha^F + \frac{U_2^F}{U_1^F} p_x\beta\right) l_i + C_A - C_H + D_i(t_H - t_A)}{(p_A + p_H)\alpha^F + 2\frac{U_2^F}{U_1^F} p_x\beta}. \tag{A-8}$$

The optimal solution for a non-resident is

$$l_{Hi}^{**} = \frac{\left(p_H\alpha^R + \frac{U_2^R}{U_1^R} p_x\beta\right) l_i + C_A - C_H + D_i(t_H - t_A)}{(p_A + p_H)\alpha^R + 2\frac{U_2^R}{U_1^R} p_x\beta}. \tag{A-9}$$

To simplify the problem, we assume $D_i = 1$ for both the farmer and the non-resident. To show the relationship between the farmer's optimal hay acreage and the non-resident's hay acreage, we calculate that $\frac{l_{Hi}^*}{l_{Hi}^{**}} = \frac{\alpha^R + \gamma_1}{\alpha^F + \gamma_2} \cdot \frac{\alpha^F + \gamma_3}{\alpha^R + \gamma_4}$, where $\gamma_1 = \frac{2\frac{U_2^R}{U_1^R} p_x\beta}{p_A + p_H}$, $\gamma_2 = \frac{2\frac{U_2^F}{U_1^F} p_x\beta}{p_A + p_H}$, $\gamma_3 = \frac{\frac{U_2^F}{U_1^F} p_x\beta + \frac{C_A - C_H + t_H - t_A}{l_i}}{p_H}$, and $\gamma_4 = \frac{\frac{U_2^R}{U_1^R} p_x\beta + \frac{C_A - C_H + t_H - t_A}{l_i}}{p_H}$.

Therefore, it is not likely that $l_{Hi}^* = l_{Hi}^{**}$. Also, it is not theoretically decided that $\frac{l_{Hi}^*}{l_{Hi}^{**}} < 1$ under the conditions specified above. Whether farmers' optimal hay acreage is less than nonresidents' optimal hay acreage is decided by a complicated formation incorporating α^F, α^R, $\frac{U_2^R}{U_1^R}$, and $\frac{U_2^F}{U_1^F}$. Even if we assume $\alpha^F > \alpha^R$ and $\frac{U_2^R}{U_1^R} > \frac{U_2^F}{U_1^F}$ (hence $\gamma_1 > \gamma_2$ and $\gamma_4 > \gamma_3$), $\frac{l_{Hi}^*}{l_{Hi}^{**}} < 1$ is still not guaranteed.

Appendix B. Assumptions required in the DID-matching estimator

The validity of the DID-matching estimator requires the following assumptions. The first assumption is *the stable unit treatment value assumption (SUTVA)*. The attributes and outcomes of one parcel are not influenced by another parcel's non-resident ownership status or change in ownership status. Farmers will manage their land to maximize agricultural profits, which depends on market prices. Since the study area belongs to a very large market including numerous farmers and farms, one classical assumption is that the market prices are not changed by the individual quantity in a competitive market, and hence in general, farmers' land cover decisions are not dependent on another parcel's change in ownership. When it comes to a non-resident owner, he may prefer to have the same land cover as his neighbors for the aesthetics brought by continuity, and thus tend to live beside another non-resident owner. But as we can see from appendix figure G2, there is little evidence suggesting clustered patterns in non-resident owners' distribution. Thus, the SUTVA assumption is satisfied in this circumstance.

The *conditional parallel trends* assumption is the DID-matching version of the unconfoundedness (or ignorable assignment) assumption. Since we are interested in an estimate of ATT, we only need a weak form of this assumption: conditional on observed attributes, if there are no non-resident property transactions, the average changes in the land cover outcome should be the same for the treated and untreated. Intuitively, this condition is to rule out cases like this: second-home buyers may select those parcels that are trending toward having more land in hay, and hence the conditional trend differences are not the exact impact of treatment (tend to be biased upwards in this setting). In practice, few buyers refer to the trends of land cover, but some issues may still exist. For example, before a property sale, farmers as sellers may target the non-residents and manage the land cover in their favor - farmers may increase the proportion in productive non-hay crops to show the land is of good quality, or increase proportion in hay if they know non-resident buyers prefer hay. In the literature, this phenomenon is called anticipation effect (Chabé-Ferret and Julie, 2013). We perform a conditional trend test (placebo test) to test this parallel trend assumption.

The last assumption is *the Common support (overlap)* assumption. To apply a DID-matching estimator, for every non-resident parcel purchase (a treated observation), there must exist a farmer-owned parcel (in the control group) having roughly the same observed attributes, and hence the same propensity score. This assumption is usually implemented loosely in empirical applications, for example, it is sometimes turned into a requirement of the same domain of the propensity scores for different groups (as in Dehejia and Wahba, 1999, 2002). To address the problem of lacking overlap, Crump et al. (2009) developed a systematic approach and suggested a simple rule of discarding propensity scores outside [.1, .9], but this strategy is not applicable in this study since it will delete too many treated observations. Huber et al. (2013), based on carefully conducted simulations, suggest that trimming observations with too much weight (or with weak support, meaning observations in the treatment group with too few corresponding observations in the control group) is important for all estimators, and that radius matching (see Dehejia and Wahba (2002)) combined with regression performs best overall.

Appendix C. Second-Home owners on farmland

We have identified the second-home counties, but to identify all the non-resident agricultural landowners in these counties is significantly harder. The large number of second-home owners from NYC makes these counties worth studying, but the second-home owners may also come from somewhere else. To be precise about the influences caused by non-resident owners on farmland, based on commuting status, we build an approach to approximately get the non-resident owners in a specific county. Taking Columbia as an example, we assign each tax bill address (zip code) into the first-home zone and second-home zone. If the tax bill of a property is sent to a zip code in the second-home zone, which is not commutable, we take it as a second home.

Table C1: Commuting destination by travel mode from Columbia County, NY

Destination	Travel mode	Workers in flow	Lower bound of flow
Kings County	Car, truck, or van: Carpooled	4	-1
Kings County	Other travel mode	7	-7
Kings County	Car, truck, or van: Drove alone	8	-6
Queens County	Car, truck, or van: Drove alone	15	1
Dutchess County	Other travel mode	20	-11
New York County	Car, truck, or van: Carpooled	49	-18
New York County	Other travel mode	140	-20
Dutchess County	Car, truck, or van: Carpooled	142	-33
New York County	Car, truck, or van: Drove alone	185	-5
New York County	Public transportation	464	76
Dutchess County	Car, truck, or van: Drove alone	1519	624

Note: The source of data "Journey to Work", see footnote n.

The Division of First-home Zone and Second-home Zone. This is a vital step to get the non-resident owners on farmland, so we want to be as precise as possible in terms of identifying a second home. In other words, we want to minimize the possibility of mistakenly including an active farmer in the non-resident group. With the commuting pattern as the first concerning confounding factor, we want to substantially limit the regular commuting possibility between areas in the second-home zone and Columbia County. Considering that a farmer may live some distance away from her farm, we set 120km as the basic commuting range to Columbia County. By setting the entirety of zip-code zones outside commuting range as the second-home zone of Columbia, we separate those farmers commuting to Columbia from non-resident owners.

If we are done here, situations existing may make the first-home zone more inclusive than it should be: tax bills sent to metropolitan areas within the commuting range may denote second-home owners. But having tax bill sent to a big city is not always a good indicator for recreational landowner since a resident can commute to a city and have his tax bills sent to the office while the farmland is still actively farmed by his family (Also, a farmer can live in a city within the commuting range and commute daily to manage his farm). To be precise about the commuting relations between nearby urban areas and Columbia, we refer to the commuting data. We find that there is no commuting flow to Columbia from nearby urban areas, so it can be safely concluded that farmers commuting from the cities within the commuting range to Columbia does not extensively exist. Moreover, with commuting data partly shown in Table A1, we are capable of identifying the commuting destinations of Columbia. There are only two destinations having significantly positive commuting flow: New York County and Dutchess County, and the travel mode of the commuting flow to New York County is by public transportation. The unique public transportation option from Columbia to New York is the Amtrak train trip, which takes two hours starting from the Hudson Station. This is not a proper option for someone who owned a farm, which is usually far away from the Hudson Station. Hence, all the urban areas except Dutchess County in the first-home range are not likely to be commuting destination for an agricultural property owner in Columbia. So, all these urban areas except those in Dutchess County are classified into the second-home zone of Columbia. As shown in appendix figure G1, a division of the zip code map between first-home zone and the second-home zone is clear now.

As stated above, if the tax bill of a property is not sent to a zip code in the first-home zone, we take it as a recreational second home. And if this property is ever used for agricultural purpose recently in the tax parcel datasets (since 1999), we take the landowner as a non-resident owner on farmland. Another issue may arise if we consider the dimension of time: the tax bill addresses may change over time. Some properties may have second-home-zone tax bills just for one year in the middle of a consecutive long-enough period, we take these cases as accidental second-home-zone tax bills and do not treat these properties as second homes. After considering all these conditions, we can define non-resident ownership versus agricultural resident ownership, as well as the ownership switch between them for each year throughout the period studied (2000-2014). For instance, we identify 201 non-residents owned agricultural properties (out of 1208) in Columba County in 2014, and they are distributed, not in a clustered pattern, throughout the agricultural districts of Columbia. The density of non-resident owners on farmland in Columbia is shown in appendix figure G2.

Through this approach, we define the agricultural properties with consecutive second-home-zone tax bills as recreational second homes on farmland, and they constitute the interested group (treatment group) of this study.

Appendix D. A preliminary regression showing the different management between farmers and non-residents

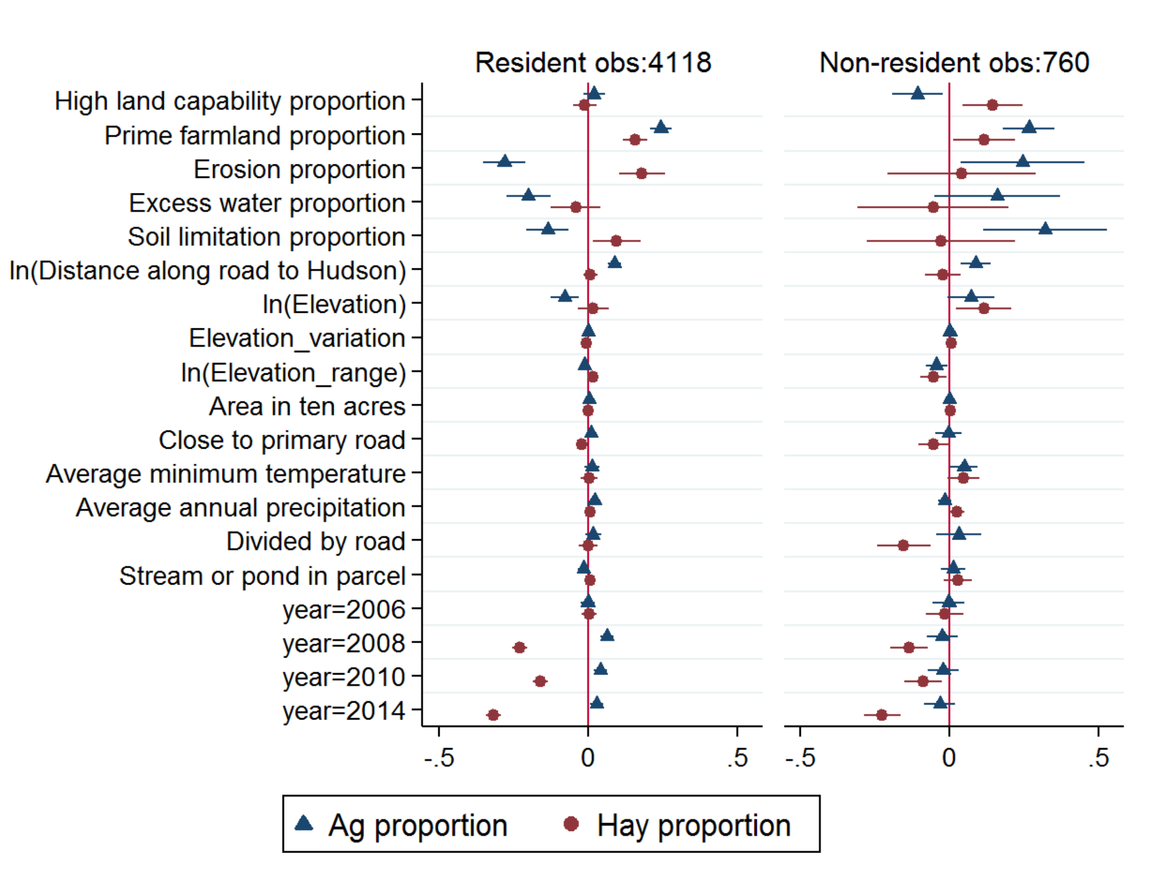


Figure D1. Different management patterns for residents and non-residents (Columbia)

Results in Figure D1 are from OLS regressions with balanced panel data, outcomes are proportion of a parcel in hay and in agricultural usage except hay. Year fixed effects are used to control price shocks. Bars represent 95% confidence interval. Constants are added but not displayed.

We can see that if the farmland is ranked with high land capability by SSURGO, a non-resident will use it for more hay and less agriculture than a farmer does. Also, residents significantly decrease their farmland in agricultural usage if their farmland is highly erodible, with excess water, or with soil limitations, while non-residents do not. More importantly, farmers tend to plant hay on lands with soil limits and erosion concerns, which confirms that hay is a low-cost-low-productivity option. Though these preliminary models are not enough to show the exact impacts of non-residents on land cover, they suggest that taking the land cover proportions as the outcome variable is applicable.

Appendix E. An alternative hypothesis test

After testing the fitness of the three conditions to our dataset, we are confident that the differences in land cover trends are caused by the second-home transactions. But are these effects brought by the fact that non-resident owners are underutilizing farmland instead of the discontinuity in management associated with the transactions or ownership changes? To rule out this alternative hypothesis, we conducted the same DID-matching approach on another group of parcels (treatment group 2), which includes all parcels that transacted from non-residents to residents from 2002 to 2010. The results are shown in Table E1, with no significant effects for all estimators across all land cover types, so it appears that the increase of farmland proportion used for hay is due primarily to the non-resident ownership and not simply the transition of ownership. Moreover, and interestingly, the impacts of treatment 2 are all to the opposite direction of treatment 1, meaning after a parcel is transacted from non-resident to resident, some hay would be turned into other agricultural use. However, the insignificance of the results suggests that the negative impact of non-residents on agriculture is generally not recoverable in a short period after a transaction in the opposite direction (non-residents to farmers).

Table E1. Alternative hypothesis test results-ATT on treatment group 2 for Columbia

Outcome: change in land cover proportion from 2001 to 2012		Treatment: transactions from non-residents to farmers[a]		
Estimator[b]	OLS coefficient	Regression Adjustment with matched sample (n=3, caliper=.01)	Regression Adjustment with matched sample (n=5, caliper=.01)	Regression Adjustment with matched sample (n=6, caliper=.01)
ATT on Ag[c] change[e]	-.0018 (.0544[d])	.0044 (.0810)	-.0016 (.0746)	.0081 (.0723)
ATT on Hay change[e]	.0027 (.0552)	.0269 (.0807)	.0181 (.0761)	-.0041 (.0741)
ATT on mismatch change[e]	.0100 (.0503)	-.0369 (.0683)	.0296 (.0638)	-.0239 (.0634)
PS-test(B/R)	-	39.1/.28	31.3/.26	23.0/.24
N treated: on support(all)	26	25(26)	25(26)	25(26)
N control: matched(all)	609	68(609)	109(609)	127(609)
N total	635	93(635)	134(635)	152(635)

Notes: a. The treatment group refers to parcels transacted from non-residents to farmers between 2002 and 2010; Control denotes farmer owner throughout the whole period. We employ propensity matching with caliper set to .01.

b. To implement the common support condition, we exclude 1 observation in treatment group 2 who does not satisfy that: each treated has at least 6 nearest neighbors in the control group within the radius .01. We carry out regression adjustment with radius matched sample (RA) on different number of near neighbors (n=3,5,6).

c. Ag denotes agricultural land cover except Hay.

d. Robust standard errors are in parentheses; * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

e. The results show us that transactions from non-residents to farmers do not have significant impacts on farmland land cover. Thus, the impact of new non-resident farmland owners is caused by the agricultural stewardship instead of the ownership change. And this impact cannot be reversed in a short period.

Appendix F. Generating A List of Second-Home Agricultural Counties

To find those urban-adjacent areas like Columbia County, we build a procedure to measure the similarity between Columbia County and each of the other counties.

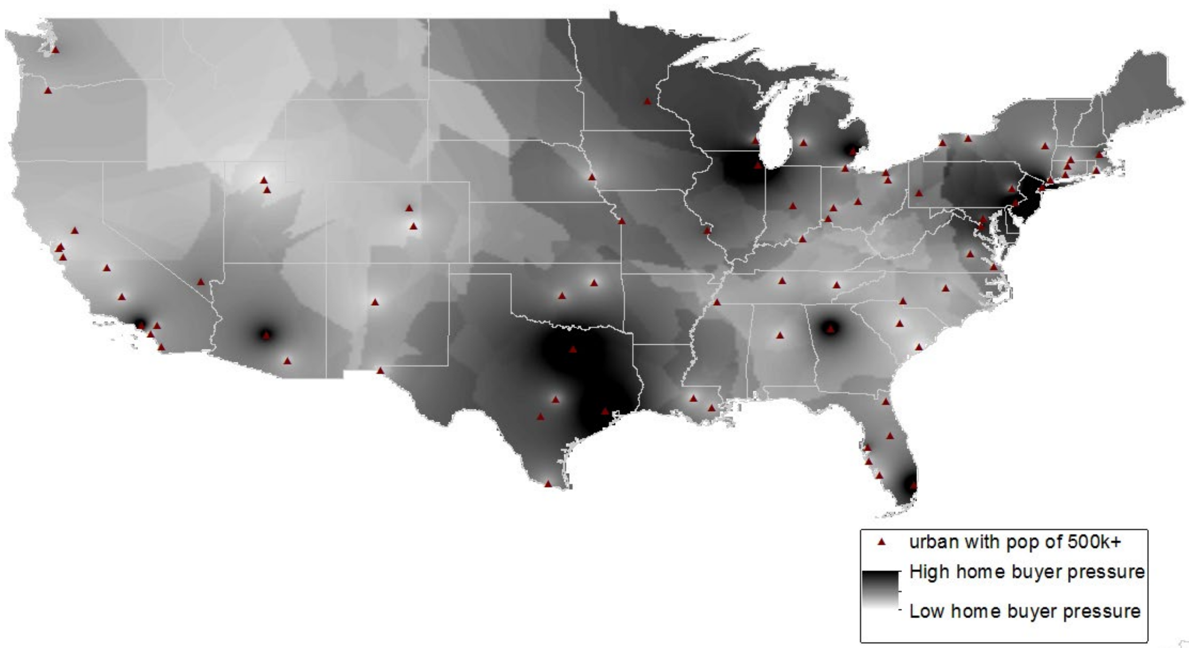


Figure F1: Homebuyer pressure of major metropolitan areas

Notes: Homebuyer pressure is derived from the urban Index of Second-Home Buyer (ISHB), and weighted by distance. The urban ISHB combines information on population, median income, and house value. Data for urban population and median income (2010) come from the Census Bureau (USCB). Data for urban median house value (2010) come from Zillow.

First, we build an index for each major urban city (a major urban city here is defined as a city with a population of more than 500k). to show their relative capacity of supplying second-home buyers, which is named Index of Second-Home Buyer (ISHB):

$$\mathrm{ISHB}_i = \alpha \cdot \frac{Pop_i \cdot W_i}{H_i},$$

where Pop_i represents the population for urban i, W_i is the five-year median income, H_i is a denominator measuring median house value, and α is a scaler with the same value for all urban areas. To give a graphical illustration of the pressure of home buyers from these cities, we conduct an inverse distance weighted (IDW) interpolation, showing in Figure F1, where the point value on the map is a weighted distance average of the point values of ISHB_i. Despite the drawbacks of IDW as a technique of interpolation (Watson and Philip, 1985), it performs fine where the source value points are densely distributed, and thus gives us a good enough numerical approximation of the second-home pressures from the nearby major urban areas for those counties not in the flyover country (West Mountain states-division 8 in USCB division), which generally isn't an area of interest for second-home buyers and for this paper.

In Figure F1, the darkest color is assigned to the major urban areas, meaning they are generating high second-home-buyer pressure to corresponding urban-adjacent areas: New York, Chicago, Houston, Dallas, Los Angeles, Miami, Detroit, Philadelphia, Phoenix, Washington, and Boston (in descending rank of ISHB). These cities are major sources of second-home buyers, who make choices, based on their preference and wealth, among places with high amenity values.

Generally, places with higher amenity value also have higher property values, as addressed by many hedonic pricing studies, and farmland is among them. However, farmland does have some differentiating characteristics, among which the most important is that policies are usually in favor of agriculture and tend to protect farmland from being converted to residential, commercial, and industrial uses. One of this kind of policy instruments is the agricultural preferential tax, from which agricultural and forested lands can gain tax reductions or exemptions based on the acreage and production. Thus, with high amenity value and preferential tax treatments, farmland in urban-adjacent areas satisfy wealthy urban dwellers' cost-effective expectation of recreational home.

Three necessary conditions to form an agricultural second-home market like Columbia County are: close enough to major sources of second-home buyers, with good amenity values attracting non-resident property buyers, and having similar agriculture with Columbia County. We collect information for all non-urban counties in the U.S. and generate variables belonging to three sets denoting these three conditions. The first set (S_1) represents the sources of non-resident owners, including distance along the highway to the closest major urban areas (among those cities with high ISHB listed before) and non-resident home buyer pressure. The second set (S_2) explains the counties' fitness to be a second-home county, including the counties' detailed information about economy (including general information about agriculture), population, geography, and houses. The third set (S_3) shows the detailed characteristics of agriculture, including the number of farms, average farm size, the acreage of cropland, average sale of farms, and the total number of operations with a direct sale to a retailer. Table F1 shows the detailed information of the variables in $S_1, S_2,$ and S_3. We exclude those counties with too short or too long distances to a major urban (We use the distribution of homeowner from NYC in appendix Figure G1 to judge the lower bound and upper bound of the tolerable distance of a second home. The distance measures here are the distance along highway, same as what's in S_1), and measure the Mahalanobis distance from all the remaining counties to Columbia County within each variable set, and expect these distances (m (S_1), m (S_2), m (S_3)) explaining how different from Columbia a county is in terms of non-resident owner sources, amenity value attracting non-resident owners, and detailed agricultural status, respectively. We take the Euclidean length[1] of $m(S_1)$, $m(S_1)$, and $m(S_1)$ to represent a county's total distance to Columbia:

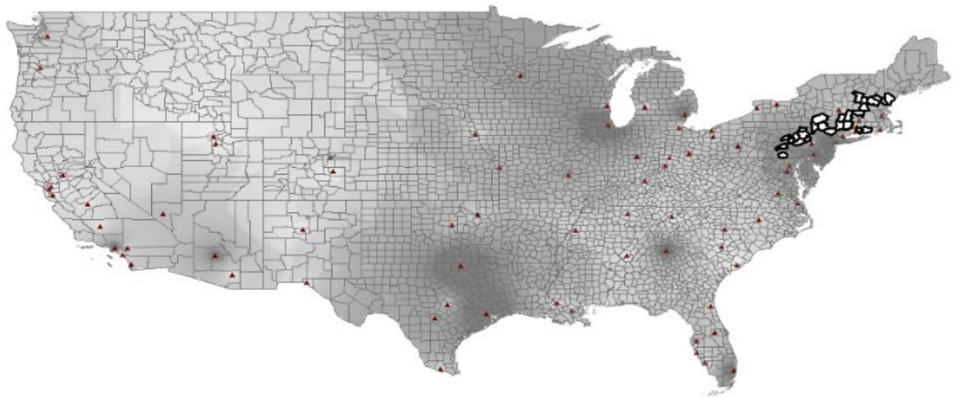
$$D(S_1, S_2, S_3) = \sqrt{[m(S_1)]^2 + [m(S_2)]^2 + [m(S_3)]^2}.$$

Finally, a ranking list of counties based on $D(S_1, S_2, S_3)$ is generated, where a higher rank suggests that the county is more likely to endure the similar impact of agricultural non-resident owners as Columbia does. Figure F2 shows the top 4%, 6%, or 8% ranked counties. We can see

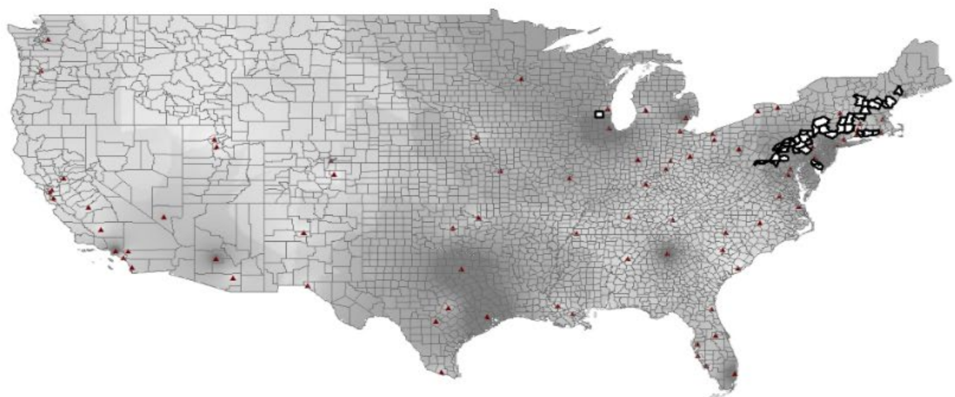
[1] To be clear, though m (S_2) and m (S_3) are obviously correlated through agricultural characters, in the analyst's view, m (S_3) measures the generalizability of the impact evaluation, which is a totally different dimension from the measure of amenity value. We take Euclidean distance because it is not only determined much by the largest difference, but also take smaller differences into consideration. Intuitively, if a county is different very much from Columbia in one dimension, it is very unlikely to have the similar impacts. However, if two counties have the same maximum difference, smaller difference matters in terms of ranking. We also test other total distance calculations like taking maximum or taking geometric mean, which yield similar orders.

the counties that are most likely to endure the same impact from non-resident owners as Columbia are generally distributed in northeastern America (Pennsylvania, New York, Massachusetts, Vermont, and New Hampshire), with Delaware and Sullivan ranked 2nd and 10th among them, respectively. When a less restrictive criterion (8%) is required, some counties in Maryland, Virginia, and Illinois will be included, and Ulster is ranked similarly (45th) to this group of counties. These results suggest that the ranking method is generally reliable, and the impacts estimated in previous sections are potentially generalizable to a considerable number of urban-adjacent counties. To be noticed, this is not to say that there is no non-resident impact on agriculture in other parts of the country. Non-farmer farmland owners, unrestrictive agricultural assessment, and other related problems may exist across the other parts of the country, which might not be similar to Columbia County.

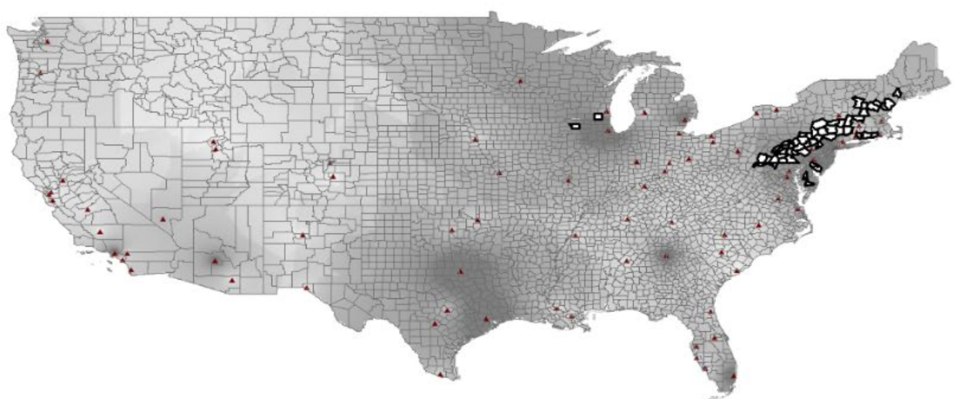
Figure F2. Counties where agricultural non-resident impacts are generalizable



Top 4% in the ranking list of counties- like-Columbia



Top 6% in the ranking list of counties- like-Columbia



Top 8% in the ranking list of counties- like-Columbia

Table F1. Variables in the calculation of total impact

	Variables	Source	Details
S_1	Distance along the highway to major urban	ArcGIS network analysis	In kilometers
	IDW index of urban home buyer pressure	ArcGIS IDW	As shown in Figure F1
S_2	Area in square miles	TIGER/Line Shapefiles	
	Average precipitation	TIGER/Line Shapefiles	Aggregated with counties in ArcGIS
	Average maximum temperature	TIGER/Line Shapefiles	Aggregated with counties in ArcGIS
	Average minimum temperature	TIGER/Line Shapefiles	Aggregated with counties in ArcGIS
	Population per square mile	Census 2012	
	White ratio	Census 2012	
	Black ratio	Census 2012	
	Median age of population	Census 2012	
	Average household size	Census 2012	
	Number of lakes	TIGER/Line Shapefiles	Counted in ArcGIS
	Having mountains	Peak counties dataset	Dummy
	Attaining mountain view	ArcGIS analysis	Categorical: -1,0,1
	Average time to work in min	Journey to work	Also exists in ACS
	Town service	ArcGIS IDW	IDW of small town population
	Medium city impact	ArcGIS IDW	IDW of city population (>30k,
	Employee	CBP 2014	During the week of March 12
	Yearly payroll per employee	CBP 2014	
	Agricultural industry employment ratio	ACS 2014	ACS 2014 displays 5-year estimation
	Whole sale employment ratio	ACS 2014	Total employee as denominator
	Retail employment ratio	ACS 2014	Total employee as denominator
	Finance employment ratio	ACS 2014	Total employee as denominator
	Retirement ratio	ACS 2014	With retirement income ratio
	With SNAP ratio	ACS 2014	
	Below poverty line family ratio	ACS 2014	
	Owner-occupied house ratio	ACS 2014	Opposite to renter occupied
	Homeowner vacancy rate	ACS 2014	
	House built between 2000 and 2009 ratio	ACS 2014	
	House with 9 or more rooms ratio	ACS 2014	Denoting large houses
	Move in from 2000 to 2009 ratio	ACS 2014	
	Move in from 2010 or later ratio	ACS 2014	
	Owner-occupied median house value	ACS 2014	In dollars
	Monthly house cost median Without	ACS 2014	Measuring cost of holding a property
S_3	Number of farms	Census 2012	
	Average size of farms	Census 2012	In acres
	Crop acreage	Census 2012	Total acreage of cropland
	Average sale of farms	Census 2012	
	Total number of operations with direct sale	USDA	Direct sale from farmers to retailers

Appendix G: Supplementary Tables and Figures

Table G1. Non-resident owners on farmland in Second-home counties of NYC

County	Ownership	2001	2002	2006	2008	2010	2014	Total
Columbia	Resident	698	696	690	685	677	672	
	Non-resident	115	117	123	128	136	141	813
	NR percent	14.1%	14.4%	15.1%	15.7%	16.7%	17.3%	
Sullivan	Resident	282	275	268	268	267	267	
	Non-resident	50	57	64	64	65	65	332
	NR percent	15.1%	17.2%	19.3%	19.3%	19.6%	19.6%	
Ulster	Resident	224	224	231	233	241	238	
	Non-resident	455	455	448	446	438	441	679
	NR percent	67.0%	67.0%	66.0%	65.7%	64.5%	64.9%	

Note: NR percent means the percentage of non-resident ownership. The statistics describe all agricultural parcels in our dataset before purging unstable parcels. Ownership between 2004 and 2010 is decided by tax bill information in the assessment data. Ownership before 2004 or after 2010 is imputed based on sales data.

Table G2. Logit model for farmland selecting into the treatment group (2002-2010 Columbia)

Dependent Variable: Farmland transferred to Non-residents (Treated)		
Variable	Coefficient	Std. Err.
Ln (Acreage)	-.5302	.3731
Ln (Assessed land value)	1.079*	.5997
Ln (Total assessed value)	-.4658	.4128
Stream or pond in parcel	.3765	.3433
Divided by road	-1.419	1.032
Close to primary road (<1km)	.4677	.3175
Ln (Distance to Hudson city)	-1.188***	.4100
Ln (Average elevation)	1.367**	.5610
Ln (Elevation range)	.4174	.3289
Annual precipitation	-.3934**	.1800
Erosion percentage	-.2697	1.323
Excess water percentage	-.7742	1.451
Soil limit percentage	-.7791	1.359
Prime farmland percentage	-.3388	.7214
Constant	-12.86	8.364
Number of observations=658	Log likelihood= -161.72	
LR chi-square (14) =25.35	p=.0312	

Note: Significance level: * p<0.1, ** p<0.05, *** p<0.01.

Table G3. Balance test for the DID matching sample (Columbia, caliper=.01, n=6)

Variables		Treated	Control (Mean)	p>|t|
Sample Size(N)	Before matching	49	609	
	After matching	46	179	
Forest ratio before treatment	Before matching	0.2742	0.2509	0.534
	After matching	0.2635	0.295	0.568
Hay ratio before treatment	Before matching	0.4291	0.474	0.29
	After matching	0.4269	0.4854	0.301
Agricultural ratio before treatment	Before matching	0.172	0.1468	0.483
	After matching	0.1791	0.1104	0.179
ln (Acreage)	Before matching	3.87	3.696	0.355
	After matching	3.858	3.823	0.879
ln (Total assessed value)	Before matching	11.78	11.61	0.397
	After matching	11.77	11.72	0.856
ln (Assessed land value)	Before matching	11.52	11.32	0.232
	After matching	11.49	11.42	0.758
Stream or pond in parcel	Before matching	0.4694	0.3941	0.301
	After matching	0.4565	0.4221	0.743
Divided by road	Before matching	0.0204	0.0805	0.127
	After matching	0.0217	0.029	0.827
Close to primary road	Before matching	0.4694	0.3842	0.24
	After matching	0.4565	0.4801	0.823
ln (Distance along roads to Hudson)	Before matching	9.657	9.764	0.196
	After matching	9.683	9.671	0.925
ln (average elevation)	Before matching	4.723	4.621	0.248
	After matching	4.722	4.723	0.994
ln (elevation range)	Before matching	3.457	3.128	.026**
	After matching	3.438	3.445	0.968
Annual precipitation	Before matching	41.87	41.74	0.557
	After matching	41.9	41.89	0.953
Excess water percentage	Before matching	0.1922	0.233	0.251
	After matching	0.1959	0.2058	0.81

Erosion percentage	Before matching	0.4437	0.3703	.100*
	After matching	0.4304	0.4714	0.534
Soil limit percentage	Before matching	0.3265	0.3637	0.44
	After matching	0.3337	0.2953	0.573
Prime farmland percentage	Before matching	0.6289	0.6468	0.663
	After matching	0.6256	0.6436	0.744
PS Matching test: B=23.8; R=1.00				

Table G4. DID-matching results: ATT for Sullivan County

Outcome: change in land cover proportion from 2001 to 2012			Treatment: transactions from farmers to non-residents[a]	
Estimators[b]	OLS coefficient	Regression Adjustment with matched sample n=3, caliper=.015	Regression Adjustment with matched sample n=5, caliper=.015	Regression Adjustment with matched sample n=6, caliper=.015
ATT on Ag[c] change	-.0972*** (.0332[d])	-.1016* (.0602)	-.1088* (.0575)	-.0945* (.0502)
ATT on Hay change	.0801* (.0409)	.0951 (.0658)	.1033 (.0650)	.0932* (.0563)
ATT on mismatch change	.1084*** (.0355)	.1169** (.0572)	.1074** (.0498)	.1100*** (.0499)
PS-test(B/R)	-	48.0/.14	44.4/.12	42.6/.13
N treated: on support(all)	24	20(24)	20(24)	20(24)
N control: matched(all)	253	50(253)	72(253)	83(253)
N total	277	70(277)	92(277)	103(277)

Notes: a. Treatment: farmland transacted from residents to non-residents between 2002 and 2010; Control: farmland owned by residents throughout the study period. (Sullivan) We employ propensity matching with caliper set to .015.

b. To implement the common support condition, we exclude 4 observations in the treatment group who does not satisfy that: each treated has at least 6 nearest neighbors in the control group within the radius .015. We carry out regression adjustment with a radius matched sample (RA) on a different number of near neighbors (n=3,5,6).

c. Ag denotes agricultural land cover except Hay.

d. Robust (for RAs) standard errors are in parentheses; * p<0.1, ** p<0.05, *** p<0.01.

Table G5. OLS Results for Columbia County

Outcome: change in land cover proportion from 2001 to 2012			
	Ag	Hay	mismatch
Treated [a]	-0.0896**	0.0724*	0.0479
	(0.0411[b])	(0.0420)	(0.0381)
Stream or pond in parcel	0.00568	-0.0229	-0.0320
	(0.0243)	(0.0249)	(0.0226)
Divided by road	-0.0521	0.0356	0.0262
	(0.0405)	(0.0415)	(0.0376)
Ln (Acreage)	-0.0780***	0.117***	0.0868***
	(0.0267)	(0.0273)	(0.0247)
Ln (Assessed land value)	0.0977**	-0.0929**	-0.0761**
	(0.0392)	(0.0401)	(0.0364)
Ln (Total assessed value)	0.00183	-0.0304	-0.0135
	(0.0246)	(0.0251)	(0.0228)
Ln (Distance to Hudson city)	0.0576**	-0.0337	-0.0539**
	(0.0286)	(0.0292)	(0.0265)
Close to primary road (<1km)	-0.0354	0.0495**	0.0204
	(0.0226)	(0.0231)	(0.0209)
Ln (Average elevation)	-0.0148	0.0580	0.0895***
	(0.0371)	(0.0380)	(0.0344)
Ln (Elevation range)	-0.0555***	0.0260	0.0227
	(0.0211)	(0.0215)	(0.0195)
Annual precipitation	0.0171	-0.0197	-0.0250**
	(0.0121)	(0.0124)	(0.0112)
Erosion percentage	0.142	-0.0282	-0.00331
	(0.0990)	(0.101)	(0.0919)
Excess water percentage	-0.0645	0.172	0.149
	(0.104)	(0.106)	(0.0965)
Soil limit percentage	0.0245	0.0502	0.0158
	(0.0999)	(0.102)	(0.0928)
Prime farmland percentage	-0.0169	-0.0346	-0.0750*
	(0.0484)	(0.0495)	(0.0449)
N	658	658	658

Notes: a. Treatment: farmland transacted from residents to non-residents between 2002 and 2010; Control: farmland owned by residents throughout the study period. (Sullivan) We employ propensity matching with caliper set to .015.

b. Standard errors are in parentheses; * p<0.1, ** p<0.05, *** p<0.01.

Figure G1: geographic attributes for the second counties of New York City

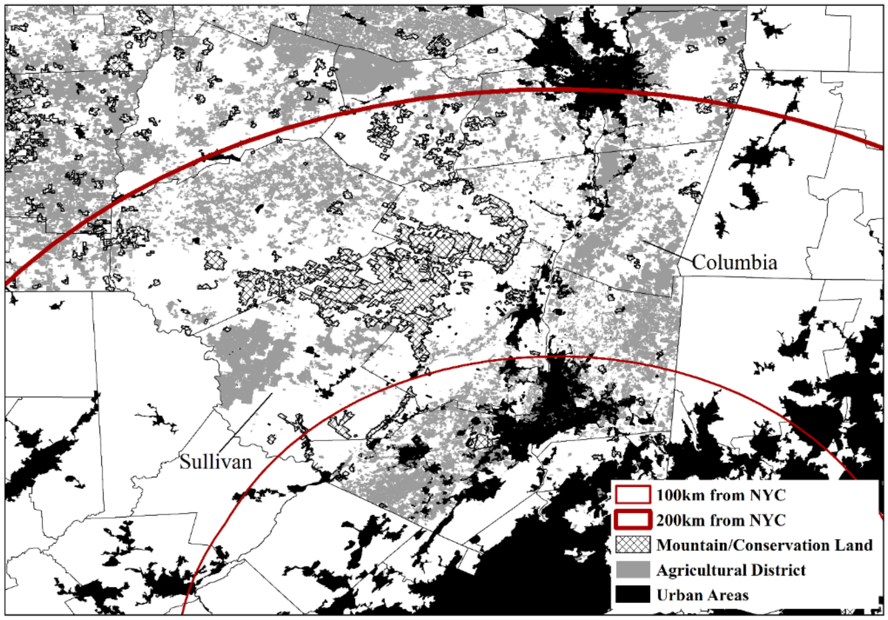


Figure G2: First-home and Second-home zip code zone for Columbia County

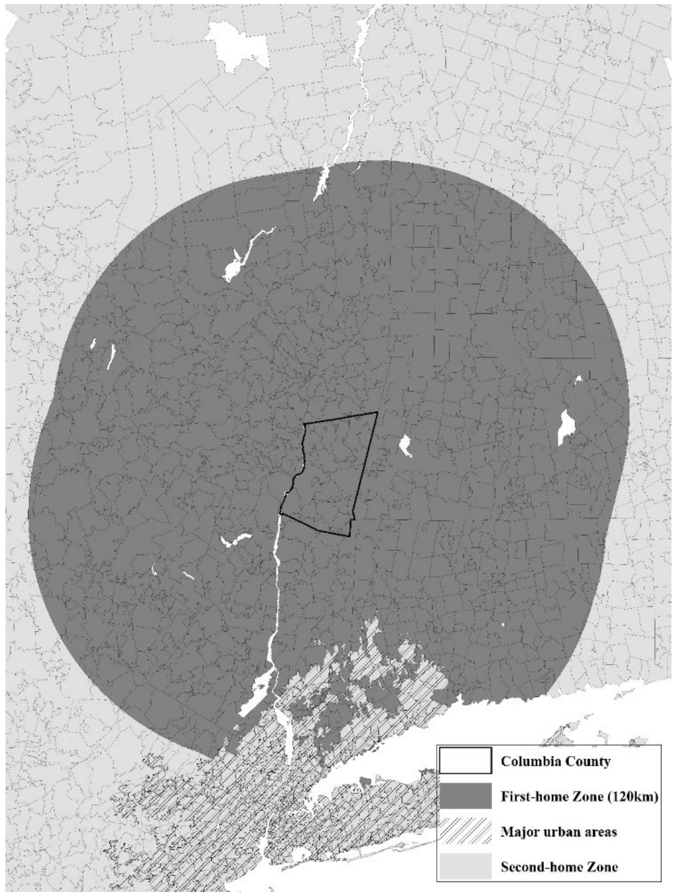


Figure G3: Distribution of Non-resident owners on farmland in Columbia County (2014)

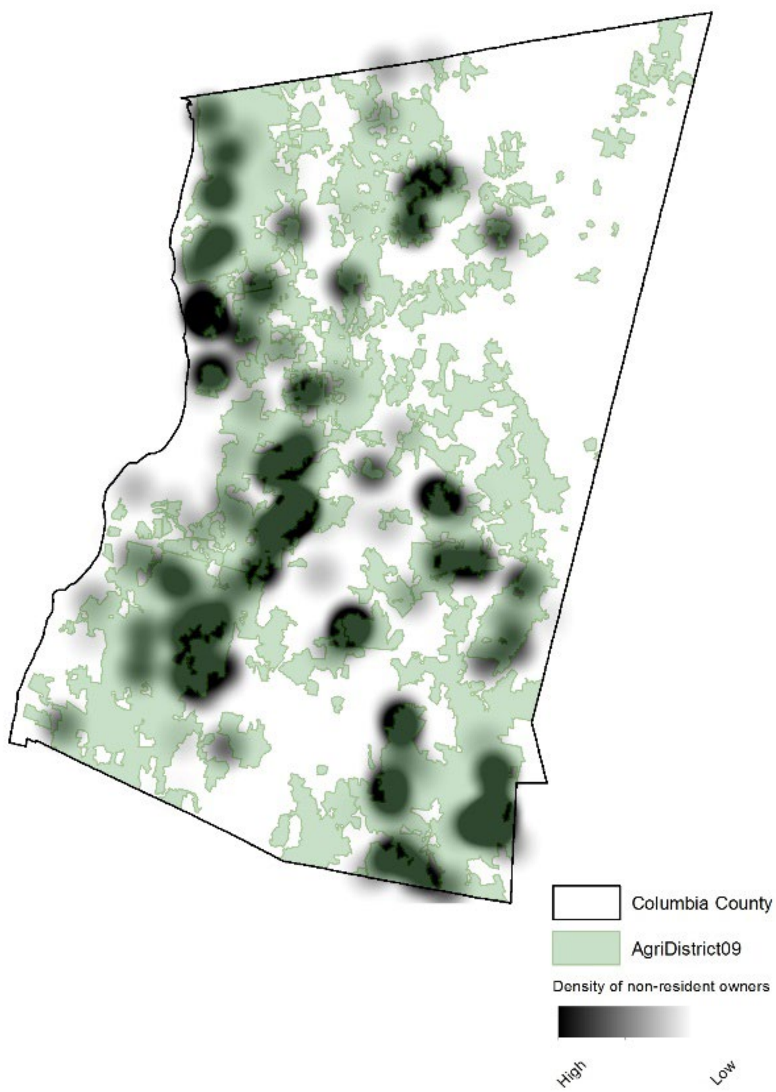


Figure G4. Propensity score histogram for Columbia (Caliper=.01, n=6)

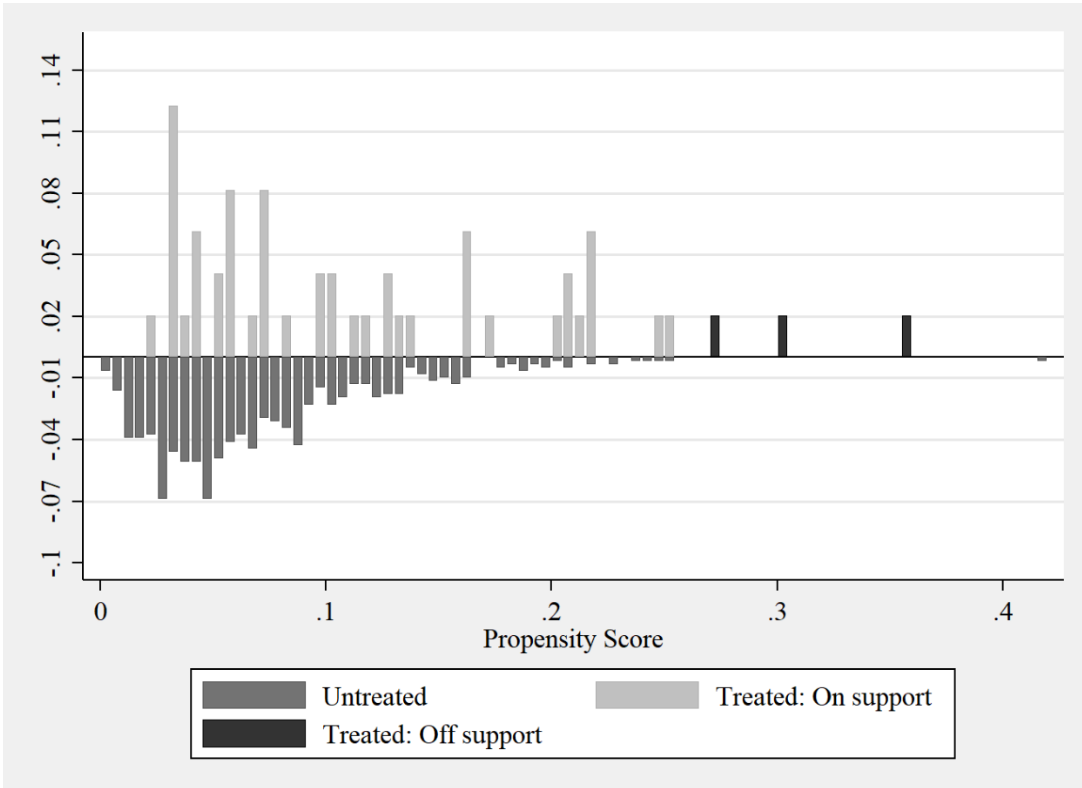
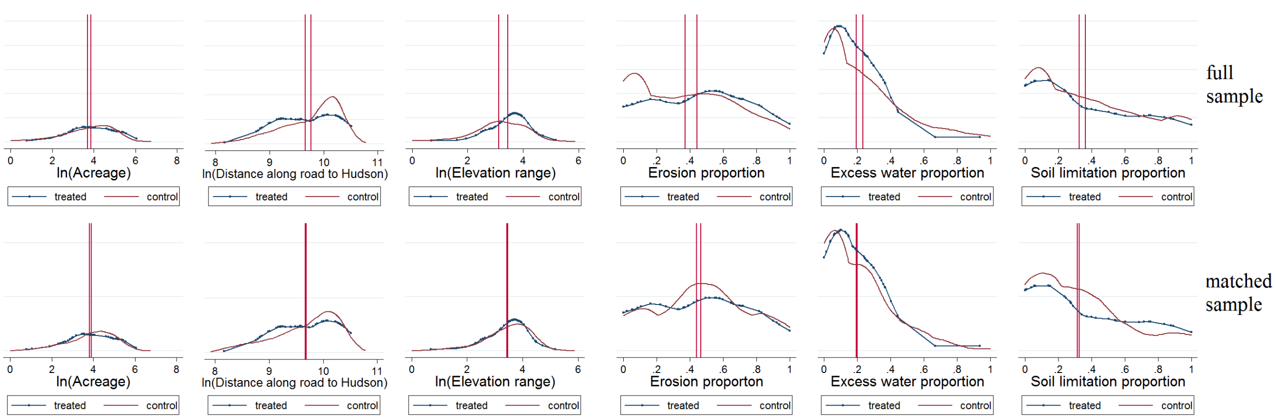


Figure G5. Kernel distributions comparing full sample and matched sample (Columbia, n=6 caliper=.01)



Note: vertical lines display the subsample means (by treated or not).