# The Role of Demand in Land Re-Development*

Felipe Carozzi*

*London School of Economics and Political Science*

**Abstract**

Vacant, previously-developed land in cities can generate negative externalities on surrounding areas, and is often the target of policies to promote re-development. This paper provides estimates of the price sensitivity of land re-development, a crucial parameter for the success of these policies. My estimates measure how prices affect long-run conversion of unused or underused previously developed land in England. The empirical strategy uses school test scores and admission district boundaries in a boundary discontinuity design to generate variation in housing demand that is orthogonal to re-development costs. Results show that the probability of re-development is effectively sensitive to housing prices. Estimates indicate that a 1% increase in prices leads to a 0.07 percentage point reduction in the fraction of hectares containing brownfield land. Price differences or substantial subsidies could lead to a significant amount of re-development in the long run. This is confirmed by observed land use changes between 2007 and 2011 being disproportionately concentrated in high price areas.

*Keywords:* Land re-development, Brownfields, Land Supply.

*JEL classification:* R14; R31

This Version: March 2020

*Address: Dept. of Geography and the Environment. London School of Economics. Houghton Street. London WC2A 2AE. Email: F.Carozzi@lse.ac.uk.

*I would like to thank André Anundsen, Peter Egger, Vernon Henderson, Christian Hilber, Ismir Mulalic, Luca Repetto, Stuart Rosenthal, Sefi Roth, Olmo Silva and Chris Timmins for useful comments and suggestions as well as participants at the 2017 SERC conference, the 2017 American meetings of the Urban Economics Association, the AQR workshop on urban economics, and the 5th workshop on urban economics at IEB. Excellent research assistance was provided by Aniket Baksy, Dibya Mishra and Koen Rutten. Funding from British Academy small grant for project "Micro Geographical Determinants of Residential Development" is gratefully acknowledged.

## 1. Introduction

Cities are areas of intense land use. Yet it is common to find vacant, idle or underused land within them. According to Pagano and Bowman (2000), up to 15% of land within US cities was vacant in 2000. In England, idle or underused previously-developed land sites represented 5.45% of total urban land in 2007 (Adams, De Sousa and Tiesdell, 2010). This prevalence is puzzling. Land demand is high in cities; in urban models this is what usually results in higher densities (Brueckner, 1987). Is the presence of these plots sensitive to local demand conditions? Knowing this is instrumental to understanding how market mechanisms shape urban density and, through density, the set of associated social, economic and environmental outcomes identified by urban economists. In addition, it is also informative about the potential success of the re-development promotion policies which have become popular over the past decades.

In this paper, I provide the first available estimates of the sensitivity of re-development to local housing prices, a key parameter linking infill site development with market forces. The analysis focuses on English previously-developed land (PdL) or *brownfield* sites, defined as land that was developed but is now vacant, derelict, or has known potential for re-development.[1] The empirical challenge to be addressed here is similar to the classical problem of estimating supply or demand elasticities from observed equilibrium outcomes. To estimate the price elasticity of PdL re-development I need a demand-shifter, a variable that affects demand but is known not to affect re-development costs or other determinants of supply. My empirical strategy will use school test scores and admission district boundaries in a boundary discontinuity design (BDD) to generate variation in housing demand that is credibly exogenous to costs. Implementation follows Gibbons, Machin and Silva (2013) in combining standard BDD methods (as in Black 1999 or Dell 2010) with spatial matching, and uses average test scores or school value-added as instruments for prices. I start by conducting a cross-sectional analysis to compare the presence of brownfield sites in high price and otherwise identical low price areas across an administrative boundary. Next, I extend this analysis to look at land use changes and estimate the effect of prices on subsequent residential construction activity.

My estimates indicate that a 1% increase in prices reduces the presence of brownfield sites in a hectare by 0.07 percentage points. While this effect appears to be small, only 1.5% of the hectares in my sample contained a PdL site in 2007. Therefore, results indicate that persis-

[1] In the United Kingdom, these are often referred to as brownfield sites. As a result, the terms brownfield and PdL site will be used interchangeably in the paper. It is important to note that, contrary to the definition prevalent in the United States, in the British definition brownfields are not necessarily polluted or affected by a hazardous substance.

tent increases in prices would lead to significant land re-development. This is consistent with separate estimates computed for ex-post land use changes in the 2007-2011 period.

By exploiting new boundaries introduced in the late 90s, as well as value-added measures of school quality, I address the potential confounding factors induced by persistent differences in administrative boundaries or household sorting, with little impact on my main estimates. In addition, I show that the price sensitivity of re-development is substantial in large cities, which are areas exhibiting a disproportionate number of PdL sites and also where local governments are more likely to engage in specific re-development policies. Using data on planning restrictiveness from Hilber and Vermeulen (2016), I show that the elasticity of re-development is lower in areas with tight planning restrictions. Finally, I find that brownfield presence is sensitive to prices even for relatively high cost sites, such as those in which land was previously used in manufacturing, mining or physical infrastructure.

My estimates inform the debate on *smart growth* policies and urban compactness in urban planning and economics. A growing consensus among urban planners and policy-makers has emerged on the desirability of achieving *compact cities* (see OECD (2012)), cities with high densities and few undeveloped patches of land within the urban footprint. The increased density is often assumed to reduce commuting time, promote productivity gains and reduce driving as well as environmental damage. While the debate over the supposed welfare enhancing effects of compactness continues (see Ahlfeldt and Pietrostefani (2019), Cheshire (2006)) several cities have embraced the smart growth agenda and its set of recommended policies. In many cases, governments have tackled this by promoting the re-development of brownfield or previously-developed land. Examples include the brownfield first policy in the United Kingdom which aims at channelling at least 60% of new developments to brownfield sites. Brownfield initiatives exist in Chicago, New York, Los Angeles and many other US cities, providing either grants or financing options to promote re-development. remediation relief for polluted sites provided by the Environmental Protection Agency in the United States often achieves a similar outcome. Appendix C provides a review of policies tailored to promote re-development and remediation in North America and Europe. The potential success of incentive-based re-development policies crucially depends on the price sensitivity of brownfield conversion. With a value of zero or close to zero, re-development is unlikely, even when prices increase substantially or a subsidy is in place. Conversely, positive and large elasticities imply high responsiveness and a potentially large effect of incentive policies such as tax breaks or remediation subsidies. Estimates reported below indicate a substantial price sensitivity, suggesting that these policies could induce a substantial amount of brownfield re-development.

This study contributes to the long literature housing supply. Housing supply elasticities are crucial to understand city-level house price volatilities (Glaeser, Gyourko and Saiz (2008), Paciorek (2013)), city systems' responses to shocks (Hornbeck and Moretti (2015)) and urban growth (Glaeser, Gyourko and Saks, 2006). The methods developed over the last twenty years combine longitudinal data on home-building with city level price indices (Malpezzi and Maclennan (2001), Green, Malpezzi and Mayo (2005)) and incorporate supply shifters such as geographical characteristics and regulation constraints to obtain city level elasticity estimates (as in Saiz (2010)). These city-level estimates are important in their own right, but they tell us very little about re-development of idle land, as city level changes in housing supply may also be affected by sprawl or changes in building heights. The estimates provided below are, to my knowledge, the first in the literature to focus specifically on re-development of idle or vacant land.

This study also contributes to the literature on land re-development. Wheaton et al. (1982) and Brueckner (1980) both provide urban models which feature re-development of housing stock when the price of redeveloped land is larger than re-development costs plus the price of land in its current use. This was successfully tested empirically in Rosenthal and Helsley (1994) using data for Vancouver and in subsequent empirical work (see for example Dye and McMillen (2007), Charles (2014), McMillen (2017) and the references therein). Most of these studies can estimate the relationship between local demand and re-development correctly under strong assumptions. My approach here is different in two dimensions. First, I focus on empty land parcels rather than teardown properties that have been the usual object of this literature. Second, I use a BDD method to obtain quasi-experimental variation in local demand that allows for cleaner estimation of my parameter of interest.

A strand of the literature on vacant land has focused on linking re-development delays to price uncertainty. Titman et al. (1985) presents a simple model showing that increasing demand uncertainty increases the option value of vacant land and reduces building activity. On the empirical front, Cunningham (2006) reports that areas with higher price uncertainty in Seattle feature higher land prices and slower development.[2] Instead of estimating the effect of uncertainty on re-development or timing to development, I attempt to credibly estimate the direct, first-order effects of demand on re-development and the pervasiveness of abandoned land.

My paper also contributes to the literature studying the effect of site clean-up policies on

[2]Cunningham (2007) shows that has significant implications regarding the effect of policies leading to uncertainty reductions such as growth boundaries or density restrictions. A review of the implications of uncertainty in future demand conditions on land values, re-development and abandonment can be found in Womack (2015).

local housing prices and neighbourhood composition. Greenstone and Gallagher (2008) use discontinuities in the assignment of clean-up funding to identify price effects of clean up policies in the US. Gamper-Rabindran and Timmins (2013) exploit the timing of clean-up and document that price effects are concentrated at the lower end of the price distribution. There is also evidence of household sorting as a response to these clean-up efforts (Gamper-Rabindran and Timmins, 2011) and of persistent negative externalities after clean up took place (Kiel and Williams, 2007). My paper follows the *opposite* direction relative to these studies, by looking at the effect of prices on conversion of PdL sites. In this sense, the estimates below should be an indication of the effectiveness of subsidies to trigger clean-up of polluted sites.

This study also relates to the literature estimating the effects of new construction on local housing prices. Ooi and Le (2013) use a hedonic model to study the effect of infill development of vacant or underused sites on local house prices in Singapore, finding positive and robust effects. Zahirovich-Herbert and Gibler (2014) study the effect of new construction and find price effects are especially high when newbuilds are well above the local mean floor area. My paper flips the direction of causality by studying how prices shape conversion and new-building at the local level. Finally, this paper relates to studies of urban decline. Both in Glaeser and Gyourko (2005) and in more recent work by Owens, Rossi-Hansberg and Sarte (2017), vacant land is a key feature of a declining city. My contribution relative to this work is to look at the link between demand and re-development in a context in which vacant land is not the result of general urban decline.

The rest of the paper is structured as follows: Section 2 presents the dataset used in the empirical analysis. Section 3 describes the empirical strategy and section 4 provides cross-sectional estimates of the effect of prices on land conversion, as well as results for *ex-post* land use changes. Section 5 provides a discussion of heterogeneity in my main estimates of interest. Section 6 presents a series of robustness checks to validate my analysis and section 7 presents the main conclusions.

## 2. Data

To conduct the empirical analysis I assemble a spatial dataset composed of over 1.6 million 100m $\times$ 100m cells or hectares located within 1 kilometre of a school admission boundary.[3] For each grid cell I know its location and related information such as the county it belongs to or

[3]The choice of cell size is arbitrary, but using postcodes or 200 $\times$ 200 metre cells as the spatial unit of analysis leads to qualitatively analogous findings. Hectare grid cells are chosen because postcode surface areas vary substantially between locations.

distance to the closest admission boundary. I use spatial data from different sources to impute information on the location of previously-developed land sites, housing prices, primary school performance, and other characteristics to each cell. The sources of this information are the following:

Data on previous developed land sites is obtained from the National Land Use Database of Previously Developed Land (NLUD-PDL). This dataset was assembled by the Department for Communities and Local Government from information provided by English Local Authorities. It includes geo-location (latitude and longitude), site area (in hectares), type of previous use and other characteristics for each site. I use the 2007 version of this dataset in this paper.[4] When conducting spatial imputation, sites are drawn as circles on a plane, centred in the coordinates of each site, with radii recovered from the surface area data. This data covers PdL sites for all of England.

I use two alternative sources of data for housing transaction prices and housing characteristics. The first source is a dataset from Nationwide, the UKs largest mortgage provider. It covers sales of properties sold under Nationwide mortgages for the years 2002-2006, with a total of 356,369 transactions. The data includes price, date of sale, postcode and a series of physical characteristics including house size (in squared metres), number of bedrooms and bathrooms, building age, and dummies indicating whether units have a garage or central heating. These characteristics are filtered out in a hedonic regression to obtain comparable price levels not driven by observed structural differences. I also use administrative data from the Land Registry including the universe of housing transactions in England for the years 2002-2006, with over 6 million sales in total. The dataset includes the date and price of each transaction as well as the postcode for each property. It also includes a small set of housing characteristics comprising property type, a newbuild dummy and a leasehold dummy.[5] Importantly, the dataset does not include information on floor size or other characteristics of the property. The lack of information about the property implies that the Land Registry data does not allow to filter hedonic characteristics. As a result, I only use this source in a robustness check (see section 6).

Data on school performance is obtained from the Department for Education for the years between 2002 and 2006. The school performance tables include data on standardized test scores for students completing key-stage 2 education level (11 years of age) as well as measures of

[4]Later versions are available for the years 2010, 2011 and 2012 which are broadly consistent with the 2007 version. I use the 2007 version here because I later use data on land use changes in an ex-post analysis for the years 2007-2011.The 2010 version of the NLUD-PDL is used in a robustness check in section 6.

[5]In the Land Registry Price Paid dataset property type is recorded as a categorical variable with types corresponding to detached, semi-detached, terraced houses and flats.

school value added, school postcodes and other school characteristics. Admission boundaries follow the boundaries of Local Educational Authority areas (LEAs). Grid cells are matched with the closest school within each LEA. All interior English LEA boundaries are used in the analysis. The boundaries of LEAs coincide with county boundaries, boundaries of Unitary Authorities and boundaries of metropolitan districts. Because the majority of these boundaries are county boundaries, and for ease of exposition, I will refer to LEA boundaries generically as county boundaries throughout the paper.

I also use data on physical land characteristics and other determinants of supply including elevation, landslide risk, planning refusal rates or agricultural land quality. Most of these variables are obtained from the British Geological Survey. They are used to validate the empirical strategy in section 3. Demographic characteristics at a disaggregated level are obtained from the 2001 census. The data is at the level of 2001 output areas (OAs) which are defined for data collection purposes as aggregations of postcodes containing roughly 140 households each. There were 165,665 OAs in England in the 2001 census, for which I have the fraction of black and Asian residents, fraction of unoccupied housing units, fraction of residents employed and unemployed. Census characteristics are included as controls in some specifications or used in validation exercises to identify residential sorting across boundaries.

When exploring the effect of demand on *changes* in land use, I use data from the Land Use Change Statistics (LUCS) recording these changes. The database is based on information from Ordinance Survey (the British national mapping service) for the years between 2007 and 2011. The LUCS data includes geo-referencing, approximate area, new and previous use of each site and year in which the change occurred. As before, I draw these sites in the plane, assuming circular shapes centred at the coordinates with radius inferred from the approximate area measure in the dataset.

When assembling my spatial dataset, I use digital maps for English OAs, postcodes, English counties (which define admission boundaries) and Unitary Authorities provided by the Office of National Statistics and Digimap.

These different sources of information are added to my grid cells by spatial imputation. Using the brownfield map, I construct a dummy taking value 1 if there is any PdL within the cell, and also a continuous variable measuring the fraction of PdL in each cell. These will constitute the main outcome variables in the analysis. I then impute to each grid cell data on hedonic-filtered house prices of the closest transaction, average test scores of the closest school within the admissions district, census characteristics from OAs, land-use changes, etc. Basic descriptive statistics for this dataset can be found in table B.1, in appendix B and further details on the dataset assembly process are presented in Appendix A.

## 3. Empirical Strategy

The empirical strategy presented here is devised to estimate the price sensitivity of re-development of previously-developed sites. To do so, we need to address the classic endogeneity problem in the estimation of supply elasticities from market outcomes. Observed variation in re-development of PdL may be driven by differences in demand or in re-development costs across locations. To obtain an unbiased estimate of supply responsiveness we need to induce variation in prices which is orthogonal to these costs. For this purpose, we can use spatial variation in an observable amenity as a demand-shifter, as long as we can safely assume that re-development costs and amenities are orthogonal. This may not be true in general. For example, in a mono-centric city, low amenity areas far from the city centre may be harder to re-develop because of reduced accessibility. I overcome this problem by using changes in accessible school test scores at school admission boundaries. School performance has been shown to influence parental location decisions, with increased demand in the catchment areas of better performing schools. Because the schools that can be accessed in different locations can vary over relatively short distances – i.e. when crossing an admission boundary – I can use the assumption that re-development costs and other confounders vary smoothly at these boundaries to identify the parameter of interest.

### *3.1. Baseline Strategy*

In estimation, I use a boundary discontinuity design. Boundary-designs have often been used to obtain estimates of willingness to pay for education quality as in Black (1999) or Gibbons, Machin and Silva (2013). I take a modified version of the specification used in these studies and use primary school performance to induce plausibly exogenous variation in log prices. Using my spatial dataset, I estimate the effect of prices on the amount of PdL in a given hectare by two stage least squares using the following specification:

$$PdL_i^{2007} = b_i + \beta ln(P_i) + County_i + \theta' X_i + u_i \quad (1)$$

where the log of prices imputed to grid cell $i$, $ln(P_i)$, is instrumented using average standardized test scores for the closest primary school on that side of the boundary. Boundary dummies $b_i$ are included so that the coefficient of interest is estimated exclusively from within-group variation for properties close, and on either side, of the boundary. Given that counties often have more than one boundary, we can also control for county specific effects $County_i$. Controlling for county effects is important because other policies may vary between counties.[6] In all specifications, I add a set of controls $X_i$ including independent terms for distance to the boundary in the high and low test scores sides, as well as latitude and longitude of each hectare centroid. Other controls relating to potential *supply-shifters* are included in some specifications, as indicated in the tables. Demographics are potentially affected by local prices, so my preferred specification does not include these as controls.[7] The error term in equation 1 is given by $u_i$. Throughout the analysis I cluster standard errors at the boundary level, so that all units on either side of one boundary between two counties are taken as belonging to the same cluster. Alternative clustering choices – for example, clustering at the county level – lead to very similar standard errors.

The interpretation of $\beta$ depends on the outcome variable used. For example, when using a dummy outcome, $\beta$ measures the effect of a 1% change in prices on the probability that a given hectare contains brownfield land. The estimated coefficient provides a measure of the price sensitivity of re-development. My hypothesis is that prices have a negative influence on the probability of having PdL land in a hectare.

The key assumption for identification is that the distribution of re-development costs – and other confounders – is the same on either side of the boundary. The credibility of this assumption hinges on selecting grid cells sufficiently close to the boundary and assuming that re-development costs vary continuously in space. Differences in policies by county would be dealt with by the county effects. Importantly, the identifying assumption in my context does *not* require that the first-stage correctly identifies the willingness to pay for educational services. As long as school performance only affects the presence of vacant land via changes in demand, the orthogonality condition should be satisfied. Issues which would compromise interpretation of the first-stage, such as household sorting in response to school performance (as in Bayer, Ferreira and McMillan (2007)), or the relative porosity of some boundaries would not, in principle, compromise identification of my parameter of interest. I discuss this specific point in section 6.

[6] For example, planning restrictiveness could differ between counties (see Hilber and Vermeulen (2016)). However, I will show below that planning restrictiveness, as measured by the fraction of rejected planning applications, is smooth across boundaries. County effects will partial out cross-sectional differences in policy by counties as long as polices are homogeneous within a county.

[7] Their inclusion has little impact on the quantitative findings. The implications of sorting for my results are discussed in detail in section 4.

### 3.2. Matching Across Boundaries

A problem with the method outlined above is that it relies on comparing hectares which are potentially far away from each other. Boundary fixed-effects ensure that estimates are obtained from within variation only and bandwidth restrictions ensure observations are not too far away from the boundaries. However, low school performance cells could still be several kilometres away from'the high performance cells to which they are compared, because I impose no further restriction on the distance between cells. To address this point, I conduct sample selection procedure based on a one-to-one matching of grid cells across the boundary. The algorithm proceeds as follows:

***Matching Algorithm***

1. Select the grid cell closest to boundary $b$.
2. Select the closest grid cell on the other side of $b$.
3. Attribute a match identifier, and a distance between grid cell centroids to the pair of cells
4. Remove the pair from the eligible cells.
5. Continue from step 1 until all cells are matched for $b$ on one side of the boundary.

Once all matches are obtained, I can estimate a spatially-differenced version of the equation above, in which all key variables are differenced within each match. Specification 1 then becomes:

$$\Delta PdL_i^{2007} = c_i + \beta \Delta ln(P_i) + \theta \Delta X_i + \Delta u_i \tag{2}$$

Spatial differencing eliminates the boundary fixed-effect $b_i$, as matches are obtained *within* each boundary. The differenced county fixed-effects can be captured by a new boundary effect $c_i$, as the spatial difference of county effects is equal for all cells on each side. Estimation of the spatially differenced equation can then follow the standard two stage least squares procedure, where $\Delta ln(P_i)$ is instrumented using the spatial difference in primary school scores.

### 3.3. Land Use Changes

I can use data on land use changes in locations containing brownfield land to study how these changes are shaped by demand conditions. This complementary analysis has two motivations. In the first place, it serves to confirm that price level differences lead to subsequent conversion. Secondly, it helps to deal with potential concerns about pre-existing differences in land use (e.g. location of manufacturing) which could be associated to school performance and give rise to brownfield areas in the long-run. If school performance differences were persistent over decades then historical land use patterns could be partially shaped by this amenity. The resulting observed cross-sectional differences in PdL site presence could not be interpreted as the result of conversion but rather as the consequence of initial differences in land use. Observing significant land use changes as a response to price differences would mitigate this concern substantially.

I modify my empirical strategy to look at the effect of price levels on *changes* in land use after 2007. I use data from LUCS to identify grid cells where at least some of the land changed towards residential use in the specified period. I then re-estimate my baseline model as specified in equation 1 but substituting the dependent variable for $LUCS_i^{07-11}$ which is the fraction of land within a grid cell $i$ which experienced a change towards residential use in the 2007-2011 period. This analysis is restricted to grid cells containing brownfield land, as defined in section 2. If high amenities lead to PdL conversion, I should find a positive estimate for $\beta$ in the modified specification. This would suggest that my cross-sectional estimates are not driven by differences in initial land use but rather by actual brownfield conversions responding to amenity differences across the boundary.

### *3.4. Validating the Empirical Strategy*

This empirical strategy used here requires that i) re-development costs vary smoothly at the boundary, and ii) average test scores have a sufficiently large effect on housing prices to warrant their use as an instrument.

I first show that supply-shifters such as physical characteristics of the terrain (elevation and landslide risk), of the soil (aquifer and agricultural land quality), presence of parks or gardens, or regulatory constraints vary continuously at the boundary. Aquifers and landslide risk are cited as important determinants of sprawl and residential density (and presumably development costs) in Burchfield et al. (2006) and Duranton and Turner (2016), respectively. Presence of parks or gardens may limit access to an area. Land parcels with higher agricultural yields may have higher alternative use and therefore be less likely to be developed. Finally, planning constraints can be seen as increasing construction costs substantially and are known to be an important obstacle to development (see Gyourko and Molloy (2015), Glaeser and Gyourko (2018) and the references therein). I measure them using average planning application refusal rates for English local planning authorities obtained from Hilber and Vermeulen (2016).

Figure 1 illustrates changes in supply-shifters at school admission boundaries. The horizontal axis measures distance to the boundary. I identify which side has the highest average test scores and sort cells according to this classification. Positive distances correspond to the high score side and negative distances correspond to the low score side of the boundary, respectively. Solid lines represent fitted values from fourth-degree polynomials estimated on either side, dashed lines correspond to 95% confidence intervals and gray circles correspond to distance bin averages. We can observe that in general there are no substantial discontinuities in supply drivers at the boundary.

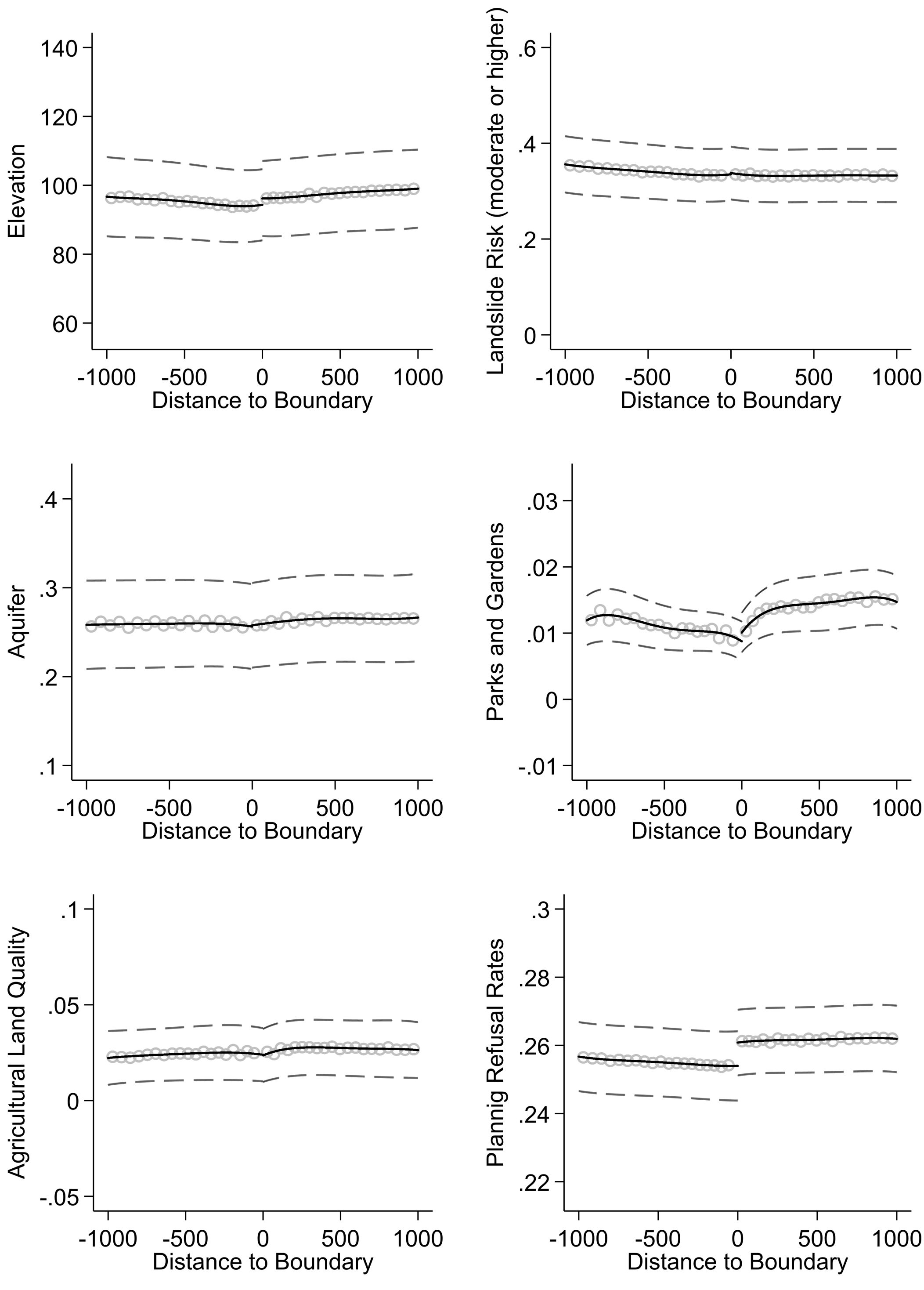
Table B.2 in Appendix B provides estimates of the partial correlations between school performance and the local supply shifters enumerated above using matching across boundaries. These estimates are a necessary complement to the illustration in Figure 1 because they incorporate more detailed variation in school performance in space.[8] We observe that in all cases the coefficients are very small and statistically insignificant in 23 out of the 24 cases. Take the example of planning refusal rates. The point estimate indicates that a 1 standard deviation increase in average test scores leads to a statistically insignificant 0.1 percentage point increase in planning refusal rates (over an average of 26%). These coefficients provide reassuring evidence that supply conditions do not jump with test scores at the boundary.

The first-stage is illustrated in Figure 2, which shows how school performance affects prices. The horizontal axis measures distance to the boundary, with cells in the low average score side having negative distances and cells in the high average score side having positive distances. The vertical axis represents mean prices in thousands of pounds (log scale). Solid lines are fitted using fourth-degree polynomials. We can observe a clear discontinuity in prices at the boundary, with high score areas having higher prices that low score areas a few metres away. There is substantial overlap between confidence intervals, largely because I am only using a fraction of the variation in average scores for this graphical illustration. A formal test of relevance of the instrument is conducted by estimating the first-stage with and without matching across boundaries, with results provided in Tables B.4 and B.3 in Appendix B. I provide estimates for different bandwidths (in columns) and different sets of controls and fixed-effects (in rows). Across specifications, we observe that the instrument is strong. Point estimates indicate that a 1 standard deviation increase in standardized scores leads to an increase of roughly 5% in housing prices. F-statistics for the instrument in question are provided in each row and are

[8]Figure 1 is drawn using a small fraction of the variation in school scores at the boundary by comparing the high and low score sides with no attention to the size of the actual jump in scores. In some boundaries this difference may be 1% and in others it may be 30%. To obtain a more precise estimate of these potential discontinuities incorporating more variation in scores, refer to appendix table B.2.

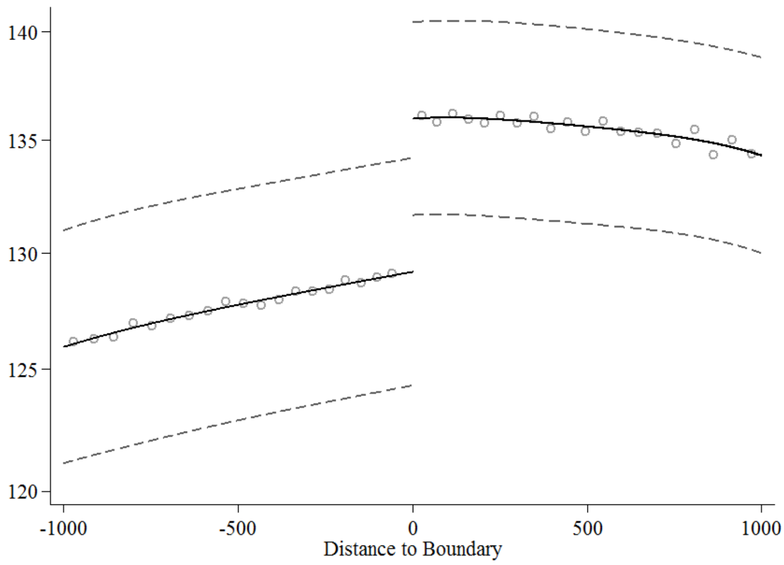
Figure 1

Covariate Balance for Determinants of Housing Supply




*Notes:* Horizontal axis represents distance to the boundary with negative distances corresponding to the county with low average test scores and positive distances corresponding to the county with high average test scores. Vertical axis corresponds to elevation above sea level (top-left), landslide risk (top right), probability of aquifer presence (centre left), probability of a park or garden in the cell (centre right), probability of high agricultural land quality (bottom left) and planning application refusal rates (bottom right). Fourth-degree polynomials fitted on the raw data represented in solid lines. 95% confidence intervals in dashed lines, with standard errors clustered at the boundary level. Gray circles correspond to averages taken within 40 distance bins.

Figure 2
First-Stage Illustration




*Notes:* Horizontal axis represents distance to the boundary with negative distances corresponding to the county with low average test scores and positive distances corresponding to the county with high average test scores. Fourth-degree polynomials fitted on the raw data represented in solid lines. 95% confidence intervals in dashed lines, with standard errors clustered at the boundary level.

all above 20.

Note that the effects estimated in the first-stage can be compared to those provided in the associated literature on willingness-to-pay for education. The closest paper in term of geographic scope and study period is Gibbons, Machin and Silva (2013). They find a 1 s.d. increase in ks1 scores (7 years of age) increases house prices by roughly 3%, and a similar effect of value added as measured by the increase in scores between ks1 and ks2 levels (between 7 and 11 years of age). My point estimates are in this range, with 3% usually falling within the associated 95% confidence interval.

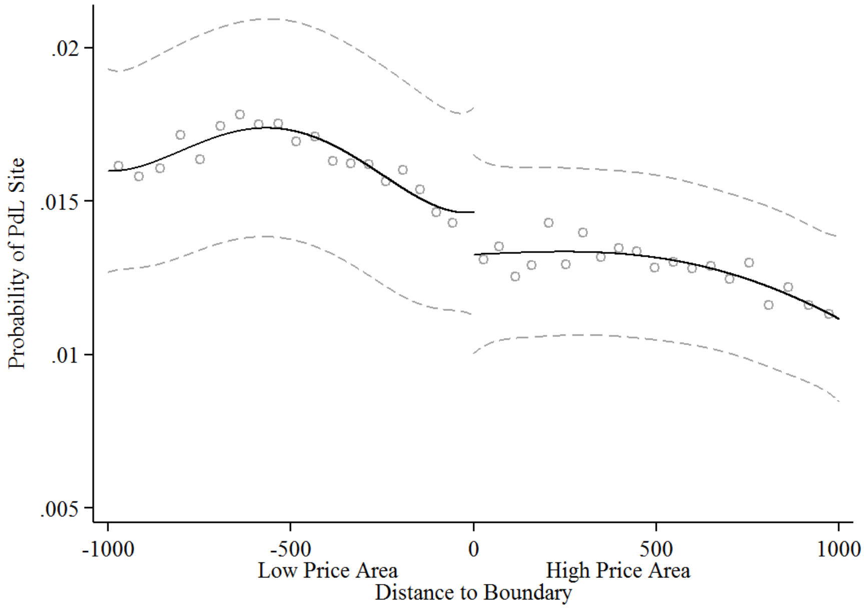
## 4. Main Results

### *4.1. Reduced-Form and Baseline Estimates*

Reduced-form effects of average test scores on brownfield presence are illustrated in Figure 3. The vertical axis now measures the average fraction of cells with some brownfield land at each distance. Fourth degree polynomials are independently estimated in both sides of the boundary, and gray circles correspond to local averages. We observe a discontinuity in the fraction of hectares containing PdL sites, with brownfield being less likely on the high price (high test scores) side of the boundary. Confidence bands are large and it is unclear whether

FIGURE 3

REDUCED-FORM GRAPH




*Notes:* Horizontal axis represents distance to the boundary with negative distances corresponding to the county with low average test scores and positive distances corresponding to the county with high average test scores. Fourth-degree polynomials fitted on the raw data represented in solid lines. Reduced-form estimates when incorporating intensive margin variation in school performance can be found in Table B.5 in Appendix B.

this discontinuity is indeed statistically significant. One reason for this is that we are only using a fraction of the variation in performance by grouping hectares on good and bad sides only. The size of the difference in scores across the admission boundary can differ from boundary to boundary, not to mention that there is also variation in school performance *within* each side. In fact, estimates from reduced-form regressions of the fraction of brownfield land on test scores, reported in appendix Table B.5, indicate school performance has a negative and significant effect on the presence of brownfields at all conventional levels.

In order to quantify the magnitude of the effect of prices on the conversion of brownfield land, I obtain 2SLS estimates using test scores as an instrument for housing prices. Results for different outcomes, spatial bandwidths and sets of controls are presented in Table 1. Columns correspond to different bandwidths, with the first column estimates obtained using cells within 1km of an admission boundary and column 4 corresponding to cells within 250m of a boundary. Estimates of the effect of prices on the probability of finding brownfield land in a grid cell – reported in panels A and B – are consistently negative and significant at all conventional levels. The point estimates are fairly stable, at roughly -0.085. This means that a 1% increase in prices reduces the probability of a PdL site by 0.085 p.p. While this effect is apparently small, it is economically significant. Recall that the baseline average in the dependent variable is 1.45% so this implies that a substantial increase in prices would make a sizeable dent in the presence of brownfield sites. Assuming that the true effect is indeed linear in log prices this would mean that a 16% increase in prices would essentially wipe out all brownfield sites. This linearity assumption is admittedly strong, but provides an easily interpretable back of the envelope figure. Estimates of the effect of prices on the *fraction* of PdL in each grid cell are reported in columns C and D of Table 1 and also indicate a substantial effect of prices on re-development. A 1% increase in prices reduces the fraction of brownfield land in a grid cell by roughly 0.04 p.p.

Table 1
Baseline IV Estimates

| | 1000m | 750m | 500m | 250m |
|---|---|---|---|---|
| | PdL Site | PdL Site | PdL Site | PdL Site |
| **A. Binary Outcome** | | | | |
| Log(Price) | -0.089*** | -0.095*** | -0.091*** | -0.087*** |
| | (0.014) | (0.016) | (0.016) | (0.017) |
| **B. Binary Outcome, Supply Controls** | | | | |
| Log(Price) | -0.077*** | -0.084*** | -0.080*** | -0.080*** |
| | (0.013) | (0.015) | (0.015) | (0.017) |
| | PdL Fraction | PdL Fraction | PdL Fraction | PdL Fraction |
| **C. Continuous Outcome** | | | | |
| Log(Price) | -0.043*** | -0.047*** | -0.045*** | -0.044*** |
| | (0.008) | (0.010) | (0.010) | (0.010) |
| **D. Continuous Outcome, Supply Controls** | | | | |
| Log(Price) | -0.036*** | -0.040*** | -0.039*** | -0.038*** |
| | (0.008) | (0.009) | (0.009) | (0.009) |
| Observations | 1350047 | 1044354 | 720359 | 375375 |

*Notes:* 2SLS estimates. In panels A and B the dependent variable is a dummy taking value 1 if there is a PdL site in the grid cell. In panels C and D the dependent variable measures the fraction of brownfield land in the grid cell. Boundary and county fixed-effects as well as controls for linear terms for distance to the boundary on either side, latitude and longitude included in all specifications. Supply-shifters included as controls in panels B and D. Columns 1 to 4 correspond to bandwidths of 1000, 750, 500 and 250 metres, respectively. S.E. clustered at the boundary level in parentheses.

### *4.2. Spatial Matching Estimates*

My baseline estimates in Table 1 are obtained from regressions including boundary fixed-effects. Even when the bandwidth around boundaries is restricted to short distances of 250 metres, grid cells used in estimation could still be quite far away from each other. In order to avoid this, I employ the matching method described in section 3.2 and use spatial differenced data to obtain my preferred estimates of the effect of prices on the incidence of previously developed land sites.

Instrumental variable estimates of the sensitivity of brownfield conversion to prices for the matched specification are provided in Table 2. As above, different columns correspond to different bandwidths around the boundary. In panels A and B the outcome variable is a dummy indicating some PdL in the grid cell. The coefficients are negative and significant across specifications. In my preferred specification, using a 250 metre bandwidth, a discrete outcome and including the supply controls, the effect of interest is -0.069 (panel B, final row). This implies that a 1% increase in prices reduces the probability of finding a PdL site in a given grid cell by $0.069$ percentage points. Under the assumption of linearity of the effect of log prices on the probability of re-development, this result implies that a 21% increase in housing prices across English locations would be necessary to prompt the conversion of most English PdL sites in the long run. Note that the baseline estimate of 0.08 from Table 1 falls within the 95% confidence interval of this estimate given by $-0.069 \pm 1.96 \times 0.029$.

Table 2
Matched IV Estimates

| | 1000 | 750 | 500 | 250 |
|---|---|---|---|---|
| | Δ PdL Site | Δ PdL Site | Δ PdL Site | Δ PdL Site |
| **A. Binary Outcome** | | | | |
| Δ Log(Price) | -0.086*** | -0.100*** | -0.094*** | -0.075*** |
| | (0.027) | (0.031) | (0.029) | (0.026) |
| **B. Binary Outcome, Supply Controls** | | | | |
| Δ Log(Price) | -0.071*** | -0.083*** | -0.084*** | -0.069*** |
| | (0.022) | (0.026) | (0.027) | (0.026) |
| | Δ PdL Fraction | Δ PdL Fraction | Δ PdL Fraction | Δ PdL Fraction |
| **C. Continuous Outcome** | | | | |
| Δ Log(Price) | -0.039** | -0.050** | -0.049** | -0.045** |
| | (0.018) | (0.021) | (0.019) | (0.018) |
| **D. Continuous Outcome, Supply Controls** | | | | |
| Δ Log(Price) | -0.027** | -0.034** | -0.037** | -0.037** |
| | (0.012) | (0.014) | (0.015) | (0.016) |
| N | 600922 | 464178 | 318984 | 165676 |

*Notes:* 2SLS estimates after spatial matching. Δ corresponds to a difference taken within matched pairs. In panels A and B the dependent variable is a dummy taking value 1 if there is a PdL site in the grid cell. In panels C and D the dependent variable is the fraction of 2007 brownfield land in the grid cell. Boundary fixed-effects as well as controls for differenced linear terms for distance to the boundary, latitude and longitude included in all specifications. Supply-shifters differenced between matched pairs included as controls in panels B and D. Columns 1 to 4 correspond to bandwidths of 1000, 750, 500 and 250 metres, respectively. S.E. clustered at the boundary level in parentheses.

Results obtained using a continuous outcome to measure brownfield presence are reported in panels C and D of Table 2. Point estimates lie between 0.027 and 0.05 depending on the bandwidth and set of controls, with the point estimate in my preferred specification being -0.037. In all cases, estimates are significant at the 5% level. The average fraction of brownfield land in cells within 1 km of a county boundary (including zeroes) is 0.64%. Therefore, my preferred estimate suggests that a 18% increase in prices in all locations would be sufficient to trigger re-development of the vast majority of PdL sites.

Taken together, the results of the cross sectional analysis reveal that there is substantial sensitivity of brownfield re-development to prices. This finding indicate that market forces can induce substantial upgrading of PdL. It is arguably also encouraging for remediation relief and other infill re-development policies such as subsidies or tax breaks which depend on high price sensitivity to be successful. Given that my estimates are cross-sectional, and that school performance is quite persistent, we can interpret them as long-run estimates of the effect of prices on re-development, with short run estimates likely to be lower.

*4.3. Unitary Authorities*

Many of the boundaries used to estimate the parameter of interest in section 4 have been administrative boundaries over a long period of time. This could pose a series of problems for the empirical strategy used here. For example, if areas with access to good schools decades ago were less likely to contain manufacturing activities, this could lead to lower brownfield presence today. If school performance is persistent, then this could induce the results in Tables 1 and 2 even in the absence of an effect of demand on ex-post conversion. Sustained differences in investment on the housing stock on different sides of the boundary, as identified by Kane, Riegg and Staiger (2006), could also compromise identification.

In order to address this issue I obtain estimates based only on recently created boundaries, established after the introduction of Unitary Authorities (UAs) by a decentralization reform in the 1990s. UAs are administrative divisions of UK local government similar to counties in their functions, including those related to public education. Because these admission boundaries are relatively new, there was no sharp discontinuity in available school performance in these areas prior to the creation of UAs between 1995 and 1998. Hence, when focusing on these boundaries, we know initial differences in land use from decades past are unlikely to be correlated with tests scores in the early 2000s.

Table 3 provides results for the restricted sample of UA boundaries. Because counties typically have only one boundary with a UA, I cannot use county fixed-effects when imposing this sample selection. I only provide estimates with and without supply-shifter controls and for different bandwidths around the boundary. We observe coefficients oscillate around $-0.08$, similar to those reported in Table 1 for the full sample. This is reassuring as it shows that my results are largely unchanged when focusing on boundaries where persistent differences in school performance cannot have generated differences in initial land use.[9]

TABLE 3
BASELINE ESTIMATES FOR UNITARY AUTHORITY BOUNDARIES ONLY

| | 1000m | 750m | 500m | 250m |
|---|---|---|---|---|
| | PdL Site | PdL Site | PdL Site | PdL Site |
| Log(Price) | -0.083*** | -0.087*** | -0.074*** | -0.079*** |
| | (0.023) | (0.026) | (0.027) | (0.029) |
| F-Stat | 54 | 53 | 52 | 48 |
| Supply Controls | N | N | N | N |
| | PdL Site | PdL Site | PdL Site | PdL Site |
| Log(Price) | -0.080*** | -0.086*** | -0.075*** | -0.084*** |
| | (0.021) | (0.025) | (0.026) | (0.030) |
| F-Stat | 48 | 46 | 44 | 39 |
| Observations | 441596 | 341965 | 236265 | 123505 |
| Supply Controls | Y | Y | Y | Y |

*Notes:* 2SLS estimates. Sample restricted to hectares around county boundaries for Unitary Authorities created before 2000. Dependent variable is a dummy taking value 1 if there is a previously-developed land plot in a grid cell. Columns 1 through 4 correspond to bandwidths of 1000, 750, 500 and 250 metres, respectively. Supply-shifters included as controls in second row. First-stage F-statistics provided below each estimate. S.E. clustered at the boundary level in parentheses. Matching estimates reported in Appendix B).

### *4.4. Land-Use Changes*

I now present results for the analysis of the effect on prices on observed land use changes over the 2007-2011 period. The purpose of the analysis is to elucidate whether the differences in brownfield presence documented in the previous sections are simply a static fact or rather whether conversions continue in high residential price areas after 2007. The sample is now restricted to all grid cells containing some brownfield land. First-stage estimates for the restricted sample have been relegated to Table B.7 in Appendix B and confirm the instrument is still strong, despite the sample restriction.

The resulting IV estimates of the effect of price levels on land use changes are presented in Table 4. The dependent variable is the fraction of the brownfield cell that experienced a change

[9]Similar estimates with spatial matching are presented in Table B.6 in Appendix B. Precision suffers because both the number of clusters and the number of observations within a cluster drop substantially after the sample restrictions associated to UA boundaries and to matching are combined. Point estimates in that table are around $-0.055$, slightly lower but not very different from those in Table 2 which oscillated around $-0.069$ for short bandwidths.

Table 4
Land Use Change & Prices - IV Estimates

| | 1000m | 750m | 500m | 250m |
|---|---|---|---|---|
| | LUC Fraction | LUC Fraction | LUC Fraction | LUC Fraction |
| Log(Price) | 0.047** | 0.046* | 0.050* | 0.070* |
| | (0.023) | (0.027) | (0.030) | (0.040) |
| County Effects | N | N | N | N |
| Supply Controls | N | N | N | N |
| Log(Price) | 0.061** | 0.063* | 0.065 | 0.121* |
| | (0.028) | (0.036) | (0.041) | (0.070) |
| County Effects | Y | Y | Y | Y |
| Supply Controls | N | N | N | N |
| Log(Price) | 0.059** | 0.061* | 0.054 | 0.098* |
| | (0.027) | (0.034) | (0.037) | (0.058) |
| Observations | 16729 | 12985 | 8861 | 4508 |
| County Effects | Y | Y | Y | Y |
| Supply Controls | Y | Y | Y | Y |

*Notes:* 2SLS Estimates. Sample restricted to hectares with PdL land. Dependent variable is the fraction of land experiencing a change towards residential use in the grid cell between 2007 and 2011. Boundary fixed-effects as well as controls for differenced linear terms for distance to the boundary, latitude and longitude included in all specifications. County fixed-effects included in second and third rows. S.E. clustered at the boundary level in parentheses.

to residential land use between 2007 and 2011. We observe positive effects across specifications, with most coefficients being significant at the 5% or 10% level. Qualitatively, this implies that cells with higher amenity-induced demand are more likely to experience a change towards residential land use, consistent with the notion that higher prices lead to brownfield conversion. Most coefficients are only marginally significant, but this is likely to be a consequence of the lack of precision induced by the small sample, as will be discussed in section 6. My preferred specification corresponds to the estimates using the shortest bandwidth in the equation with county effects and controls, yielding an estimate of $0.098$. Over the 2007-2011 period, areas with higher amenity values experienced higher rates of brownfield conversion. A 1% increase in prices increases the fraction of a hectare changing towards residential use by 0.1 percentage points. Two conclusions can be extracted from these results. In the first place, they show that the conversion of brownfields takes time, as differences in school performance were still leading to conversions after the 2002-2006 period. The fact that the estimated effects reported in Table 4 are relatively small suggests that much of the conversion had already taken place in 2007. Secondly, these results indicate it is unlikely that cross-sectional estimates are driven by initial differences in land use.

I conduct an additional exercise to further validate the results reported in Table 4. As has been argued above, access to schools with good test scores can be seen as an amenity affecting demand of housing in a specific area. While other amenities such as transport access or clean air may affect demand for different land uses, local access to good schools is likely to have an effect only on households, and hence, should only attract re-development towards residential activities. With this intuition in mind, I replace a measure of land use changes to commercial, industrial and other non-residential uses as dependent variable and estimate the effect of prices on these changes. Results for this exercise are provided in Table 5. Across specifications we obtain small and often negative effects of house prices on changes towards non-residential use. Coefficients are statistically insignificant in all cases. The zero coefficients are consistent with the notion that school performance is a residential-specific amenity.

Table 5

Non Residential Land Use Changes

| | 1000m | 750m | 500m | 250m |
|---|---|---|---|---|
| | Chng.Use Other | Chng.Use Other | Chng.Use Other | Chng.Use Other |
| Log(Price) | -0.018 | -0.028 | -0.014 | 0.019 |
| | (0.03) | (0.04) | (0.04) | (0.04) |
| County Effects | N | N | N | N |
| Supply Controls | N | N | N | N |
| Log(Price) | -0.002 | -0.025 | -0.019 | 0.023 |
| | (0.06) | (0.07) | (0.08) | (0.08) |
| County Effects | Y | Y | Y | Y |
| Supply Controls | N | N | N | N |
| Log(Price) | 0.022 | -0.004 | 0.005 | 0.059 |
| | (0.09) | (0.11) | (0.11) | (0.11) |
| Observations | 22774 | 17795 | 12174 | 6175 |
| County Effects | Y | Y | Y | Y |
| Supply Controls | Y | Y | Y | Y |

*Notes:* 2SLS Estimates. Sample restricted to hectares with PdL land. Dependent variable is the fraction of the grid cell that experienced a change towards a non-residential use (commercial, industrial, etc.) between 2007 and 2011. Boundary fixed-effects as well as controls for distance to the boundary, latitude and longitude included in all specifications. County fixed-effects included in second and third rows. S.E. clustered at the boundary level in parentheses.

## 5. Heterogeneous Effects

In this section, I discuss factors that could affect the sensitivity of re-development to demand conditions. The potential effects of an increase in local demand – or an incentives-based brownfield remediation policy – could differ substantially depending on the origin and characteristics of different types of sites, the policy environment in the form of planning restrictions, or the location of the site in the city system. I explore whether the price sensitivity of brownfield presence is shaped by these conditions.

*5.1. Previous Land Use*

The responsiveness of re-development to demand may depend on the previous use of abandoned or derelict sites. It is reasonable to think, for example, that the cost of re-development is higher for sites that used to be devoted to manufacturing, mining or physical infrastructure as these may contain pollutants or require specialized machinery. This difference in the cost distribution could lead to a difference in price sensitivity. To study this, I split brownfields between those with a manufacturing origin and those with a different origin (residential, commercial or other) and construct variables $PdL_i^{Manuf}$ and $PdL_i^{Other}$ taking value 1 if there is PdL of manufacturing or other origins in a grid cell, respectively.[10] I then obtain separate estimates of the effect of log prices on each type of brownfield.

Results for this exercise are provided in Table 6. We observe that the estimates are very similar for both types of brownfields. For example, baseline estimates of the effect of log price on the probability of development are -0.043 and -0.037 for manufacturing and other brownfields, respectively. This shows the responsiveness of brownfield re-development to prices is unaffected by previous land use and, presumably, the associated cost of re-development. If this is indeed the case, then a subsidy is likely to have a similar effect on conversion for different types of sites. This is relevant because remediation policies are usually tailored to high cost sites containing or presumably containing chemicals, pollutants or heavy structures such as those found in ports or large manufacturing plants.[11] Moreover, it is these sites that are likely to generate the largest negative externalities and, therefore, the ones for which re-development can have a larger effect on welfare. It is encouraging that the re-development of these sites is also sensitive to demand conditions.

*5.2. Planning Restrictiveness*

I now turn to study whether the price sensitivity of brownfield re-development varies with planning restrictiveness. Planning policy is perhaps the single most relevant policy affecting

[10] Alongside manufacturing I also consider sites with mining, or infrastructure origin as these are expected to also have high or very high re-development costs. the information on previous use comes from the NLUD-PDL source.

[11] See Appendix C for a policy review including details on site eligibility conditions and examples of government funded re-development projects.

Table 6
Previous Land Use

| | PdL Manufacturing | | PdL Other Sites | |
|---|---|---|---|---|
| Log(Price) | -0.040*** | -0.043*** | -0.043*** | -0.037*** |
| | (0.008) | (0.009) | (0.010) | (0.011) |
| Observations | 1044354 | 375375 | 1044354 | 375375 |
| Bandwidth | 750m | 250m | 750m | 250m |
| Matching | N | N | N | N |
| | PdL Manufacturing | | PdL Other Sites | |
| $\Delta$ Log(Price) | -0.040** | -0.029* | -0.043*** | -0.040** |
| | (0.017) | (0.017) | (0.015) | (0.017) |
| Observations | 464178 | 165675 | 464178 | 165675 |
| Bandwidth | 750m | 250m | 750m | 250m |
| Matching | Y | Y | Y | Y |

*Notes:* Dependent variable is a dummy taking value 1 if there is a previously-developed site of manufacturing origin (columns 1 and 2) or of other origins (columns 3 and 4) in the grid cell. Columns 1 and 3 correspond to estimates using bandwidths of 750 metres and columns 2 and 4 use bandwidths of 250 metres. All specifications include county effects and control for supply-shifters. S.E. clustered at the boundary level in parentheses. Second row corresponds to estimates obtained using spatial matching of cells. S.E. clustered at the boundary level in parentheses.

land use. Understandably, its link with land supply, prices, or the allocation of land to different uses has been studied thoroughly. Planning policy can sometimes operate by imposing restrictions to development.[12] Does this have an impact of the sensitivity of re-development to prices? In order to test whether this is the case, I use a measure of planning refusals borrowed from Hilber and Vermeulen (2016). This records the fraction of all applications that were rejected in the planning process between 1979 and 2008 for all English planning authorities. I use this refusal rate as a cross-sectional proxy for the different levels of restrictiveness of these planning authorities across locations. I calculate the average refusal rate taken over *both sides* of each boundary. I then split my sample around the median of these rates and obtain separate estimates for each sub-sample of boundaries.[13]

Results are provided in the first two rows of Table 7. Panel A displays estimates for low refusal rate areas - areas with low planning restrictiveness - and panel B displays those for high refusal areas. Columns 1 and 2 provide estimates of the effect of prices on the probability of finding a PdL site in a grid cell. Estimates in columns 3 and 4 measure the effect of prices on

[12]See Hilber (2017) for a discussion of the restrictions on development imposed by the English planning system.
[13]I split the sample into a high refusal and a low refusal boundaries. This ensures that grouped bands of hectares around each boundary are kept together when selecting the sample.

the fraction of developed land within a cell. We observe that prices affect brownfield presence both for areas with low and high refusal rates. However, point estimates are roughly 25% larger (in absolute value) in high refusal rate areas that in low refusal rate areas across specifications.

To incorporate more of the variation in refusal rates in the analysis I estimate the baseline equation adding an interaction between log prices and the average refusal rate in the boundary to the second stage. I add in as an instrument an interaction between average test scores and the refusal rate. This helps to deal with endogeneity of the price variable but refusal rates may still be endogenous and, therefore, interpretation should proceed with caution. Panel C provides estimates of both the coefficient of log price and the interaction term. The refusal rate has been normalized to have mean zero and a s.d. of 1 therefore the first coefficient estimates price sensitivity at the mean and the interaction term indicates how this sensitivity changes for a 1 s.d. increase in restrictiveness. The coefficient on the interaction term is positive in all specifications and statistically significant in three out of four cases. Looking at columns 1 and 2, this indicates that in areas with planning restrictiveness 1 s.d. above the mean, the price responsiveness is about 2/3 of the magnitude at the mean.

I interpret this as suggesting that planning restrictiveness prevents brownfield re-development. This also indicates that an incentive based policy for re-development (e.g. a subsidy) could have larger effects in places with relatively more lenient planning processes.

### *5.3. Cities*

Previously developed land sites occur in urban and rural environments, but sensitivities to re-development could differ. For example, the re-development cost distribution could be different in these different areas because of lack of access to employment or road transport in remote locations. To explore this I study different subsamples, focusing only on urban spatial cells, the 20 largest English cities or the metropolitan area of London. Sample restrictions applied here are based on Travel-to-work Areas, which are metropolitan areas for selected cities defined based on spatial commuting patterns.

Estimates for different subsamples are provided in table 8, where the first row presents baseline estimates and the second presents estimates after spatial matching. In both cases I measure PdL sites with a dummy taking value 1 if there is any brownfield land within the hectare.[14] The first column presents the baseline and matched estimates that have been reported above for comparison purposes. In column 2, the sample is restricted to hectares in urban areas

[14] Estimates using the fraction of brownfield land in a grid cell are reported in Table B.8 in appendix B.

### Table 7
Planning Restrictiveness

| | PdL Site | PdL Site | PdL Fract. | PdL Fract. |
|---|---|---|---|---|
| **A. Low Refusal Rates** | | | | |
| Log(Price) | -0.097*** | -0.099*** | -0.044*** | -0.048*** |
| | (0.020) | (0.021) | (0.013) | (0.013) |
| Observations | 242004 | 126072 | 242004 | 126072 |
| **B. High Refusal Rates** | | | | |
| Log(Price) | -0.068*** | -0.064** | -0.037*** | -0.031** |
| | (0.023) | (0.026) | (0.014) | (0.013) |
| Observations | 478355 | 249303 | 478355 | 249303 |
| **C. Interactions** | | | | |
| Log(Price) | -0.072*** | -0.070*** | -0.035*** | -0.033*** |
| | (0.015) | (0.016) | (0.009) | (0.009) |
| Log(Price) × Refus. Rate | 0.020** | 0.023** | 0.008 | 0.012** |
| | (0.010) | (0.010) | (0.006) | (0.006) |
| Observations | 720359 | 375375 | 720359 | 375375 |
| Bandwidth | 500m | 250m | 500m | 250m |
| County Effects | Y | Y | Y | Y |
| Controls | Y | Y | Y | Y |

*Notes:* Dependent variables are a dummy taking value 1 if there is a PdL site in the hectare in columns 1 and 2, and the fraction of PdL land in the grid cell in columns 3 and 4. Panel A displays estimates obtained with the sub-sample of boundaries with below median refusal rates, panel B displays estimates for the sub-sample of boundaries with above median refusal rates, and panel C uses the full sample and adds an interaction term between imputed prices and average refusal rates. Bandwidths around the boundary are 250 metres in columns 1 and 3 and 500 metres in columns 2 and 4. All specifications include boundary and county fixed-effects as well as controls for supply conditions. S.E. clustered at the boundary level in parentheses.

based on the urban/rural classification by the ONS. We see that the point estimates increase relative to those obtained with the full sample. Column 3 reports results excluding London. The point estimates are almost unchanged, indicating results are not driven solely by the English capital, which accounts for roughly 5% of my sample. Results in columns 4 and 5 are obtained focusing exclusively on London and on the 20 largest cities in England, respectively.[15]

Together with the results for all urban hectares, these estimates show that the sensitivity of unused land to prices is higher in cities. This occurs mainly because the baseline probability of having a PdL site in a grid cell is generally larger in urban areas. Differences in amenities will generate very limited changes in brownfield presence in areas which have almost no brownfields to begin with.[16] That being said, it is important that the effect of demand conditions on

[15]The abnormally high coefficient for London may be driven by the fact that cells containing brownfield land are roughly five times more common in the British capital than in the rest of the sample.

[16]The fraction on hectares containing PdL in my full sample is 1.45%, while the same fraction is as high as 2.8%

brownfields is clearly present and large in cities given that it is it is usually city or local governments in urban areas that are more concerned with remediation and PdL regeneration policies. An additional message from the estimates reported in Table 8 is that the main results in section 4 are not driven by a few geographical areas.

Table 8
Cities

| | Full Sample | Urban Only | Excl. London | London | 20 Largest |
|---|---|---|---|---|---|
| | PdL Site | PdL Site | PdL Site | PdL Site | PdL Site |
| Log(Price) | -0.080*** | -0.112*** | -0.070*** | -0.352** | -0.085*** |
| | (0.017) | (0.025) | (0.017) | (0.147) | (0.026) |
| Matching | N | N | N | N | N |
| | Δ PdL Site | Δ PdL Site | Δ PdL Site | Δ PdL Site | Δ PdL Site |
| Δ Log(Price) | -0.069*** | -0.132*** | -0.061** | -0.354 | -0.083** |
| | (0.026) | (0.049) | (0.026) | (0.298) | (0.032) |
| Matching | Y | Y | Y | Y | Y |

*Notes:* 2SLS estimates for different sub-samples including cities and groups of cities. First column corresponds to the full sample, second column excludes the London metropolitan area, third column restricts the sample to cells in the London metropolitan area, and fourth column restricts the sample to the 20 largest English metropolitan areas by size of workforce. Dependent variable is the a dummy taking value 1 if the cell contains PdL. In all specifications, the bandwidth is 250 metres. First row corresponds to baseline estimates. Second row corresponds to estimates obtained using spatial matching. S.E. clustered at the boundary level in parentheses.

## 6. Robustness Checks

In this section, I present a series of robustness checks to show my main results are not driven by i) household sorting in response to brownfield sites contaminating the instrument, ii) details of the specification such as bandwidth choice or functional form assumptions, or iii) measurement error in brownfield location.

### *6.1. Household Sorting and School Value-Added*

Previous work on estimating willingness to pay for education has shown households of different characteristics sort in response to differences in school performance (see Bayer, Ferreira and McMillan (2007)). While this may be a concern in the estimation of the willingness-to-pay for school performance, it would only have limited impact here: That the first-stage fails to identify a potential parameter of interest does not preclude the use of the instrument for our purposes, as long as the orthogonality condition is still met.

in the 20 largest English urban areas.

What *could* suppose a threat to identification in this paper is sorting of households in response to brownfield sites subsequently affecting educational outcomes. Previous research has documented the presence of negative externalities for at least some of these sites (Greenstone and Gallagher (2008), Kiel and Williams (2007)), as well as sorting in response to these externalities (Gamper-Rabindran and Timmins (2011)).[17] Household sorting could affect school scores, particularly if sorting is income-related and household income can affect students' performance (see for example Blanden, Gregg and Macmillan (2007)). In that case, there could be reverse causality from the dependent variable to the instrument, thus compromising the credibility of my estimates. Moreover, this would bias the estimates away from zero.

As is shown in the first two columns of Table B.9 in Appendix B, there are significant discontinuities in several household characteristics at school admission boundaries. Higher school scores correlate positively with owner-occupation and the fraction of the local population with college degrees. Likewise, scores correlate negatively with fraction of black residents and percentage unemployed. No sorting is observed for household size, percentage of Asian residents or fraction of unoccupied dwellings. This sorting can be problematic if it occurs in response to brownfield location. To avoid this potential problem, I propose two solutions. In the first place, I substitute school value added for the average school score as my instrument. School value-added measures *increases* in standardized test scores between school entry and 11 years of age (between *key-stage 1* and the end of *key-stage 2*, in the English Department for Education terminology). Insofar as student household quality effects are already present at school entry, this instrument will be less affected by household sorting. This can be seen in columns 3 and 4 of Table B.9. There is substantially less sorting with respect to value added measures. Coefficients on household demographics are halved when using school value-added.[18] The p-values of significance tests also experience a substantial increase when using this variable. While some sorting remains, this is reduced significantly. Therefore, comparing estimates using value-added or average scores as instruments can inform us on whether sorting is driving my findings.

Results from estimation using the value added instrument and a PdL site dummy outcome are reported in Table 9. Results from estimation using a PdL fraction outcome are reported in Table B.10 in appendix B. In both cases we observe that coefficients continue to be negative and

[17]This literature has largely focused on the United States, where the definition of brownfield requires some degree of existing or perceived contamination. Perhaps these types of externalities are weaker in the context of this paper, were most sites will feature no reported contamination.

[18]Both variables have been normalized to take value 0 and have a s.d. equal to 1 so the scales are comparable.

significant. Estimates are generally close to the point estimates reported in tables 1 and 2 using average school score as an instrument. This is reassuring. We would expect the bias mentioned above to be reduced substantially if it was present at all. However, point estimates actually move *away* from zero, and not by much, indicating sorting does not drive my second-stage results.

TABLE 9
VALUE-ADDED INSTRUMENT

| | 1000m | 750m | 500m | 250m |
|---|---|---|---|---|
| | PdL Site | PdL Site | PdL Site | PdL Site |
| Log(Price) | -0.090** | -0.102** | -0.101** | -0.111** |
| | (0.035) | (0.040) | (0.044) | (0.049) |
| County Effects | N | N | N | N |
| Supply Controls | N | N | N | N |
| Log(Price) | -0.095** | -0.112** | -0.111** | -0.122** |
| | (0.039) | (0.044) | (0.047) | (0.053) |
| County Effects | Y | Y | Y | Y |
| Supply Controls | N | N | N | N |
| Log(Price) | -0.100** | -0.118** | -0.118** | -0.139** |
| | (0.043) | (0.049) | (0.051) | (0.063) |
| Observations | 1350476 | 1044667 | 720563 | 375471 |
| County Effects | Y | Y | Y | Y |
| Supply Controls | Y | Y | Y | Y |

*Notes:* 2SLS estimates using school value-added as an instrument for prices. Dependent variable is a dummy taking value 1 if there is a previously-developed site in the grid cell. Boundary fixed-effects as well as controls for distance to the boundary on either side, latitude and longitude included in all specifications. County effects included for second and third row of estimates. Supply-shifters included as controls in the third row of estimates. S.E. clustered at the boundary level in parentheses.

As a final check, I have estimated equation 2 by using the initial instrument and controlling for census characteristics of households in each grid cell. Results are essentially unaltered, again showing that *contemporaneous* household sorting appears not to be affecting my main estimates.[19]

### *6.2. Bandwidth Choice & Functional Form*

A potential concern regarding the robustness of the results in section 4 relates to the choice of bandwidths. Several methods are available to determine one *optimal* bandwidth, following the bandwidth selection literature that has emerged in the last decade. However, it is also

[19]Results not reported but available upon request.

straightforward to reproduce results for different bandwidths close to the threshold and observe whether estimated effects experience substantial changes. I follow the latter approach here. Panel A of Figure 4 displays the different coefficients obtained by estimating the baseline specification in section 3 under several different bandwidths. We observe the coefficients are stable and slightly above 0.08, as reported in Table 1. The coefficient is similar even for smallest bandwidth of only 100 metres. I conclude that baseline results are robust to bandwidth choice.

Panel B contains a similar graph for the coefficients obtained when using matching across the boundary. Again, the estimated coefficients are negative and significant as expected. The point estimates appear to be smaller for shorter boundaries, but the difference is generally not significant.

Finally, panel C displays the coefficients measuring the effect of log prices on the probability of a Land Use Change in the period 2007-2011, as discussed in section 3.3. The coefficients are fairly stable and between 0.05 and 0.1, with point estimates for smaller bandwidths being slightly larger. It is worth noting that this increase in point estimates is accompanied by noticeable differences in the size of confidence intervals, which grow substantially for smaller bandwidths. This is in all likelihood a consequence of the relatively smaller sample sizes.[20] Given that most changes in the significance of coefficients in panel C of Figure 4 are a consequence of larger confidence intervals and not of smaller point estimates, I conclude that the weak significance of coefficients in Table 4 are the result of lower statistical power only.

I next evaluate the robustness of my estimates regarding the choice of functional form restrictions implicit in the baseline analysis. In all specifications in the paper using a dummy outcome, the second-stages can be interpreted as a linear probability model. As has been argued in Horrace and Oaxaca (2006), OLS estimates of linear probability models can exhibit significant bias when the average of the dummy dependent variable is either close to 0 or close to 1. In my case, the probability of having PdL Site $= 1$ is only 1.46%, so I cannot rule out that a significant small sample bias may exist. In order to check the relevance of this possible bias on my results, I re-estimate the baseline specification using the two-step estimator in Newey (1987), which specifies the second stage as a probit model. In this way, I avoid the somewhat problematic linear specification of the conditional distribution of a limited dependent variable. Results are provided in Table B.11 of appendix B. We observe that the probit specification still leads to a significant effect of prices on the presence of brownfield sites.

[20] Recall that my estimates for land use changes are obtained restricting my sample to grid cells containing brownfield land only.

*6.3. Measurement Error*

Throughout the analysis, data for brownfield locations in 2007 is obtained from the NLUD-PDL, as discussed in section 2. This dataset records the geolocation and surface area of each brownfield site as reported by local authorities. It is likely that there is some degree of misreporting, which adds to potential issues with the spatial imputation of PdL sites to grid cells. In addition, as described in section 2, the housing price information is derived from data on housing transactions which is imputed on grid cells based on spatial location. The spatial imputation mechanism necessarily induces some measurement error in our price measure. Moreover, the data from the mortgage provider may not be representative of all transactions, which could result in further measurement error in the instrumented variable.

I expect these sources of measurement error to have a limited impact on estimated effects. Regarding the problem with imputing brownfield land to cells, this would lead to measurement error in the dependent variable only. Under classical measurement error this would only affect precision, which appears not to be an issue with most of my estimates. Regarding measurement error in prices, because I use an IV strategy, I do not expect this to have a substantial effect on results.

That being said, I provide two different robustness checks to ensure measurement error is not a significant problem in this context. In the first place, I validate the PdL site location data by obtaining alternative estimates using the 2010 edition of the NLUD. The 2010 data was part of a three year effort to have a consistent PdL atlas after the last NLUD-PDL edition in 2007. The broad methodology of data collection based on local authorities is similar but significant efforts were made to increase data accuracy.[21] Estimates using the 2010 brownfield measures are provided in panel B of Table 10 in Appendix B. Estimates of price sensitivity are similar to those obtained with the 2007 measures, negative and significant across specifications. The point estimates are slightly larger than for the 2007 data but all fall comfortably within the corresponding 95% C.I..

In order to evaluate the potential measurement error in prices, I also obtain baseline estimates using Land Registry prices. These are spatially imputed as above but, because they include all transactions in England and Wales, the distances involved in spatial imputation are significantly lower. On the other hand, the amount of housing characteristics present in the Land Registry data is quite limited so the hedonic adjustment is not as thorough as when us-

[21] The metadata for the 2010 NLUD-PDl reads "In 2010, detailed reviews of current site intelligence for brownfield sites took place in several local authorities across England to improve the accuracy, currency and completeness of data. "

ing the Nationwide dataset. Panel C of Table 10 indicates our estimates are still negative and statistically significant. Again, point estimates are almost exactly the same as those reported in panel A using the 2007 NLUD-PDL, and the Nationwide prices.

I conclude from this discussion and the associated estimates that measurement error is unlikely to have a substantial effect on my results.

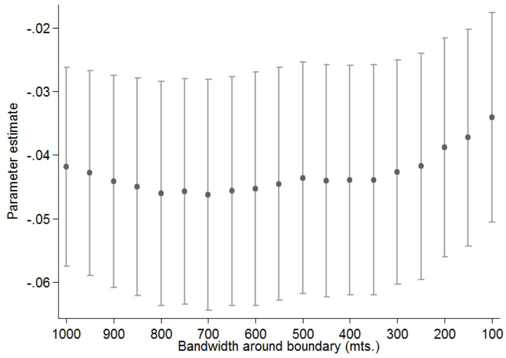
Table 10

Measurement

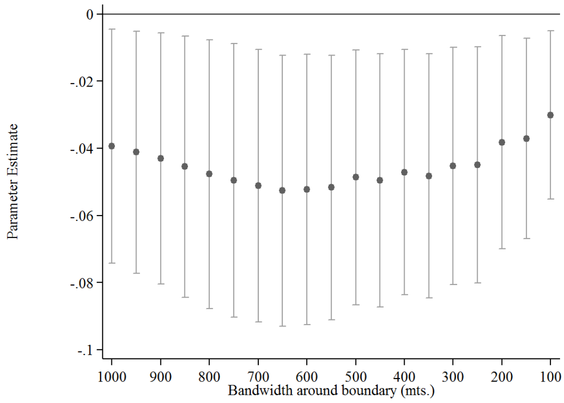
| | Δ PdL Fract. | Δ PdL Fract. | Δ PdL Fract. | Δ PdL Fract. |
|---|---|---|---|---|
| **A. Matched Estimates** | | | | |
| Δ Log(Price) | -0.027** | -0.034** | -0.037** | -0.037** |
| | (0.012) | (0.014) | (0.015) | (0.016) |
| **B. Matched Estimates for 2010 PdLs** | | | | |
| Δ Log(Price) | -0.032** | -0.038** | -0.042*** | -0.043*** |
| | (0.013) | (0.015) | (0.016) | (0.016) |
| **C. Matched Estimates using Land Registry Prices** | | | | |
| Δ Log(Price) (LR) | -0.023** | -0.030** | -0.035** | -0.038** |
| | (0.010) | (0.012) | (0.014) | (0.016) |
| Observations | 600922 | 464178 | 318984 | 165676 |
| Bandwidth | 1000m | 750m | 500m | 250m |
| Controls | Y | Y | Y | Y |
| Boundary Effects | Y | Y | Y | Y |

*Notes:* Panel A reproduces panel D of Table 2 for comparison purposes. In panel B, the dependent variable is a dummy taking value 1 if there is previously developed land in a grid cell according to the 2010 NLUD-PDL database. In panel C, the instrumented variable is replaced for imputed prices based on Land Registry Transactions. In panels A and C, the dependent variable is the fraction of previously-developed land in the grid cell according to the 2007 NLUD-PDL database. All specifications include boundary effects and controls for supply conditions. S.E. clustered at the boundary level in parentheses.

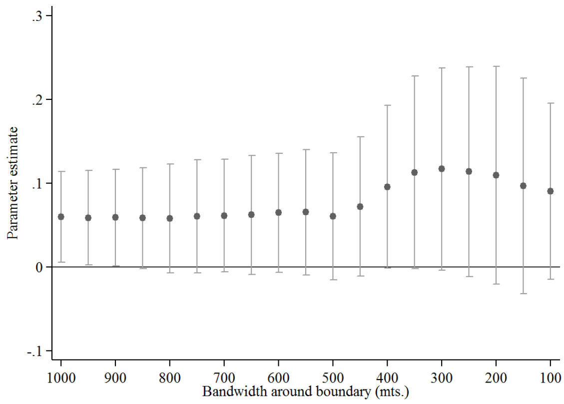
## Figure 4
### Coefficients by Bandwidth




(A) Baseline Estimates for Different Bandwidths




(B) Matched Estimates for Different Bandwidths




(C) Land Use Change Estimates for Different Bandwidths

*Notes:* Panel A: Baseline estimates for different bandwidths using a continuous outcome (see section 3). Panel B: Matching estimates for different bandwidths using a continuous outcome (see section 3.2). Panel C: Land use change estimates for different bandwidths using a continuous outcome (see section 3.3). Parameter values indicated in the vertical axis. Bandwidths indicated in the horizontal axis. Baseline and land use change specifications include county fixed-effects. Matching specification includes boundary fixed-effects. 95% confidence intervals represented in vertical lines.

## 7. Conclusion

This paper provides the first estimates of the price sensitivity of vacant or idle land re-development. I find that high demand reduces the prevalence of brownfield land in cities. Estimates indicate that a 1% increase in local prices reduces the number of hectares with brownfield land in 0.07 percentage points. Given that only about 1.5% of hectares in the sample contain PdL sites, this figure is economically significant. I find substantial effects of prices on ex-post land use changes in 2007 sites for the period 2007-2011. Therefore, I conclude that the presence of vacant or idle land is clearly sensitive to demand conditions, with price differences inducing re-development in the long run. Market forces are responsible for a substantial amount of re-development of derelict or vacant sites.

The results provided here are also relevant to understand the potential effect that remediation relief, tax breaks and price growth can have on re-development of vacant, derelict or underutilized sites. Estimates indicate that increased demand-side incentives can have moderate effects on re-development, especially in cities. I also find that this is the case regardless of the previous land use of these sites, with relatively higher cost sites such as those previously used in manufacturing or mining also being sensitive to demand conditions. Finally, I provide suggestive evidence indicating that planning restrictiveness has a negative impact on the price elasticity of re-development.

A more systematic decomposition of the price elasticity of housing supply into its different sources is an interesting avenue for future research. Estimates such as the ones presented here could be provided for changes in supply resulting from changing building heights, or development in greenfield land. It may also be possible to account for externalities across sites and estimate their influence on re-development. Finally, the empirical strategy provided here can be implemented to study shifts in land use between commercial, industrial and residential.

## References

**Adams, David, Christopher De Sousa, and Steven Tiesdell.** 2010. "Brownfield development: A comparison of North American and British approaches." *Urban Studies*, 47(1).

**Ahlfeldt, Gabriel M, and Elisabetta Pietrostefani.** 2019. "The economic effects of density: A synthesis." *Journal of Urban Economics*, 111: 93–107.

**Bayer, Patrick, Fernando Ferreira, and Robert McMillan.** 2007. "A unified framework for measuring preferences for schools and neighborhoods." *Journal of political economy*, 115(4): 588–638.

**Black, Sandra E.** 1999. "Do better schools matter? Parental valuation of elementary education." *The Quarterly Journal of Economics*, 114(2): 577–599.

**Blanden, Jo, Paul Gregg, and Lindsey Macmillan.** 2007. "Accounting for intergenerational income persistence: noncognitive skills, ability and education." *The Economic Journal*, 117(519).

**Brueckner, Jan K.** 1980. "A vintage model of urban growth." *Journal of Urban Economics*, 8(3): 389–402.

**Brueckner, Jan K.** 1987. "The structure of urban equilibria: A unified treatment of the Muth-Mills model." *Handbook of regional and urban economics*, 2: 821–845.

**Burchfield, Marcy, Henry G Overman, Diego Puga, and Matthew A Turner.** 2006. "Causes of sprawl: A portrait from space." *The Quarterly Journal of Economics*, 121(2).

**Charles, Suzanne Lanyi.** 2014. "The spatio-temporal pattern of housing redevelopment in suburban Chicago, 2000–2010." *Urban Studies*, 51(12): 2646–2664.

**Cheshire, Paul.** 2014. "Turning houses into gold: Don't blame the foreigners, it's we Brits who did it." *Centrepiece*, 19(1): 14–18.

**Cheshire, Paul C.** 2006. "Resurgent cities, urban myths and policy hubris: what we need to know." *Urban studies*, 43(8): 1231–1246.

**Cohen, Rachel, Javier Garciadiego, Tyler Gumpright, Rose Martinez, and Sam Levy.** 2016. "Evaluation of the New York State Brownfield Opportunity Areas Program." NYU Robert F. Wagner Graduate School of Public Service.

**Comission, European.** 2013. “Science for the Environment Policy. Thematic Issue: Brownfield Regeneration.” Directorate-General Environment, European Comission.

**Cunningham, Christopher R.** 2006. “House price uncertainty, timing of development, and vacant land prices: Evidence for real options in Seattle.” *Journal of Urban Economics*, 59(1).

**Cunningham, Christopher R.** 2007. “Growth controls, real options, and land development.” *The Review of Economics and Statistics*, 89(2): 343–358.

**Dell, Melissa.** 2010. “The persistent effects of Peru's mining mita.” *Econometrica*, 78(6).

**Dixon, Timothy, and David Adams.** 2008. “Housing supply and brownfield regeneration in a post-Barker world: is there enough brownfield land in England and Scotland?” *Urban Studies*, 45(1): 115–139.

**Duranton, Gilles, and Matthew A Turner.** 2016. “Urban form and driving: Evidence from US cities.” *Processed, Wharton School, University of Pennsylvania.*

**Dye, Richard F, and Daniel P McMillen.** 2007. “Teardowns and land values in the Chicago metropolitan area.” *Journal of Urban Economics*, 61(1): 45–63.

**Gamper-Rabindran, Shanti, and Christopher Timmins.** 2011. “Hazardous waste cleanup, neighborhood gentrification, and environmental justice: Evidence from restricted access census block data.” *The American Economic Review: P&P*, 101(3): 620–624.

**Gamper-Rabindran, Shanti, and Christopher Timmins.** 2013. “Does cleanup of hazardous waste sites raise housing values? Evidence of spatially localized benefits.” *Journal of Environmental Economics and Management*, 65(3): 345–360.

**Gibbons, Stephen, Stephen Machin, and Olmo Silva.** 2013. “Valuing school quality using boundary discontinuities.” *Journal of Urban Economics*, 75: 15–28.

**Glaeser, Edward, and Joseph Gyourko.** 2018. “The economic implications of housing supply.” *Journal of Economic Perspectives*, 32(1): 3–30.

**Glaeser, Edward L, and Joseph Gyourko.** 2005. “Urban decline and durable housing.” *Journal of Political economy*, 113(2): 345–375.

**Glaeser, Edward L, Joseph Gyourko, and Albert Saiz.** 2008. “Housing supply and housing bubbles.” *Journal of Urban Economics*, 64(2): 198–217.

**Glaeser, Edward L, Joseph Gyourko, and Raven E Saks.** 2006. "Urban growth and housing supply." *Journal of Economic Geography*, 6(1): 71–89.

**Green, Richard K, Stephen Malpezzi, and Stephen K Mayo.** 2005. "Metropolitan-specific estimates of the price elasticity of supply of housing, and their sources." *The American Economic Review*, 95(2): 334–339.

**Greenstone, Michael, and Justin Gallagher.** 2008. "Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program." *The Quarterly Journal of Economics*, 951–1003.

**Gyourko, Joseph, and Raven Molloy.** 2015. "Regulation and housing supply." In *Handbook of regional and urban economics*. Vol. 5, 1289–1337. Elsevier.

**Higgins, Jessica.** 2008. "Evaluating the Chicago Brownfields Initiative: The Effects of City-Initiated Brownfield Redevelopment on Surrounding Communities." Northwestern Journal of Law & Social Policy.

**Hilber, Christian.** 2017. "The UK Planning System-Fit for Purpose." *Birkbeck L. Rev.*, 5: 121.

**Hilber, Christian AL, and Wouter Vermeulen.** 2016. "The Impact of Supply Constraints on House Prices in England." *The Economic Journal*, 126(591): 358–405.

**Hornbeck, Richard, and Enrico Moretti.** 2015. "Who Benefits from Productivity Growth? The Local and Aggregate Impacts of Local TFP Shocks on Wages, Rents, and Inequality." Mimeo.

**Horrace, William C, and Ronald L Oaxaca.** 2006. "Results on the bias and inconsistency of ordinary least squares for the linear probability model." *Economics Letters*, 90(3): 321–327.

**Kane, Thomas J, Stephanie K Riegg, and Douglas O Staiger.** 2006. "School quality, neighborhoods, and housing prices." *American law and economics review*, 8(2): 183–212.

**Kiel, Katherine A, and Michael Williams.** 2007. "The impact of Superfund sites on local property values: Are all sites the same?" *Journal of Urban Economics*, 61(1): 170–192.

**Mace, Alan, Fanny Blanc, Ian Gordon, and Kath Scanlon.** 2016. "A 21st century metropolitan green belt."

**Malpezzi, Stephen, and Duncan Maclennan.** 2001. "The long-run price elasticity of supply of new residential construction in the United States and the United Kingdom." *Journal of Housing Economics*, 10(3): 278–306.

**McMillen, Daniel.** 2017. "Teardowns and Demolitions in Chicago, 2000–2014: A Conditionally Parametric Approach." *Revue économique*, 68(3): 511–535.

**Newey, Whitney K.** 1987. "Efficient estimation of limited dependent variable models with endogenous explanatory variables." *Journal of Econometrics*, 36(3): 231–250.

**OECD.** 2012. *Compact City Policies: A Comparative Assessment.* Matsumoto, Tadashi and Sanchez-Serra, Daniel and Ostry, Adam.

**Ooi, Joseph TL, and Thao TT Le.** 2013. "The spillover effects of infill developments on local housing prices." *Regional Science and Urban Economics*, 43(6): 850–861.

**Owens, Raymond E, Esteban Rossi-Hansberg, and Pierre-Daniel G Sarte.** 2017. "Rethinking Detroit." *Unpublished Paper.*

**Paciorek, Andrew.** 2013. "Supply constraints and housing market dynamics." *Journal of Urban Economics*, 77: 11–26.

**Pagano, Michael A, and Ann O'M Bowman.** 2000. *Vacant land in cities: An urban resource.* Center on Public and Metropolitan Policy.

**Rosenthal, Stuart S, and Robert W Helsley.** 1994. "Redevelopment and the urban land price gradient." *Journal of Urban Economics*, 35(2): 182–200.

**Saiz, Albert.** 2010. "The geographic determinants of housing supply." *The Quarterly Journal of Economics*, 125(3): 1253–1296.

**Salet, Willem.** 2014. "The ladder of sustainable urbanization." *disP-The Planning Review*, 50(4): 4–5.

**Titman, Sheridan, et al.** 1985. "Urban land prices under uncertainty." *American Economic Review*, 75(3): 505–514.

**Van der Krabben, Erwin, and Harvey M Jacobs.** 2013. "Public land development as a strategic tool for redevelopment: Reflections on the Dutch experience." *Land Use Policy*, 30(1): 774–783.

**Wheaton, William C, et al.** 1982. “Urban spatial development with durable but replaceable capital.” *Journal of Urban Economics*, 12(1): 53–67.

**Womack, Kiplan S.** 2015. “Real Options and Urban Land Values: A Review of the Literature.” *Journal of Real Estate Literature*, 23(1): 53–64.

**Zahirovich-Herbert, Velma, and Karen M Gibler.** 2014. “The effect of new residential construction on housing prices.” *Journal of Housing Economics*, 26: 1–18.

# Online Appendices

## A. Data Sources and Dataset Assembly

### *A.1. Data Sources*

Data on previously-developed land sites is obtained from the **National Land Use Database of Previously Developed Land** (NLUD-PDL). Most of the analysis uses the 2007 version of the database which was published by the Department of Communities and Local Government in 2008, and is currently held at the UK National Archives. Later versions of the database were released in 2010, 2011 and 2012. I use 2010 in a validation exercise in section 6.3.

Data on housing transactions is obtained from **Nationwide**, a British building society and one of the largest providers of household mortgages in the United Kingdom. The advantage of this dataset lies in that it includes detailed housing characteristics which allow to control for structural attributes of the property in a hedonic regression before spatial imputation of prices to grid cells. Section 6.3 shows that, using alternative data from the Price Paid database made public by the **Land Registry** leads to comparable results.[22]

Data on school performance is obtained from the **school performance tables**, made available by the Department of Education at https://www.gov.uk/school-performance-tables. I use school performance tables for the years 2002 to 2006. These includes several measures of school performance for primary schools in England and Wales such as average standardized test scores or school level measures of value-added.

Data on land use changes is obtained from the **Land Use Change Database** (LUCS). This data is available since 1985 and 2011. While land use change statistics exist for the period after 2011, there was a substantial methodological break in the regular surveys that year. That is why I focus on land changes in the period 2007-2011 only in section 3.3. Trends in land use changes for the sample period can be found at "Land Use Change Statistics in England: 2011" published by the Department for Communities and Local Government. [23]

Variables for potential supply shifters, used as controls in most specifications and in the balancing tests displayed in Table B.2 and Figure 1 are obtained from different sources. The data on elevation above sea level is based on a combination of Ordinance Survey Terrain 50 which records elevation data for the British territory in a 50 metre grid, imputed to postcodes

[22]Data produced by Land Registry © Crown copyright 2015.
[23]See https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/267551/LUCS_Stats_Release__Dec_2013_FINAL_.pdf.

based on their centroids. The postcode grid-cell match in my dataset is based on spatial assignment based on postcode centroids. The landslide risk measure and the data on underground aquifers are obtained from the **British Geological Survey**. In the case of the landslide risk, it is specifically obtained from the GeoSure 5km Hexagonal Grid. The data on aquifers is obtained from the Hydrogeogology map (scale 1:625,000) and spatially matched with grid cell centroids. The aquifer variable used here takes value 1 if the grid cell centroid falls in an aquifer identified as "highly productive". Data on agricultural land quality is based on the 1988 Agricultural Land Classification of England and Wales, elaborated by the Ministry of Agriculture, Fisheries and Food and the Welsh Office Agriculture Department at the time. Data on the location of registered parks and gardens is obtained from the Historic England shapefile recording these locations. Historic England is a public body devoted to caring about England's historic environment. Spatial imputation of parks and gardens to grid cells is again based on grid cell centroids. Data on planning application average refusal rates at the local planning authority level for the period 1979-2008 obtained from Hilber and Vermeulen (2016).

*A.2. Dataset Assembly*

The dataset assembly process relied heavily on combining spatial data using Geographic Information Systems. In the first place, I use a shapefile of counties (polygons) to obtain a set of county boundaries (lines). Boundaries with the sea or boundaries with Wales or Scotland are removed from this set. I build a buffer area of 1km around the remaining boundaries and create a grid of hectares within those buffer areas. Only whole hectares for which the distance between their centroid and the corresponding boundary is less than 1 km are kept in the analysis. Furthermore, I remove from the analysis all grid cells that are crossed by a county boundary. The resulting grid constitutes my sample, with each hectare-cell being one observation (see panel A of Figure A.1). Each cell is identified with its closest boundary.

The next step is to impose the location of brownfield sites. As discussed above, these are created as circular polygons around the centroids reported in the NLUD-PDL data. An overlay identifies which grid cells contain brownfield land and which do not (see panel B of Figure A.1). In addition, I compute the fraction of each grid cell that is covered by these circular polygons. This provides the continuous measure of PdL in a grid cell.

I next impute the hedonic-filtered prices to each hectare. In the first place, I run a regression of the logarithm of each observed transaction prices on the property's characteristics (floor area, number of bedrooms, number of bathrooms, garage presence dummy and central heating dummy). For each transaction I recover the residual and add back the constant term to obtain a measure of price after accounting for unit characteristics. Next, I impute this price to grid cells based on closeness within the corresponding county. Recall county boundaries operate as school admission boundaries in the United Kingdom (see panel C of Figure A.1). A total of 118,000 transactions for the period 2002-2006 are used in the imputation process when using the Nationwide data. Importantly, school scores are not used in the imputation of prices to grid cells. Moreover, complementary results indicate that distance between a cell and its imputed transaction do not change at the boundary (not shown).

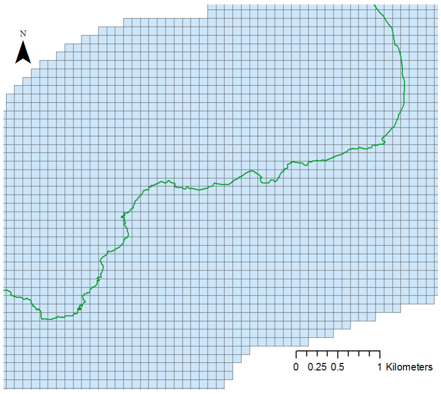
Finally, I impute school scores to grid cells by attaching to each cell the average score and value-added measure of the closest school on its side of the boundary.

Other variables such as those recording potential supply shifters or census characteristics are imputed using grid cell centroids.

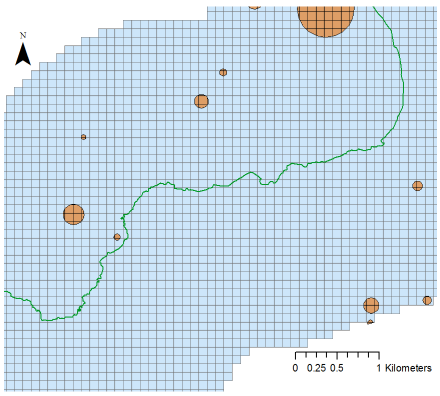
As a sanity check to show the dataset assembly process proceeded correctly, I show that latitude and longitude vary smoothly across boundaries, and that school scores jumps when moving from the low performance to the high performance side of the county border. The graphs, displayed in Figure A.2 of Appendix B, merely show that the dataset has been adequately constructed. As expected, both latitude and longitude vary smoothly at the boundary and school performance exhibits a sharp discontinuity, with test scores in the *good* side exhibiting being 0.6 of an s.d. higher than in the *bad* side.

FIGURE A.1

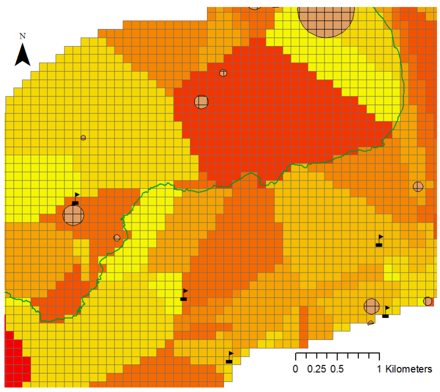
DATASET ASSEMBLY




(A) COUNTY BOUNDARY AND HECTARE GRID

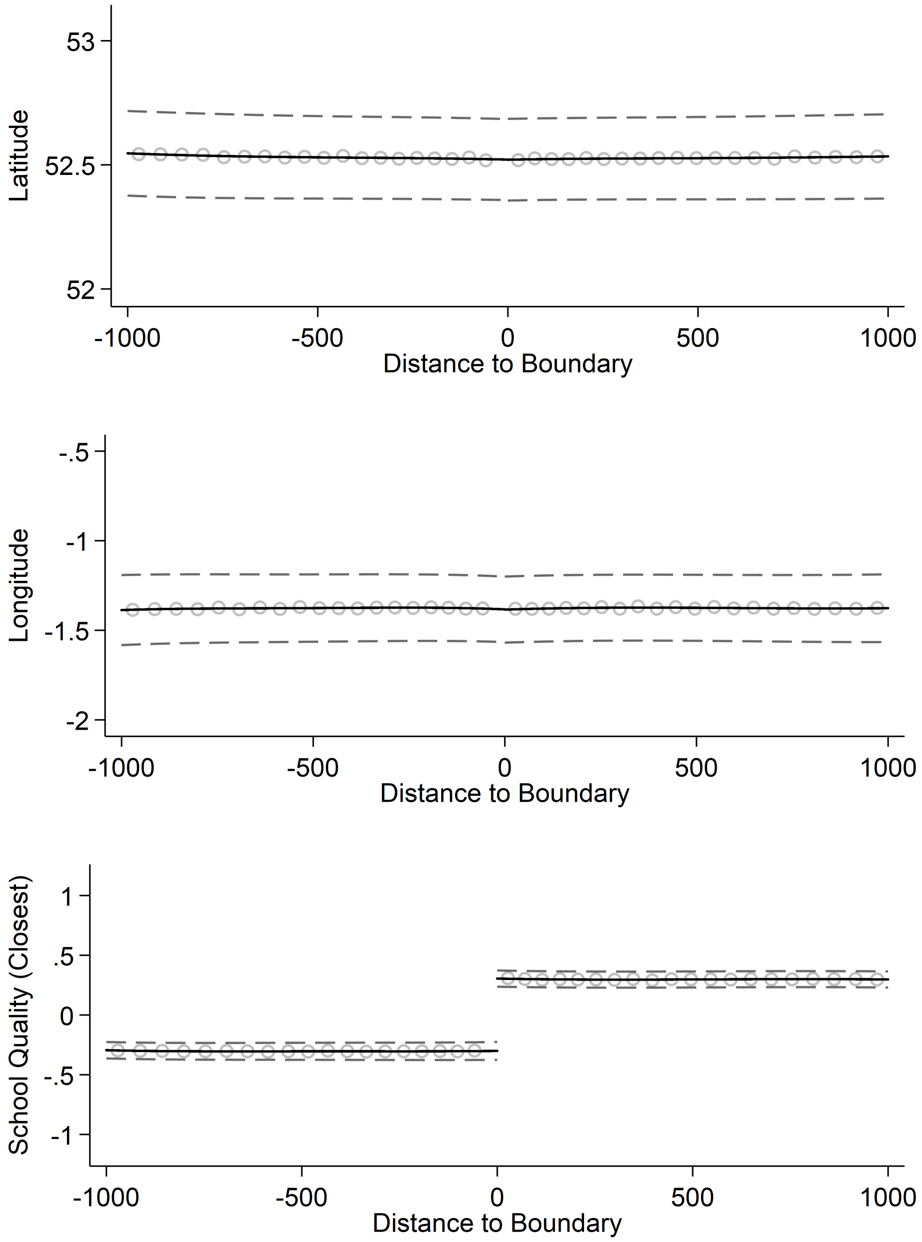



(B) OVERLAY OF BROWNFIELD LOCATIONS




(C) IMPUTING SMOOTHED HOUSING PRICES AND SCHOOLS

**FIGURE A.2**
COVARIATE BALANCE (SANITY-CHECK)




*Notes:* Horizontal axis represents distance to the boundary with negative distances corresponding to the county with low average test scores and positive distances corresponding to the county with high average test scores. Vertical axis corresponds to latitude (top panel), longitude (middle panel) and closest school test scores (bottom panel). Third degree polynomials fitted on the raw data represented in solid lines on either side of the boundary. 95% confidence intervals in dashed lines, with standard errors clustered at the boundary level. Gray circles correspond to averages taken within 40 distance bins.

## B. Additional Tables and Figures

This Appendix presents a series of tables and figures complementing those in the main text. These includes descriptive tables, matching or baseline estimates and results for alternative outcome variables.

TABLE B.1

DESCRIPTIVES

| | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|
| Any Brownfile in this cell (Dummy) 2007 | 0.015 | 0.120 | 0 | 1 |
| Fraction Brownfiled in this cell (2007) | 0.006 | 0.071 | 0 | 1 |
| LUCS Site. To Residential | 0.021 | 0.142 | 0 | 1 |
| School Score | 27.956 | 1.448 | 21.500 | 31.850 |
| Latitude | 52.529 | 1.127 | 50.332 | 55.188 |
| Longitude | -1.377 | 1.095 | -4.549 | 1.739 |
| Elevation above sea level | 96.423 | 84.406 | -10.000 | 570.000 |
| House Price (Smoothed) | 141703 | 56145 | 13268 | 1477636 |
| Population (2001 census) | 1588 | 310 | 1000 | 4569 |
| Observations | | 1,566,798 | | |

*Notes:* Descriptive statistics for the full sample. Grid cells (hectares) within 1 km of a county (school admission) boundary.

## Table B.2

Covariate Balance Estimates of Determinants of Housing Supply Across Boundaries

| | 1000m | 750m | 500m | 250m |
|---|---|---|---|---|
| | Elevation | Elevation | Elevation | Elevation |
| Δ School Qty. | -0.269 | -0.335 | -0.481 | -0.554 |
| | (0.719) | (0.710) | (0.701) | (0.690) |
| | Landslide | Landslide | Landslide | Landslide |
| Δ School Qty. | -0.002 | -0.002 | -0.001 | -0.001* |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| | Aquifer | Aquifer | Aquifer | Aquifer |
| Δ School Qty. | 0.002 | 0.002 | 0.002 | 0.002 |
| | (0.002) | (0.002) | (0.002) | (0.001) |
| | Ag. Quality | Ag. Quality | Ag. Quality | Ag. Quality |
| Δ School Qty. | 0.001 | 0.000 | -0.000 | -0.001 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| | Refusal Rate | Refusal Rate | Refusal Rate | Refusal Rate |
| Δ School Qty. | 0.001 | 0.001 | 0.001 | 0.001 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| | Parks | Parks | Parks | Parks |
| Δ School Qty. | 0.001 | 0.001 | 0.001 | 0.001 |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| Observations | 789986 | 610861 | 420880 | 219202 |

*Notes:* Grid cell level regressions using differenced estimates based on matched grid pairs. Dependent variable is the spatial difference in each variable as indicated in the row heading. The variables are elevation above sea level (in metres), landslide risk (dummy taking value 1 if moderate or high), fraction of planning application refused, a dummy taking value 1 if agricultural land quality is high and a dummy taking value 1 if the grid cell contains a park or garden. Columns 1 through 4 correspond to bandwidths of 1000, 750, 500 and 250 metres, respectively. All specifications include boundary fixed-effects to account for differenced county effects. S.E. clustered at the boundary level in parentheses.

Table B.3

First-Stage Baseline Estimates

| | 1000m | 750m | 500m | 250m |
|---|---|---|---|---|
| | Log(Price) | Log(Price) | Log(Price) | Log(Price) |
| School Score | 0.061*** | 0.061*** | 0.060*** | 0.059*** |
| | (0.005) | (0.006) | (0.006) | (0.006) |
| County Effects | N | N | N | N |
| Controls | N | N | N | N |
| School Score | 0.058*** | 0.058*** | 0.058*** | 0.057*** |
| | (0.005) | (0.006) | (0.006) | (0.006) |
| County Effects | Y | Y | Y | Y |
| Controls | N | N | N | N |
| School Score | 0.055*** | 0.055*** | 0.055*** | 0.055*** |
| | (0.006) | (0.006) | (0.006) | (0.006) |
| Observations | 1350047 | 1044354 | 720359 | 375375 |
| County Effects | Y | Y | Y | Y |
| Controls | Y | Y | Y | Y |

*Notes:* Grid-cell level regressions. Dependent variable is the log of housing prices imputed to a hectare. School score normalized to have mean 0 and s.d. equal to 1. Boundary fixed-effects and separate linear terms for distance to the boundary on either side, latitude and longitude included in all specifications. Other fixed-effects and controls as indicated in the table. S.E. clustered at the boundary level in parentheses.

## Table B.4

### First-Stage Matched Estimates

| | 1000 | 750 | 500 | 250 |
|---|---|---|---|---|
| | Log(Price) | Log(Price) | Log(Price) | Log(Price) |
| Δ School Score (closest) | 0.050*** | 0.050*** | 0.050*** | 0.048*** |
| | (0.01) | (0.01) | (0.01) | (0.01) |
| F-Stat | 67 | 59 | 53 | 47 |
| Boundary Effects | N | N | N | N |
| Controls | N | N | N | N |
| | Log(Price) | Log(Price) | Log(Price) | Log(Price) |
| Δ School Score (closest) | 0.041*** | 0.041*** | 0.041*** | 0.040*** |
| | (0.006) | (0.007) | (0.007) | (0.007) |
| F-Stat | 41 | 36 | 33 | 29 |
| Boundary Effects | Y | Y | Y | Y |
| Controls | N | N | N | N |
| | Log(Price) | Log(Price) | Log(Price) | Log(Price) |
| Δ School Score (closest) | 0.040*** | 0.040*** | 0.039*** | 0.039*** |
| | (0.007) | (0.007) | (0.008) | (0.008) |
| F-Stat | 33 | 28 | 25 | 23 |
| Observations | 600922 | 464178 | 318984 | 165676 |
| Boundary Effects | Y | Y | Y | Y |
| Controls | Y | Y | Y | Y |

*Notes:* Grid-cell level regressions after differencing within matched pairs. Dependent variable is the difference in log of housing prices. School score normalized to have mean 0 and s.d. equal to 1. Linear terms for distance to the boundary on either side as well as latitude and longitude included in all specifications. Boundary effects included for second and third row of estimates. Differenced supply-shifters included as controls in the third row of estimates. S.E. clustered at the boundary level.

### Table B.5
Reduced-Form Estimates

| | 1000m | 750m | 500m | 250m |
|---|---|---|---|---|
| | PdL Fraction | PdL Fraction | PdL Fraction | PdL Fraction |
| School Score | -0.005*** | -0.006*** | -0.005*** | -0.005*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| County Effects | N | N | N | N |
| Controls | N | N | N | N |
| School Score | -0.005*** | -0.006*** | -0.005*** | -0.005*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| County Effects | Y | Y | Y | Y |
| Controls | N | N | N | N |
| School Score | -0.004*** | -0.005*** | -0.004*** | -0.004*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Observations | 1350047 | 1044354 | 720359 | 375375 |
| County Effects | Y | Y | Y | Y |
| Controls | Y | Y | Y | Y |

*Notes:* Grid-cell level regressions. Dependent variable is a dummy taking value 1 if there is a PdL site in the grid cell. School performance is measured using average test scores of the closest primary school within the county and normalized to have mean 0 and s.d. equal to 1. Boundary fixed-effects and linear terms for distance to the boundary on either side, as well as latitude and longitude, included in all specifications. County effects included for second and third row of estimates. Supply-shifters included as controls in the third row of estimates. S.E. clustered at the boundary level in parentheses.

### Table B.6
Spatial-Matching Estimates for Unitary Authority Boundaries

| | 1000 | 750 | 500 | 250 |
|---|---|---|---|---|
| | Δ PdL Site | Δ PdL Site | Δ PdL Site | Δ PdL Site |
| Δ Log(Price) | -0.066** | -0.063* | -0.056* | -0.056* |
| | (0.033) | (0.034) | (0.033) | (0.031) |
| F-Stat | 22 | 22 | 20 | 17 |
| Controls | N | N | N | N |
| | Δ PdL Site | Δ PdL Site | Δ PdL Site | Δ PdL Site |
| Δ Log(Price) | -0.063* | -0.063* | -0.053 | -0.052 |
| | (0.035) | (0.037) | (0.037) | (0.034) |
| F-Stat | 20 | 20 | 18 | 15 |
| Observations | 191965 | 167164 | 119622 | 62022 |
| Controls | Y | Y | Y | Y |

*Notes:* Grid cell level regressions using differenced estimates based on matched grid pairs. Sample restricted to boundaries of Unitary Authorities created between 1995 and 2000. Dependent variable is the spatial difference in a dummy taking value 1 if there is previously-developed land in a grid cell. Columns 1 through 4 correspond to bandwidths of 1000, 750, 500 and 250 metres, respectively. S.E. clustered at the boundary level in parentheses.

## Table B.7
### First-Stage: LUCS Sample

| | 1000m | 750m | 500m | 250m |
|---|---|---|---|---|
| | Log(Price) | Log(Price) | Log(Price) | Log(Price) |
| School Score | 0.045*** | 0.045*** | 0.047*** | 0.040*** |
| | (0.009) | (0.009) | (0.010) | (0.012) |
| County Effects | N | N | N | N |
| Controls | N | N | N | N |
| School Score | 0.038*** | 0.034*** | 0.037*** | 0.032** |
| | (0.008) | (0.009) | (0.010) | (0.012) |
| County Effects | Y | Y | Y | Y |
| Controls | N | N | N | N |
| School Score | 0.042*** | 0.041*** | 0.045*** | 0.042*** |
| | (0.007) | (0.008) | (0.009) | (0.011) |
| Observations | 16729 | 12985 | 8861 | 4508 |
| County Effects | Y | Y | Y | Y |
| Controls | Y | Y | Y | Y |

*Notes:* Grid-cell level regressions. Sample restricted to cells containing brownfield land in 2007. Dependent variable is the log of housing prices. School score normalized to have mean 0 and s.d. equal to 1. Boundary fixed-effects and separate linear terms for distance to the boundary on either side as well as latitude and longitude included in all specifications. County effects included for second and third row of estimates. Supply-shifters included as controls in the third row of estimates. S.E. clustered at the boundary level in parentheses.

## Table B.8
### Cities (Fraction)

| | Full Sample | Urban Only | Excl. London | London | 20 Largest |
|---|---|---|---|---|---|
| | PdL Fract. | PdL Fract. | PdL Fract. | PdL Fract. | PdL Fract. |
| Log(Price) | -0.038*** | -0.050*** | -0.033*** | -0.221*** | -0.044*** |
| | (0.009) | (0.014) | (0.010) | (0.085) | (0.015) |
| Matching | N | N | N | N | N |
| | Δ PdL Fract. | Δ PdL Fract. | Δ PdL Fract. | Δ PdL Fract. | Δ PdL Fract. |
| Δ Log(Price) | -0.037** | -0.060** | -0.033** | -0.234 | -0.043* |
| | (0.016) | (0.030) | (0.016) | (0.175) | (0.022) |
| Matching | Y | Y | Y | Y | Y |

*Notes:* 2SLS estimates for different sub-samples including cities and groups of cities. First column corresponds to the full sample, second column excludes the London metropolitan area, third column restricts the sample to cells in the London metropolitan area, and fourth column restricts the sample to the 20 largest English metropolitan areas by size of workforce. Dependent variable is the fraction of land in a grid cell that is covered by brownfield. In all specifications, the bandwidth is 250 metres. First row corresponds to baseline estimates. Second row corresponds to estimates obtained using spatial matching. S.E. clustered at the boundary level in parentheses.

## Table B.9
### Household Sorting by Boundaries

| | | | | |
|---|---|---|---|---|
| | Δ % Owner Occupiers | | | |
| Δ School Av.Score | 1.925*** | 1.917*** | | |
| | (0.000) | (0.000) | | |
| Δ School Value-Added | | | 0.628** | 0.670** |
| | | | (0.014) | (0.011) |
| | Δ % Household Size | | | |
| Δ School Av.Score | 0.011 | 0.009 | | |
| | (0.160) | (0.266) | | |
| Δ School Value-Added | | | 0.003 | 0.002 |
| | | | (0.634) | (0.737) |
| | Δ % Black Residents | | | |
| Δ School Av.Score | -0.106** | -0.090* | | |
| | (0.036) | (0.078) | | |
| Δ School Value-Added | | | -0.042 | -0.032 |
| | | | (0.338) | (0.488) |
| | Δ % Asian Residents | | | |
| Δ School Av.Score | -0.003 | -0.007 | | |
| | (0.938) | (0.883) | | |
| Δ School Value-Added | | | -0.001 | -0.004 |
| | | | (0.980) | (0.930) |
| | Δ % Higher Education | | | |
| Δ School Av.Score | 1.089*** | 1.038*** | | |
| | (0.000) | (0.000) | | |
| Δ School Value-Added | | | 0.600*** | 0.580*** |
| | | | (0.000) | (0.000) |
| | Δ % Unemployed | | | |
| Δ School Av.Score | -0.201*** | -0.197*** | | |
| | (0.000) | (0.000) | | |
| Δ School Value-Added | | | -0.083*** | -0.083*** |
| | | | (0.009) | (0.010) |
| Observations | 335371 | 204143 | 335371 | 204143 |
| Bandiwdth | 500m | 250m | 500m | 250m |
| Matching | Y | Y | Y | Y |

*Notes:* Grid cell level regressions using differenced estimates based on matched grid pairs. Dependent variable in each specification indicated on the top of each row. All estimates obtained using spatial differences within matched pairs across boundaries. Specifications in columns 1 and 2 use average school standardized test scores to measure educational quality. In columns 3 and 4, this is measured with school value-added, as reported in the school performance tables. Columns 1 and 3 correspond to estimates using bandwidths of 500 metres and columns 2 and 4 use bandwidths of 250 metres. All specifications include county effects and control for supply-shifters. **p-values for coefficient significance tests in parentheses.**

## Table B.10
### Value-Added Instrument - Continuous Outcome

| | 1000m | 750m | 500m | 250m |
|---|---|---|---|---|
| | PdL Fraction | PdL Fraction | PdL Fraction | PdL Fraction |
| Log(Price) | -0.038* | -0.047** | -0.050* | -0.065** |
| | (0.019) | (0.023) | (0.026) | (0.030) |
| County Effects | N | N | N | N |
| Controls | N | N | N | N |
| | PdL Fraction | PdL Fraction | PdL Fraction | PdL Fraction |
| Log(Price) | -0.049** | -0.059** | -0.061** | -0.077** |
| | (0.023) | (0.027) | (0.029) | (0.034) |
| County Effects | Y | Y | Y | Y |
| Controls | N | N | N | N |
| | PdL Fraction | PdL Fraction | PdL Fraction | PdL Fraction |
| Log(Price) | -0.051** | -0.061** | -0.066** | -0.085** |
| | (0.025) | (0.029) | (0.032) | (0.039) |
| Observations | 1350476 | 1044667 | 720563 | 375471 |
| County Effects | Y | Y | Y | Y |
| Controls | Y | Y | Y | Y |

*Notes:* Two-stage least square estimates using school value-added as an instrument for prices. Dependent variable measures the fraction of 2007 brownfield land in the grid cell. Boundary fixed-effects and linear terms for distance to the boundary on either side as well as latitude and longitude included in all specifications. County effects included for second and third row of estimates. Supply-shifters included as controls in the third row of estimates. S.E. clustered at the boundary level in parentheses.

## Table B.11
### Robustness Checks - Probit

| | 1000 | 750 | 500 | 250 |
|---|---|---|---|---|
| | PdL Site | PdL Site | PdL Site | PdL Site |
| Log(Price) | -1.904*** | -1.947*** | -1.931*** | -1.863*** |
| | (0.219) | (0.226) | (0.245) | (0.257) |
| Boundary Effects | Y | Y | Y | Y |
| Controls | N | N | N | N |
| | PdL Site | PdL Site | PdL Site | PdL Site |
| Log(Price) | -1.760*** | -1.834*** | -1.777*** | -1.795*** |
| | (0.239) | (0.248) | (0.268) | (0.292) |
| Boundary Effects | Y | Y | Y | Y |
| Controls | Y | Y | Y | Y |

*Notes:* Coefficients obtained by maximum likelihood using boundary fixed-effects. Dependent variable is a dummy taking value 1 if there is a previously-developed site in the grid cell (2007). Columns 1 through 4 correspond to bandwidths of 1000, 750, 500 and 250 metres, respectively. Second row of coefficients obtained including potential supply-shifters as controls. S.E. clustered at the boundary level in parentheses.

### C. Re-development Policies for Brownfield and Previously Developed Sites

Idle or vacant land plots within cities, often containing the remnants of previous developments, are a common feature of cities worldwide. While an internationally harmonized definition for previously developed unused or underused sites is not available, several individual studies bare witness to this fact. In England, Adams, De Sousa and Tiesdell (2010) claim previously developed sites amount to 5.45% of total urban developed land. In the Greater London area alone, there are over 2000 hectares of land identified as brownfield sites with potential for re-development.[24] Estimates from the European Commission indicate there are over 3 million brownfield sites across Europe located and well connected to urban boundaries, with 500,000 hectares of brownfield land estimated to be available for development (Comission, 2013).

In the case of the United States, up to 15% of urban land is classified as vacant (Pagano and Bowman, 2000). The US definition of brownfields is restricted only to property where expansion, re-development or reuse may be complicated by "the presence or potential presence of a hazardous substance, pollutant or contaminant".[25] The EPA estimates that there are over 450,000 brownfields in the USA according to this definition, though some authors increase the figure to over 1 million, covering 6% of urban areas (Adams, De Sousa and Tiesdell, 2010).

There are several types of policies deployed to promote the development of vacant land, previously developed land and contaminated sites within cities. The policies themselves vary substantially by jurisdiction, because of, among other things, the different administrative frameworks applying to these sites. Policy objectives encompass the urban densification, clean-up of contaminated sites (especially, but not exclusively, in the US) and removing financial barriers to re-development. The policies themselves can be classified into four broad categories:

1. Grants and Subsidies such as clean-up grants
2. Financial Instruments such as low interest loans to developers.
3. Public Ownership schemes in which local governments buy the land, conduct part of the re-development efforts and sell it out to developers.
4. Differential Planning schemes aiming to target new developments to previously developed sites.

[24]Calculations based on data from the London Brownfields Sites Review, accesible at https://data.london.gov.uk/dataset/london-brownfield-sites-review.

[25]EPA's Brownfields and Land revitalization Programs. Properties with New Purpose. https://www.epa.gov/sites/production/files/2015-09/documents/oblr_brochure_weblayout_508.pdf

I use this taxonomy to classify the most salient brownfield policies in North America and Europe. This policy review is not meant to be exhaustive. It provides an overview of the policy levers currently in use for this purpose. As I will argue below, many of these policies can be linked, in one form or another, to the sensitivity of re-development to demand conditions.

***Grants and Subsidies***

Grants and subsidies for clean up of pollutants, as well as for re-development of derelict sites are one common mechanism to foster brownfield conversion. Low interest loan can be also seen as a form of subsidy, but will be treated separately. I provide three examples here for the United Kingdom and the United States. Other state and city level programs are available in the USA (e.g. Los Angele's Citywide Brownfields Program, City of Chicago's Brownfield Initiative).

*United Kingdom - Department for Environment, Food & Rural Affairs (DEFRA)*

DEFRA funding had been available for local authorities from 2000 until 2017. The amount of funding available peaked at GBP 17.5 million in 2009-2010 and was gradually faced out until 2017. The funding was made available via small grants with an average value of GBP 38,000. This was directed to clean up and other remediation activities carried out jointly by local authorities and land owners/occupiers who provided 17% of all funding for remediation efforts in this context. The phase out of the program has led to discussions about the ability of English local governments to meet their statutory obligations in aiding the process of land remediation. [26]

*United States - Environmental Protection Agency Grants*

The EPA is the primary enforcer of environmental statutes and regulations in the United States. EPA has launched the Brownfields and Land Revitalization Programs to revitalize contaminated land and return properties to productive use.

There are broadly two types of grants, Assessment grants and Clean Up grants. Both can fund up to 200,000 USD for plans lasting up to 3 years. The Assessment grants are meant to fund evaluation of clean up costs, including detection of hazardous substances on site. Clean up grants apply only to sites owned by the grantee and can only cover up to 80% of the total clean up costs. The EPA also runs an area-wide planning grant program directed to local governments, in their role as planning authorities, which is meant to aid in the development of

[26] For further reference see https://publications.parliament.uk/pa/cm201617/cmselect/cmenvaud/180/18005.htm.

planning processes to assess, clean and reuse brownfield sites.

The 2016 budget of the EPA's Brownfield programs amounted to 110 million USD funding over 151 cooperative programs with municipalities, clean up costs over 142 sites and assessments for over 3000 sites. The EPA also provides technical assistance to lower level government bodies.[27] The assessment and clean-up grants can be seen as subsidies for re-development.

*New York - Brownfield Opportunity Areas (BOA) Program*

The BOA program was launched in 2003 through the New York (NY) State Brownfields Reform Act and is administered by the NY State Department. It is targeted to brownfield re-development in poor communities and provides grants of up to 90% of the eligible project costs to finalise revitalization plans and ultimately lead to brownfield re-development. The goal is to reduce re-development costs by removing uncertainty regarding site conditions, ownership structure or future feasible uses. Eligible applicants are not private developers but rather municipal governments and community-based organizations. That being said, the program also includes a 2% tax credit bonus for proposed development projects on sites that are part of the program. Note that, in terms of net present value of a project, a tax credit operates as a subsidy.

The budget allocated to this program has varies substantially over the years, from 32 million USD in 2011 to 45 million in 2016. The program is still operational. A formal evaluation of the effect of this program on nearby housing prices was conducted in Cohen et al. (2016).

***Financial Instruments***

The supply of appropriate financial instruments to promote brownfield and infill re-development has been a popular policy approach, especially in the European Union. These provide both low-interest conventional loans as well as equity loans and other financial engineering tools.

*European Union - Joint European Support for Sustainable Investment in City Areas (JESSICA)*

JESSICA is a European Comission initiative developed jointly with the European Investment Bank and the Council of Europe development Bank to support urban development and urban regeneration schemes. The initiative is implemented as part of the European Regional Development Fund. The objective of the policy is to provide financial assistance in the form of equity, loans or guarantees channelled through public-private partnerships. This is meant to cover insufficient availability of equity from private investors, or to compensate for low re-

[27] For further referent visit https://www.epa.gov/brownfields/overview-brownfields-program and https://www.epa.gov/sites/production/files/2015-02/documents/fy_2016_bib_combined_v5.pdf.

turns that need to be leveraged to attract private investors (due to contamination of property or poor local infrastructure). Guarantees are meant to enable developers and other private participants to secure funding from third parties (e.g. banks). Potential proceeds from these financing operations are meant to be re-invested in new urban development projects.

*European Union - GINGKO Fund*

European Investment Bank and Edmond de Rothschild Group funding dedicated to acquiring a portfolio of brownfield sites, including also other private investors. The goal of the fund is to signal investment appeal to other private equity funds interested in regeneration and re-development of brownfield sites in its role as an environmental remediation specialist. It conducts supplementary environmental conditions studies before acquisition and analyses economic feasibility of operations to bolster its signalling effect. The fund's activity has been mostly concentrated in Belgium and France where most re-development has been publicly led. The total assets in the fund amounted to € 140 million in 2016. The actions of GINGKO fund can be seen as reducing actual and perceived costs or re-development for investors.[28]

*California Environmental Protection Agency: Brownfields Initiative*

California's environmental authority in charge of restoring, protecting and enhancing the environment launched the initiative to remedy brownfields. The policy consists of a fund meant to provide low cost loans to developers, property owners, NGOs and local government agencies which own brownfield sites with potential for re-development. The budget consists of 2.7 million USD available for low interest loans which operate both as a financing tool and as an effective subsidy given the low rates. Government and non-for profits can only request loans up to USD 200,000.

An example of a development using these funds is the Third Street Project in the South Market Area in San Francisco. The program provided a 1.6 million USD loan used to clean-up lead contamination of the soil in the site.[29]

### ***Public Ownership***

In some jurisdictions, brownfield remediation is carried out in the context of public ownership of the sites themselves. In its simplest form, a public authority simply acquires a site, pays the cost of remediation, and sells it of to developers or other private agents.

[28] For further information please see http://www.eib.org/infocentre/publications/all/eib-information-1-2011-n141.htm.

[29] For further information, see http://www.dtsc.ca.gov/SiteCleanup/Brownfields/Loans_Grants.cfm.

*re-development in the Netherlands*

Dutch land policy is often characterized by an important role of public ownership and command-and-control tools. The Netherlands has a "public land development strategy, [that] involves public purchase, ownership and servicing of land and active planning for land use before land is released for actual development to the private sector. This guarantees building developments according to public policies, it realizes full cost recovery of all public works via the sale of building plots and it captures at least part of the surplus value of the land after a change in use" (Van der Krabben and Jacobs, 2013). This model is also applied to brownfield projects and requires strategic land acquisitions prior to regeneration. Categorized as a public comprehensive top-down model: a public body (usually the municipality) acquires the land for future development, services that land and re-parcels it into building plots that can be sold of for cost recovery and value capture. For further reference, see Van der Krabben and Jacobs (2013), which also discusses the potential risks associated to this type of top-down approach to land use and re-development policies.

An example of a large scale regeneration project carried out by public authorities acting as leaders in the development process can be found in the recent regeneration of areas within the Rotterdam inner-city harbour. The national government provided a €31 million subsidy for this purpose, which adds to other €27 million from a public fund and additional municipal resources. Municipalities, who lead the project, were allowed to modify zoning laws to accommodate this re-development. While the endeavour is publicly controlled, private parties are actively incorporated to cover for the additional costs at an early stage. The municipality and Havenbedrijf Rotterdam first acquire the subsidies, permits, property management, and spatial legislation before the project was commissioned to private developers. The regeneration project includes housing development, revitalising economic activity in the harbours, and investment in the local innovation system.

*Chicago - Brownfields Initiative*

The City of Chicago launched the Brownfields Pilot in 1990 with a 2 million investment from General Obligation Bonds to redevelop brownfield sites. The pilot project was a success, and was leveraged into the Brownfields Initiative in 1993 with additional loan guarantees from the US Department of Housing and Urban Development and funds from the EPA. The city's Departments of Environment, Planning and Development coordinates the program in collaboration with several other organisations. The initiative recycles neglected properties to reuse the land for the creation of green and open areas, affordable housing, office space and economic re-development. The City acquires contaminated sites to add to the city's investment portfolio. After the sites have been assessed, enrolled in Illinois EPA's Site remediation Program and cleaned, the sites are marketed by the City for re-development.

Since 1990 funding for brownfields re-development has been leveraged from several sources: 2 million USD from the General Obligations Bonds, 74 million from the HUD Section 108 Loan Guarantee, $691.000 from the Brownfields Showcase Community Designation, etc. [30]

***Urban Planning Tools***

In some cases planners try to use planning guidelines to provide incentives for brownfield re-development. This can involve relaxation of planning restrictions (as in the case of the Rotterdam harbour mentioned above), tightening of restrictions on planning development or a combination of both.

*United Kingdom - Brownfield First & Green Belts*

The Brownfield First policy was launched in 1998 with the goal of ensuring that 60% of new urban development in the United Kingdom happens within the urban footprint. The argument motivating the policy is that developers do not pay the social cost of greenfield development because local governments are the ones in charge of providing transportation, sanitation and other infrastructure. Presumably, the cost of providing this infrastructure is much lower for properties located within the urban footprint. The policy led local authorities to factor these priorities into their planning guidelines (for example in Local Development Frameworks). By 2008, 80% of developments was happening on previously developed land sites.

Another British planning policy that was crucial in directing new development to PdL sites is the widespread use of urban greenbelts. These were introduced in London in 1935 and generalized for other cities in the 1947 Town and Country Planning Act. This sets out an area - the greenbelt - around cities where development is forbidden. Greenbelts cover roughly 12% of England and are usually placed around urban areas. They are meant to contain sprawl, operate as a sort of urban lung and safeguard the countryside among other uses. Its presence is a highly debated issue, as many authors argue that green belts amount to a tight restriction on urban residential construction which results in high housing prices Cheshire (2014). The combined imposition of greenbelts and the brownfield first policy implies most new development in the United Kingdom effectively happens on previously developed land. For further reference, see

[30] For further reference, see Higgins (2008) and https://www.cityofchicago.org/city/en.html.

Dixon and Adams (2008) and Mace et al. (2016).

*Netherlands - "Ladder for Sustainable Urbanization"*

The Dutch minister of Infrastructure and Environment introduced the "ladder for sustainable urbanization" in 2012. This process requires planning agencies to go through three steps before planning approval. Municipalities or regional government organizations have to first document there is demand for new development in their area. If, even accounting for current and future supply, demand for development is still identified, then the agencies pushing for development should identify appropriate sites, giving priority to sites in existing urban areas. If, for example because of high re-development costs, development cannot take place in existing urban areas, the planning application moves to the third step. Approval for greenfield development can be provided if no financially viable site within the city is available and a good case can be made. Using the ladder to justify actions is obligatory, even though compliance is not monitored or enforced by the central government. The general guidelines are simple but can lead to very different arguments depending on the context.

One of the goals of this policy framework is to ensure that a proper case is made before development happens outside of brownfield land. It also attempts to ensure an efficient use of urban land (were the term efficient is loosely defined) and promote densification of existing urban areas. For further reference, see Salet (2014).