

Koster, Hans; Zabihidan, Mohammad Saeed

**Working Paper**

## The Welfare Effects of Greenbelt Policy: Evidence from England

Tinbergen Institute Discussion Paper, No. TI 2019-023/VIII

**Provided in Cooperation with:**

Tinbergen Institute, Amsterdam and Rotterdam

*Suggested Citation:* Koster, Hans; Zabihidan, Mohammad Saeed (2019) : The Welfare Effects of Greenbelt Policy: Evidence from England, Tinbergen Institute Discussion Paper, No. TI 2019-023/VIII, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:

<https://hdl.handle.net/10419/205313>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

TI 2019-023/VIII  
Tinbergen Institute Discussion Paper

# The Welfare Effects of Greenbelt Policy: Evidence from England

*Hans R.A. Koster<sup>1</sup>*

*Mohammad Saeed Zabihidan<sup>2</sup>*

<sup>1</sup> Department of Spatial Economics, Vrije Universiteit Amsterdam

<sup>2</sup> Faculty of Economics, Razi University

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: [discussionpapers@tinbergen.nl](mailto:discussionpapers@tinbergen.nl)

More TI discussion papers can be downloaded at <https://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam  
Gustav Mahlerplein 117  
1082 MS Amsterdam  
The Netherlands  
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam  
Burg. Oudlaan 50  
3062 PA Rotterdam  
The Netherlands  
Tel.: +31(0)10 408 8900

# The Welfare Effects of Greenbelt Policy: Evidence from England\*

Hans R. A. Koster<sup>†</sup>      Mohammad Saeed Zabihidan<sup>‡</sup>

August 2018

**Abstract** — We measure the economic effects of *urban growth boundaries* or *greenbelts* that prohibit new construction beyond a predefined urban fringe. We focus on England, where 13% of the land area is designated as greenbelt land. Using spatial differencing, we show that the external effects of these regulations are substantial (about 15-20%) but very local (within 1km). In contrast to the previous literature, we find no evidence for internal or ‘own-lot’ effects. We further show that supply effects are important: greenbelt policy reduces housing construction in greenbelts by about 80%, thereby increasing prices throughout the housing market by about 4%. We show that greenbelt policy implies a negative welfare cost of about £7.5 billion a year (0.5% of England’s GDP). We further find evidence that greenbelts are no popular recreational destinations (proxied by geocoded pictures), and do not imply longer commutes or more housing CO<sub>2</sub> emissions.

**Keywords** — housing, supply constraints, greenbelts, urban growth boundary, open space

**JEL codes** — G10, R30

---

\*This work has benefited from financial support from the Netherlands Organisation for Scientific Research (VENI grant). Paul Cheshire is thanked for particularly helpful discussions. We also thank Sofia Franco, Steve Gibbons, Christian Hilber and Tobias Seidel, as well as seminar participants in Düsseldorf, London and Vancouver, for useful input.

<sup>†</sup>Corresponding author. Department of Spatial Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105 1081 HV Amsterdam, email: [h.koster@vu.nl](mailto:h.koster@vu.nl). Hans is also research fellow with the Higher School of Economics (Russian Federation), the Tinbergen Institute and the Centre for Economic Policy Research.

<sup>‡</sup>Faculty of Economics, Razi University, Tagh-e-Bostan, 6714967346, Kermanshah-Iran, email: [zabihidansaeed@yahoo.com](mailto:zabihidansaeed@yahoo.com).

# 1 Introduction

In many countries urban growth leads to an increasing pressure on developable land in and around cities. Most cities aim to regulate urban development by the imposition of several constraints on *e.g.* building height or type of land use. Local governments may also restrict the expansion of urban areas in order to prevent urban sprawl. These *urban growth boundaries* or *greenbelts* reduce the land available for development at the urban fringe. Greenbelts are important and cover about 13% of the total area in England, and are surrounding most of the larger cities. Also many U.S. cities, such as Portland (OR), Miami, Minneapolis Saint-Paul, and San Jose (CA), have urban growth boundaries.

Land use regulation does not necessarily lead to welfare losses, because constraints may reduce negative land use externalities and frictions associated with development. Greenbelt policy intends to protect agricultural land and secure amenity benefits from open space (Brueckner 2001). However, when regulatory constraints do not adjust to changes in demand, regulation may also imply substantial economic losses. It indeed has been argued that greenbelt policy should be relaxed to mitigate the ‘housing affordability crisis’ because restrictions on housing supply supposedly lead to strong price appreciation (Cheshire 2014, Economist 2017). However, despite the importance of greenbelts restricting the growth and development of cities and the potentially pronounced impacts on housing markets, to the best of our knowledge no study has attempted to evaluate the different welfare effects of greenbelt policy.

Our study focuses on the effects of greenbelt policy on the housing market and aims to identify both *local* and *aggregate* effects. Following Turner et al. (2014), we make a distinction between three *local* effects of greenbelt policy: an internal, an external and a supply effect.

- The internal effect reflects the cost of greenbelts to restrictions on the consumption of housing. In greenbelts it may be more costly for home owners to alter their properties as it may be more costly to get planning permission. This may lead to a lower house price. Internal effects may also be reflected in housing consumption; houses may be either larger or smaller in greenbelts.
- The external effect is the (dis)amenity benefit on one’s neighbours. If greenbelts are

predominantly used for intensive livestock farming, this external effect may be negative (Bontemps et al. 2008). On the other hand, greenbelts ensure that houses are closer to open space, which may generate positive benefits (Irwin 2002, Anderson & West 2006, Brander & Koetse 2011).

- The supply effect of greenbelt policy constitutes the reduction in the amount of developable land. This implies that plots that are in the greenbelt cannot be developed, leading to (much) lower land revenues for land owners.

Greenbelt policy may also have *aggregate* effects. When regulation restricts the number of dwellings that are constructed, this leads to higher prices for residential real estate throughout the housing market (Hilber & Vermeulen 2016). As a result, households may occupy smaller homes. We then put the local and aggregate results together by adopting a simple welfare framework based on the changes in prices of residential properties in England.

We use data on more than 10 million housing transactions in England between 1995 and 2017. We link these data to information on housing characteristics – and importantly – the exact location so that we can identify whether a property is in or close to one of the greenbelts. Our empirical set-up relies on two identification strategies to measure the local and aggregate effects, respectively. Local effects are identified by focusing on observations close to the inner and outer greenbelt boundaries. We include detailed fixed effects to control for unobserved locational characteristics. This ensures that – locally – the greenbelt boundary can be considered as random. Moreover, we estimate specifications excluding observations close to greenbelt boundaries that intersect with roads, rivers and administrative boundaries (*e.g.* school districts) to make sure that the greenbelt boundary is not capturing something else. The results imply that prices can be up to 20% higher in greenbelts, which confirms that there is indeed a positive external effect of the greenbelt policy because of proximity to open space. In contrast to Turner et al. (2014) we do not find robust evidence for an internal effect, although it seems that houses closer to greenbelts are somewhat larger. We further show that supply effects are paramount: greenbelt policy reduces the availability of dwellings with almost 80%.

The second set of regressions focuses on the aggregate price and housing consumption effects of

this reduction in housing supply by calculating the price and level of housing consumption net of the internal and external effect, and compare price and house size differences between housing markets with more and less greenbelt land. To mitigate the issue of unobserved characteristics that may be correlated to the amount of greenbelt land, we construct so-called counterfactual greenbelts. We exploit the fact that greenbelt boundaries in England have hardly changed since their imposition in 1955. More specifically, we use information on the population in small areas before implementation of greenbelt policy in 1955 and construct counterfactual greenbelts using information on population density and city size in 1951. We then compare prices in areas with actual greenbelts with areas where one would expect a greenbelt based on the historic population distribution. We find that a 50% increase in the share of greenbelts within 15km of the property increases house prices by about 7.5%. We also find evidence that greenbelt policy implies that households live in smaller houses. This effect is of about the same magnitude, so the elasticity of price with respect to house size is about 1; in other words, households seem to substitute away from house size when housing becomes more expensive.

We then put these results together to approximate the welfare effects. Our welfare measure takes into account the higher price and resulting shift in demand when housing supply is restricted. We show that greenbelt policy implies a negative welfare effect, because of a supply effect that strongly reduces the number of available housing units in greenbelt land. The supply effect is about three times as large as the external effect. Overall, greenbelt policy implies a negative welfare cost of about £7.5 billion a year – 0.5% of England’s GDP.<sup>1</sup>

One may argue that the measurement of the welfare effects does not include other general equilibrium effects, *e.g.* via visiting people that may live further away from the greenbelt, an inefficient urban structure that leads to longer commutes, or to more housing energy consumption because of suboptimal densities. Using ancillary data on geocoded pictures, arguably capturing locational attractiveness, we find that greenbelts do attract fewer tourists and natives than nearby areas. Using data on commuting and housing CO<sub>2</sub> emissions, we do not find that greenbelt policy implies longer commutes or more CO<sub>2</sub> emissions.

---

<sup>1</sup>It is important to not interpret this as the total welfare effects of land use restrictions, because regulation in England is ubiquitous and regulation is also pronounced in areas outside greenbelts (Hilber & Vermeulen 2016). Hence, in case one would interpret our estimates as estimates of regulation more generally, they serve as a (strong) underestimate.

**Related literature.** Most studies on the effects of land use regulation so far concentrate on housing supply restrictions and show that supply constraints are associated with increasing housing costs, a strong reduction in new construction and rapid price growth (Mayer & Somerville 2000, Glaeser et al. 2005, Green et al. 2005, Ihlanfeldt 2007). This effects is particularly pronounced for cities in England, in which land use regulation is very restrictive (Hilber & Vermeulen 2016). Recent evidence for England by Cheshire et al. (2018) show that land use regulation may also lead to higher vacancy rates and longer commutes. Effects of land use regulation can also be local. Glaeser & Ward (2009) find that local constraints do not increase the price of that good because of close substitutes (so within a small area). On the other hand, they find that density levels in Massachusetts are too low from a welfare point of view. Koster et al. (2012) finds that internal effects of regulation may be substantial (up to 10%) of the housing value. They find limited evidence for sizeable external effects. Turner et al. (2014), which is the paper that is the closest to our work, evaluate the internal and external effects of land use regulation on the value of land and on welfare in the U.S., but they pay limited attention to welfare implications of the reduction in housing supply and do not take into account aggregate effects of regulation on housing consumption.

The plan for the remainder of the paper is as follows. In Section 2 we discuss the theoretical framework. In Section 3 we explain how greenbelts were designated, introduce the datasets and provide descriptives. Section 4 outlines the empirical methodology. In Section 5 we report the results with respect to local effects and aggregate effects of greenbelt policy. We put these results together by means of a counterfactual analysis, followed by a summary of the sensitivity analysis. Section 6 concludes.

## 2 Theoretical considerations

### 2.1 Local effects: internal, external and supply effects

We improve on the approach by Turner et al. (2014) to identify external, internal and supply effects of land use regulation by estimating supply effects and allowing for aggregate changes in prices and housing consumption. For a fuller discussion on the basic set-up we refer to their paper.



We assume that there are two areas, one containing urban land with regulatory restrictiveness  $z^U$  stretching from  $-H$  to  $0$ , and an area being the greenbelt stretching from  $0$  to  $B$  with  $z^G$  denoting the stringency of regulation in greenbelts. We assume dwellings at each location  $i$  are uniformly distributed between  $-H$  to  $B$ . Agents receive a wage  $w$  and pay housing price  $p_i$  per  $m^2$  of house size  $h_i$  to occupy a property at location  $i$ . They also consume a composite good  $c$  for a price that is normalised to one. We assume that all residents are indifferent between locations and receive  $\theta$ . Let us assume that the outside option is heterogeneous, which is distributed as  $\theta \in [0, \infty)$ .  $g(\theta)$  is the *p.d.f.* of households of type  $\theta$  and the *c.d.f.* is  $G(\theta^*) = H + B(z^G)$ . We can fill the area with households for whose outside option is worse than  $\theta^*$ , where  $\theta^*$  is the marginal household that decides to live in the area. Following [Turner et al. \(2014\)](#), let us for convenience assume that the amount of developable land is removed from the interior of the greenbelt. In absence of regulation, the amount of land is given by  $H + B$ , while  $\partial B / \partial z^G < 0$  implying that the amount of developable land, and therefore the number of dwellings, in the greenbelt is decreasing in regulatory restrictiveness.

We further assume that construction companies can freely adjust housing outside greenbelts. On the contrary, we assume that inside greenbelts housing consumption is more expensive with  $\zeta(z_i) \geq 1$  because it is may be more costly to make adjustments to properties.

We assume that utility is then given by a standard Cobb-Douglas function, so that:

$$\theta^* = h_i^\alpha c^{1-\alpha} e^{\beta v_i} \quad \text{s.t.} \quad w = p_i \zeta(z_i) h_i + c, \quad (1)$$

where  $v_i$  is the locational quality of  $i$  related to greenbelt policy,  $\beta > 0$  is a preference parameter for greenbelts, which we expect to be positive when households value greenbelt land, and  $\alpha$  is a preference parameter. The optimality conditions associated with (1) imply that the bid rent at a certain location is given by  $\log p_i = -\log \zeta(z_i) + \frac{\beta}{\alpha} v_i + \frac{1}{\alpha} \log \left( \frac{w}{\theta^*} \right) + \tilde{c}$ , where  $\tilde{c} = \log \alpha + \frac{1-\alpha}{\alpha} \log(1-\alpha)$ . Housing consumption at  $i$  is then given by:  $\log h_i = -\frac{\beta}{\alpha} v_i + \frac{1}{\alpha} \log \theta^* + \frac{\alpha-1}{\alpha} \log((1-\alpha)w)$ .

We can make a distinction between an internal price effect of regulation and an external price effect  $v_i(z_k)$ . The internal effect arises because for house owners and developers it is more

expensive to adjust the house leading to higher housing costs. For the internal effect we would expect a negative discrete jump in prices at the greenbelt boundary when  $\zeta(z_i) > 1$ .

The net external benefits of greenbelts derived at a certain location can be given as:

$$v_i(z_k) = \delta \int_{-H}^B f(\delta, d_{ik}) g(z_k^G) dk \quad (2)$$

where  $d_{ik}$  is the distance between  $i$  and  $k$ ,  $g(\cdot)$  indicates whether a location  $i$  is in the greenbelt,  $\delta$  is a decay parameter and  $f(\cdot)$  is a distance decay function, where it holds that  $\partial f(\cdot)/\partial d_{ik} < 0$  and  $H + B$  is the total land area and  $g_k$ . Hence,  $v_i$  (and therefore  $p_i$ , see equation (??)) is higher when there is more greenbelt land in the vicinity of  $i$ .

We illustrate the different effects in Figure 1. Suppose that greenbelt land offers appreciable amenities, in line with the literature that shows that households value open space (Irwin 2002, Rouwendal & Van der Straaten 2008, Vermeulen & Rouwendal 2014). In absence of an internal price effect, the dotted black line in Figure 1 shows that prices will always be higher in the greenbelt, in particular when  $i$  is further away from the greenbelt boundary. If housing consumption is indeed more expensive in the greenbelt (so when  $\zeta(z_i) > 1$ ), the prices may be lower in the greenbelt if the internal effect is large enough. Once greenbelt policy also have an impact on the supply of housing so that the amount of developable land is reduced from  $H + B$  to  $H + B(z_2^G)$  the price gradient shifts up, because fewer houses are available and only households with a low enough  $\theta^*$  will locate in this area.

## 2.2 Aggregate effects

A reduction in housing supply is expected to raise prices in the whole housing market, soup to  $-H$ . Moreover, it can be expected that when houses become more expensive due to limited supply, households will reduce housing consumption and occupy smaller houses ( $\bar{h}_2 < \bar{h}_1$ ) throughout the housing market. Let us denote  $\bar{p} = p_{-H}$  and  $\bar{h} = h_{-H}$  respectively as the housing price and housing consumption net of external and internal effects. We then will use a separate identification strategy to estimate aggregate changes in house price and housing consumption due to greenbelt policy.

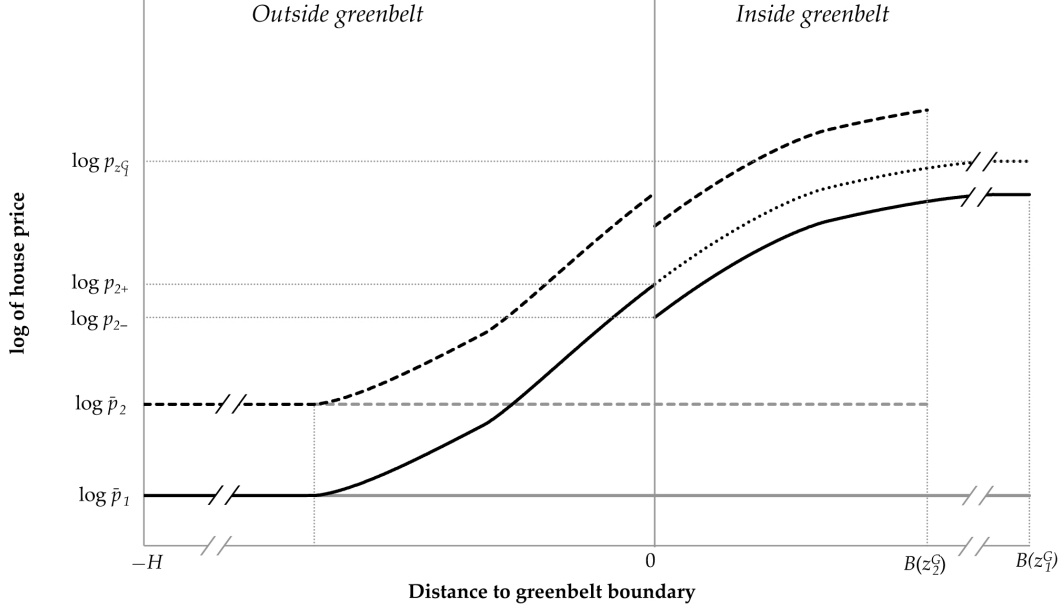


FIGURE 1 – PRICE GRADIENTS AROUND THE GREENBELT BOUNDARY

### 2.3 Welfare effects

Before we introduce the econometric framework to estimate the relative sizes of the different effects on house prices and housing consumption we aim to investigate the changes in welfare through changes in land use regulation. We examine a change from  $z^U$  to  $z^G$ , with  $z_1^G = z^U$  and  $z_2^G > z^U$ . Recall that  $h_i$  denotes housing consumption. The aggregate rents from housing consumption is then used as a proxy for the producer's surplus. Following [Turner et al. \(2014\)](#), the producer and consumer surplus are respectively given by:

$$\int_{-H}^B p_{it} h_{it} di \quad \text{and} \quad \int_0^{\theta_t^*} \check{p}_{it}(\theta) \check{h}_{it}(\theta) d\theta - \int_{-H}^B p_{it} h_{it} di, \quad (3)$$

where  $t = 1, 2$  and  $\check{p}_{it}(\theta) \check{h}_{it}(\theta)$  indicates (unobserved) bid rents of households for dwellings. The total change in welfare given by:

$$\begin{aligned} \mathcal{W}_2 - \mathcal{W}_1 = & \underbrace{\int_{-H}^{B(z_2)} p_{i2} h_{i2} - \bar{p}_2 \bar{h}_2 di}_{\text{internal and external effect}} + \\ & \underbrace{\int_0^{\theta_2^*} \check{p}_{i2}(\theta) \check{h}_{i2}(\theta) - p_{i2} h_{i2} d\theta - \int_0^{\theta_1^*} \check{p}_i(\theta) \check{h}_{i1}(\theta) - p_{i1} h_{i1} d\theta}_{\text{supply effect}} \end{aligned} \quad (4)$$

The first term captures the change in land rents due to the combined internal and external effect. The second and third term capture the total difference in bid rents between a situation with and without greenbelt policy. Note further that the loss in consumer surplus due to higher prices exactly cancels out by the gain for home owners. For the marginal household it holds that  $\check{p}_{i2} = p_{i2}$ , so that we can simplify (4) to:

$$\begin{aligned}
\mathcal{W}_2 - \mathcal{W}_1 &= \underbrace{\int_{-H}^{B(z_2)} p_{i2} h_{i2} - \bar{p}_2 \bar{h}_2 di}_{\text{internal and external effect}} + \\
&\quad \underbrace{\int_{\theta_2^*}^{\theta_1^*} \bar{p}_1(\theta) \bar{h}_1(\theta) d\theta + \int_{\theta_2^*}^{\theta_1^*} p_{i2}(\theta) h_{i2}(\theta) - \bar{p}_1(\theta) \bar{h}_1(\theta) d\theta}_{\text{supply effect}} \\
&\approx \underbrace{\int_{-H}^{B(z_2)} p_{i2} h_{i2} - \bar{p}_2 \bar{h}_2 di}_{\text{internal and external effect}} - \underbrace{\Delta(\bar{p}_1 \bar{h}_1) - \frac{\Delta}{2}(\bar{p}_2 \bar{h}_2 - \bar{p}_1 \bar{h}_1)}_{\text{supply effect}}
\end{aligned} \tag{5}$$

The first term in (5) is again the change in land rents due to the internal and external effect. The second term captures the main supply effect, which is the price of dwellings ( $p_i h_i$ ) integrated over the number of foregone dwellings. The third term captures the aggregated change in prices due to greenbelt policy integrated over the number of foregone dwellings.

We make the alternative assumption that demand is approximately uniformly distributed between  $\theta_2^*$  and  $\theta_1^*$ . Let  $\Delta$  be the (absolute) difference in the number of dwellings before and after the greenbelt policy. In the above equation we show that (5) can be further simplified.

It is apparent that for the estimation of external and internal effects one does not need information on prices before greenbelt policy was implemented. However, to estimate the second part of (5), we need information on prices and housing consumption before and after the imposition of greenbelts ( $\bar{p}_1 \bar{h}_1$  and  $\bar{p}_2 \bar{h}_2$ ). If one assumes that prices and housing consumption are more or less similar ( $p_{-H1} h_{-H1} \approx p_{-H2} h_{-H2}$ ), this leads to:

$$\mathcal{W}_2 - \mathcal{W}_1 \approx \underbrace{\int_{-H}^{B(z_2)} p_{i2} h_{i2} - \bar{p}_2 \bar{h}_2 di}_{\text{internal and external effect}} + \underbrace{-\Delta(\bar{p}_2 \bar{h}_2)}_{\text{supply effect}} \tag{6}$$

which is equivalent to the welfare effect by [Turner et al. \(2014\)](#).<sup>2</sup> However, when demand for

---

<sup>2</sup>Please note that the above equation is only dependent on house prices after greenbelt policy has been

housing changes drastically due to a reduction in housing supply, equation (5) may lead to a substantially different estimate of the supply effects of greenbelts than equation (6).

### 3 Context, data and descriptives

#### 3.1 Institutional setting

There is a long-standing tradition in England to restrict urban growth. In the 1920s, proposals were put forward by the London Society and the Campaign to Protect Rural England (CPRE) to prevent development in a continuous belt within 2km of London. In the 1947 Town and Country Planning Act, local authorities (LAs) were for the first time allowed to take planning decisions and to incorporate greenbelt proposals in their development plans.

In 1955, Duncan Sandy, who was then the Minister of Housing, encouraged local authorities around the country to consider protecting land around cities by the formal designation of well-defined greenbelts. In a statement in the House of Commons he wrote:

*“I am convinced that for the well-being of our people and for the preservation of the countryside, we have a clear duty to do all we can to prevent the further unrestricted sprawl of the great cities. The Development Plans submitted by the local planning authorities for the Home Counties provide for a Green Belt, some 7 to 10 miles deep, all around the built-up area of Greater London. [...]. No further urban expansion is to be allowed within this belt.”*

and:

*“In other parts of the country, certain planning authorities are endeavouring, by administrative action, to restrict further building development around the large urban areas. But I regret that nowhere has any formal Green Belt as yet been proposed. I am accordingly asking all planning authorities [...] to submit to me proposals for the creation of clearly defined Green Belts, wherever this is appropriate.”*

Greenbelts were introduced soon after around almost all the big cities (London, Birmingham, Liverpool and Manchester), but also around smaller cities (*e.g.* Bournemouth, York, Oxford  

---

implemented ( $\bar{p}_2 \bar{h}_2$ ).

and Cambridge). Almost all cities that put forward proposals for greenbelt land had at least a population of 100,000 inhabitants at that time and so qualified as “large urban areas”. Since then no new greenbelts have been introduced and the total amount of greenbelt land essentially has not changed in the last 25 years.<sup>3</sup> Currently, greenbelts cover about 13% of all land in England (for comparison, built-up land covers about 10%) and should, according to the National Planning Policy Framework in 2012, offer appreciable amenities to the urban population by improving access to the open countryside, by providing opportunities for outdoor sport and recreation, and by retaining attractive landscapes close to urban areas. However, critics claim that greenbelt land is often used for intensively farmed agricultural land, rather than used for recreational and nature purposes, and is therefore unlikely to offer amenity value (Cheshire 2014). Indeed, about one third of greenbelt land is used for agriculture, while only about 7% of the land is classified as accessible open spaces, parks or gardens.<sup>4</sup>

In Figure 2 we show the 14 greenbelts in England. In particular the Metropolitan Greenbelt around London is large, but also greenbelts around Birmingham (West-Midlands) and Manchester/Liverpool (North-West) are substantial. In the map we also display national parks and Areas of Outstanding Natural Beauty (AONBs). It is likely that in those areas new development is also restricted. However, because those areas are usually in rural low-density areas, the effects on the housing market are expected to be an order of magnitude smaller.

### 3.2 Data

We make use of three datasets. The first dataset contains the universe of housing transactions from England from the *Land Registry* between 1995 and 2017. These data provide information on the transaction price as well as the housing type, the date of the transaction and the ownership structure (leasehold or freehold). A disadvantage of the Land Registry data is the limited amount of information on housing characteristics, most importantly size of the house and number of rooms.

We therefore merge the *Land Registry* data to additional characteristics using data on *Energy*

---

<sup>3</sup>For example, the total hectares of greenbelt land in 1997 was 1,652,310, while it was 1,638,610 in 2013, a change of less than 1%, which may as well be due to measurement error.

<sup>4</sup>In our empirical analysis we will therefore include specifications where we make a distinction between accessible vs. non-accessible and agricultural vs. non-agricultural greenbelt land.



FIGURE 2 – GREENBELTS IN ENGLAND

*Performance Certificates* (EPCs). Since 2007 an EPC has been required whenever a home is constructed or sold. The dataset contains all EPCs issued since October 1, 2008. The data provide information of the energy performance of buildings and their characteristics that are obtained by a physical inspection of the interior and exterior of the home by an independent assessor. This provides us with the floor area of the property, number of rooms, as well as the energy performance.

Our merging strategy is to sequentially match individual sales to the EPC data using the full address or a subset of the address and the date of the sale and certificate.<sup>5</sup> About one third of the sales in the *Land Registry* remains unmatched, so we drop them from the analysis. We also drop transactions that are matched to multiple EPCs (about 15%).<sup>6</sup>

Information on greenbelts in 2012 is obtained from the *Department of Communities and Local Governments* (DCLG). Each local authority digitised land use information and DCLG merged these separate datasets. We aim to identify internal, external and supply effects using the inner and outer boundaries of greenbelts, as these capture the actual urban containment boundaries. Hence, we determine the inner and outer boundaries and calculate the distance of each property to the nearest inner or outer boundary of a greenbelt.

We further gather data from *Ordnance Survey* (OS) on roads and rivers to investigate whether the greenbelt boundaries intersect with the course of rivers and roads. From OS we also obtain information on parks. From the *Land Cover* dataset we get information on agricultural land and developed land in each postcode.

We drop observations of prices that are above £1.5 million or below £15,000 (less than 0.5% of the data). Because greenbelt boundaries have hardly changed between 1995 and 2017, we use the full temporal extent of the data.<sup>7</sup> This leaves us with 10,210,717 sales. For the final data on *EPCs* we have 13,543,673 observations.

---

<sup>5</sup>Specifically, we first match a sale to an EPC using the primary address object name, secondary address object name, street name, and postcode. We then keep the certificate that is closest in days to the sale. We repeat this exercise for unmatched properties but allowing one of the address identifiers to be different. Our final round of matching matches on the full postcode.

<sup>6</sup>The matching is harder for flats that often share an address, implying that the proportion of flats is reduced from 23% in the *Land Registry* sample to 3% in the final sample. Our analysis therefore mainly focuses on single-family homes.

<sup>7</sup>However, we will make sure that when focusing on observations in 2012 only this does not change the results (see Section 5.4 for more details).



For the analysis of the aggregate effects we aim to calculate how the amount of greenbelt land affects aggregate price and housing consumption levels. We cannot use standard definitions of housing markets used in the literature, such as local authorities or travel-to-work-areas, because those areas are often delineated by greenbelts. We therefore create a dataset at the Output Area (OA) level – the lowest geographical level at which census estimates are provided (the median size of an OA is only 6.6ha). For each OA we calculate the share of greenbelt land within 15km, which corresponds to the average commuting distance in England of 15.18km (based on the 2011 Census). We further calculate the current and 1951 population density within 15km, as well as the share of counterfactual greenbelt land within 15km. To control for geographical constraints, we calculate the share of water bodies, the share of developed land, and the standard deviation of altitude, all within 15km (see *e.g.* Saiz 2010). Using information from OS, we also calculate the share of AONBs and national parks, in which there are potential restrictions on new home construction as well. Finally, we obtain data on OA characteristics from the 2011 Census, such as average number of rooms, the share of owner-occupied housing and the share of properties with central heating.

### 3.3 Descriptive statistics

Table 1 reports descriptive statistics for the housing transactions data. On average 3.6% of the transactions are in the greenbelt. We show that the average price per m<sup>2</sup> of floor space is £1769 and the average floor size is 87m<sup>2</sup>. In greenbelts this is respectively £2410 and 97m<sup>2</sup>. Not surprisingly, the share of developed land in postcodes is high (0.83); however, this share is only 0.34 for dwellings in greenbelts. In contrast, the share of agricultural land for properties in greenbelt is much higher (0.085 versus 0.289).

We also report descriptive statistics at the output area level in Table 2. On average, 27% of the land is greenbelt land. This is somewhat higher than the total share of land used for greenbelts (13%), which shows that greenbelts are close to urban areas. The share of AONBs and national parks within 15km of OAs is much lower.

TABLE 1 – KEY DESCRIPTIVE STATISTICS FOR HOUSING TRANSACTIONS

	(1)	(2)	(3)	(4)
	mean	sd	min	max
Price per m <sup>2</sup>	1,768.9130	1,277.2650	100.0000	10,000.0000
Distance to greenbelt boundary ( <i>in km</i> )	16.4492	29.2659	0.0000	296.8268
Share of postcode in greenbelt	0.0356	0.1624	0.0000	1.0000
Share of greenbelt land in vicinity, $\delta = 3.767$	0.0966	0.1865	0.0000	1.0000
Housing type – flat	0.0317	0.1753	0.0000	1.0000
Housing type – terraced	0.4118	0.4922	0.0000	1.0000
House newly built	0.0690	0.2535	0.0000	1.0000
Share developed land in postcode	0.8291	0.3093	0.0000	1.0000
Share parks in postcode	0.0331	0.1143	0.0000	1.0000
Share agricultural land in postcode	0.0849	0.2011	0.0000	1.0000
Distance to the city centre ( <i>in km</i> )	35.7809	33.5061	0.0802	313.6746

*Notes:* The number of observations is 10,210,717. The full list of descriptive statistics are provided in Tables A1 and A2 in Appendix A.

TABLE 2 – DESCRIPTIVE STATISTICS FOR OUTPUT AREAS

	(1)	(2)	(3)	(4)
	mean	sd	min	max
Share greenbelt land, < 15km	0.2695	0.2432	0.0000	0.8987
Population density per km <sup>2</sup> , < 15km	14.0211	16.4982	0.0175	78.5325
Standard deviation of altitude, < 15km	43.5435	25.5685	0.8131	211.0376
Share AONB < 15km	0.0595	0.1351	0.0000	0.9964
Share national park < 15km	0.0147	0.0689	0.0000	0.9990
Share developed land < 15km	0.2407	0.2014	0.0000	0.8296
Share water land < 15km	0.0157	0.0172	0.0000	0.1505
Household size	2.3709	0.3734	1.0526	6.1091
Share single households	0.2946	0.1176	0.0000	0.9645
Share lone parents	0.1159	0.0700	0.0000	0.5938
Share non-adults (< 18 years)	20.8611	6.6508	0.0000	79.0000
Share elderly people	17.4092	10.0176	0.0000	96.7000
Share highly-educated people	0.3880	0.1465	0.0349	0.9586
Share Buddhists	0.0041	0.0074	0.0000	0.4960
Share Hindus	0.0132	0.0388	0.0000	0.9218
Share Muslims	0.0381	0.0934	0.0000	0.9803

*Notes:* The number of Output Areas is 161,336. The number of observations in the regressions in Section 5.2 is the number of OA times the number of years, while ignoring the areas for which we do not observe prices or house size.

## 4 Methodology

### 4.1 Identifying local effects: internal, external and supply effects

We make a distinction between internal, external and supply effects. First, we are interested in the local internal and external effects of greenbelt policy using information on house prices. Following [Turner et al. \(2014\)](#) we re-emphasise that the presence of internal effects may lead to a discrete price change at the greenbelt boundary: within the greenbelt more stringent regulations apply that prevent house owners from making substantial changes to their properties. Let then  $p_{it}$  be the house price in postcode  $i$  in output area  $j$  in year  $t$  and  $g_i$  be the share of the own postcode that is in a greenbelt,

We capture external effect  $v_i$  using a distance decay function. Let  $\delta$  denote a decay parameter,  $d_{ik}$  the distance between  $i$  and  $k$ . Locations  $k = 1, \dots, K$  form a grid of one hectare per cell  $k$ , and  $\tilde{g}_k$  is a dummy that equals one when the centroid of cell  $k$  is in the greenbelt. The empirical analogue of equation (2) is then:

$$v_i = \frac{\sum_{k=1}^K e^{-\delta d_{ik}} \tilde{g}_k}{200\pi/\delta}. \quad (7)$$

Hence,  $v_i$  equals one when all the land in vicinity is in the greenbelt.<sup>8</sup> Note that a higher  $\delta$  implies that we put more weight on land in the vicinity. In [Figure A1](#) in [Appendix A](#) we show the relative weight of an hectare of greenbelt land for different  $\delta$ . In the empirical application, we choose  $\delta = 3.767$  implying that the effect of greenbelts is mainly important within 1km. To justify this choice, we discuss supporting evidence in the sensitivity analysis, [Section 5.4](#).

One may argue that the amount of greenbelt land in the vicinity is correlated to housing attributes; houses with particular price-increasing characteristics may be predominantly located in greenbelts. For example, because of historic city limits, properties in greenbelts may be disproportionally detached, while houses outside greenbelts may come in the form of apartments or terraced housing. To mitigate this problem we include housing characteristics, denoted by  $c_i$ .

To control for (changes in) unobservable locational attributes, we include OA  $\times$  year fixed effects  $\rho_{jt}$ . Because OAs are small, the fixed effects are likely capturing many unobserved characteristics

---

<sup>8</sup>It can be shown that when  $i$  is fully surrounded by greenbelt land  $\sum_{k=1}^K e^{-\delta d_{ik}} \tilde{g}_k \rightarrow 200\pi/\delta$ .

that could be correlated to the share of greenbelt land, such as accessibility and neighbourhood composition. It may, nevertheless, still be preferable to control for additional neighbourhood characteristics  $m_i$ , including distance to the nearest city centre of a city larger than 250 thousand inhabitants, distance to roads and rivers and land use in the postcode. Hence:

$$\log p_{ijt} = \alpha_1 g_i + \alpha_2 v_i + \alpha_3 m_i + \alpha_4 c_i + \rho_{jt} + \epsilon_{ijt}. \quad (8)$$

where  $\epsilon_{it}$  is an independently and identically distributed error term. Note that  $\alpha_1$  captures the internal price effect at the greenbelt boundary, while  $\alpha_2$  captures the importance of the external effect.

A concern with the above specification is that greenbelt boundaries are not random and for example follow the course of rivers. If we find a positive external effect of greenbelts, this may be caused by the fact that areas close to the greenbelt boundary are close to natural amenities that are not necessarily a result of greenbelt policy. First, we therefore exclude observations that are within 2.5km of a greenbelt boundary that intersects with the course of rivers, major roads and, importantly, administrative boundaries. By excluding observations near greenbelt boundaries that intersect with Local Authority boundaries, we control for differences in *e.g.* local taxation, school quality, and the provision of other public goods, which may be correlated to being in or outside the greenbelt. Second, we will employ a boundary design where we include fixed effects for every straight greenbelt boundary segment.<sup>9</sup> Further, we only include properties that are within  $\bar{d}$ km to greenbelt boundaries, so that –locally– the boundary can be considered random. The specification to be estimated is then:

$$\log p_{ijt} = \alpha_1 g_i + \alpha_2 v_i + \alpha_3 m_i + \alpha_4 c_{it} + \rho_{jtb} + \epsilon_{ijt} \quad \text{if } d_{ib} < \bar{d}, \quad (9)$$

where  $\rho_{jtb}$  are OA  $\times$  year  $\times$  boundary segment fixed effects,  $d_{ib}$  is the distance to the nearest boundary segment and  $\bar{d}$  is some threshold distance.

We also test whether local housing consumption is impacted by  $v_i$ , as suggested by equation

---

<sup>9</sup>The average length of a boundary segment is only 8.7m, so we control in a very detailed way for heterogeneity in the greenbelt boundary.

(??). We basically repeat equation (9), but replace  $p_{it}$  by  $h_{it}$ , so that we have:

$$\log h_{ijt} = \beta_1 g_i + \beta_2 v_i + \beta_3 m_i + \beta_4 c_{it} + \sigma_{jtb} + \epsilon_{ijt} \quad \text{if } d_{ib} < \bar{d}, \quad (10)$$

where  $\sigma_{jtb}$  are OA×year×boundary segment fixed effects, and  $\beta_1$  is expected to be zero. We emphasise that we might find that house size is restricted and related to greenbelt policy, but that this not necessarily implies that there are negative price effects, because heterogeneous households may sort themselves in properties of different sizes.

Greenbelt policy also restricts the number of dwellings on greenbelt land. We then employ a similar boundary design as above and estimate the following regression:

$$n_{ij} = \exp(\gamma_1 g_i + \gamma_2 m_i + \lambda_{jb} + \epsilon_{ij}) \quad \text{if } d_{ib} < \bar{d}, \quad (11)$$

where  $n_i$  is the number of dwellings in a postcode in 2011,  $g_i$  is the share of the postcode in the greenbelt,  $m_i$  are neighbourhood characteristics and  $\lambda_{jb}$  are OA×boundary segment fixed effects.<sup>10</sup> Because  $n_i$  is a count variable, we estimate equation (11) by a Poisson model, so that the  $\gamma$ 's are interpreted as semi-elasticities.

#### 4.2 Identifying aggregate effects

To measure aggregate effects on prices and housing consumption we use the estimated  $\hat{\rho}_{jt}$  and  $\hat{\sigma}_{jt}$  from equations (9) and (10) respectively, so that  $\log \bar{p}_{jt} \equiv \hat{\rho}_{jt}$  and  $\log \bar{h}_{jt} \equiv \hat{\sigma}_{jt}$ . Hence,  $\hat{\rho}_{jt}$  and  $\hat{\sigma}_{jt}$  represent prices and housing consumption net of the internal and external effects. We then calculate the share of greenbelt land within 15km, denoted by  $G_j$ .<sup>11</sup> Of course, prices and housing consumption are not only influenced by greenbelt policies, but also by the demand for a certain location and other policies that may restrict new housing construction. We therefore control for location attributes  $M_j$ , including the share of AONBs, national parks, developed land and water bodies, as well as the standard deviation of the altitude within 15km.  $M_j$  also includes the population density within 15km and output area demographics. We further will

---

<sup>10</sup>Note that, next to the fixed effects, we also control for distance to the nearest city centre, which should absorb any density gradient.

<sup>11</sup>The choice of 15km should capture meaningful housing markets and is line with the average commuting distance. However, we provide robustness to this assumption by changing this to 10 and 50km, respectively, in Appendix B.

include 9 (government) region×year fixed effects  $\mu_{jt}$  to control for unobserved trends in the regional housing market.

When identifying aggregate effects of greenbelt policy, including a sheer endless string of control variables is likely to mitigate, but not solve, the problem of omitted variable bias. We therefore aim to construct counterfactual greenbelts so that we can compare price and housing consumption levels in areas with actual and counterfactual greenbelts. Areas with counterfactual greenbelt land should then be similar in unobservables to areas with greenbelt land. To construct counterfactual greenbelts we gather data on the population in Parishes in 1951, which was just after the 1947 Town and Country Planning Act was implemented, but before greenbelt policy was implemented.<sup>12</sup> Greenbelts were predominantly implemented around larger cities of at least 100 thousand inhabitants. To identify urban areas we first select Parishes with a population density of at least 10 people per ha, which applies to about 5% of the areas. The next step is to amalgamate areas that are adjacent to each other and keep 37 amalgamated areas that have a population of more than 100,000. The next step is that we draw circles of 15km around each of these urban areas but erase the parishes in the counterfactual greenbelts with a density of at least 10 persons per ha.

In Figure A2 in Appendix A we display a map with the counterfactual greenbelts. The correlation between the share of greenbelt land within 15km and the share of counterfactual greenbelt land within 15km is reasonably high and around 0.5. In general, we are able to predict the location of most greenbelts quite well, such as the greenbelts around London, Birmingham, Manchester and Liverpool. Only the Cheltenham/Gloucester greenbelt, as well as the greenbelt around Cambridge are not included in the counterfactual greenbelt sample because those cities had a population lower than 100,000 in 1951. On the other hand, reasonably large cities such as Leicester, Norwich, Middlesbrough and Plymouth do not have greenbelts, although one would expect a greenbelt around those areas based on the 1951 population distribution.

We then estimate specifications where we control flexibly for the share of counterfactual greenbelt land  $C_j$ . As long as  $f(C_j)$  is flexible enough, this implies that we compare prices and housing consumption between areas with an identical  $C_j$ , but with a different share of actual greenbelt

---

<sup>12</sup>Parishes are quite small. The median size is 847 hectares.

land. To further improve on this we also will flexibly control for the population in 1951 within 15km, denoted by  $g(H_j)$ . We estimate  $f(\cdot)$  and  $g(\cdot)$  by Cubic Smoothing Spline functions with 10 knots. The following equation is then estimated:

$$\hat{\rho}_{jt} = \delta_1 G_j + \delta_2 M_i + f(C_j) + g(H_j) + \mu_{jt} + \epsilon_{jt}, \quad (12)$$

where  $\delta_1$ ,  $\delta_2$  and  $\mu_{jt}$  are parameters to be estimated. We repeat the same exercise but replace  $\hat{\rho}_{jt}$  by  $\hat{\sigma}_{jt}$  to measure the aggregate effects of greenbelt policies on housing consumption.

One may be worried that if  $v_{it}$  is misspecified and do not capture all the external effects of greenbelts,  $\hat{\rho}_{jt}$  may still include positive external effects of greenbelts, leading us to find a positive effect of greenbelts on aggregate prices. To mitigate this bias, we run specifications where we exclude all observations in greenbelts and within 5km of a greenbelt boundary.

We think that the identification of the aggregate effects of greenbelt policies is probably less convincing than the identification of the internal, external and supply effects with detailed fixed effects, because the construction of ‘perfect’ counterfactual greenbelts is hardly possible. However, we note that if we do not control at all for  $f(C_j)$  and  $g(H_j)$ ,  $\delta_1$  is about 40% lower, suggesting that controlling imperfectly for unobserved characteristics of the local housing markets leads to an underestimate of the true effect of supply restrictions on house prices. This is likely because areas with a lot of greenbelt land tend to be near the urban fringe in areas with a lower housing demand. Furthermore, our welfare estimates strongly suggest that the supply effects using equations (6) or (??) are very similar, suggesting that our results do not rely on the exact identification strategy based on counterfactual greenbelts.

## 5 Results

### 5.1 Local effects

**Internal and external effects.** We first show the results for the internal and external effects, where we regress house prices on the share of greenbelt land in the vicinity of the house. Table 3 reports the results.

In column (1) we estimate a naïve specification where we regress prices on the share greenbelt

TABLE 3 – REGRESSION RESULTS: EXTERNAL AND INTERNAL EFFECTS  
*(Dependent variable: the log of house price per m<sup>2</sup>)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	LA fixed effects	+ OA fixed effects	+ Housing attributes	+ Location attributes	Boundary design			
					Obs.<2.5km	Obs.<1.0km	Obs.<0.5km	
Share greenbelt land in postcode, $g_i$	-0.0212*** (0.047)	0.0204** (0.0049)	0.0101** (0.0041)	-0.0215*** (0.0045)	-0.0227** (0.0069)	-0.0106 (0.0167)	0.0109 (0.0224)	0.0023 (0.0293)
Share greenbelt land in vicinity, $v_i$	0.3172*** (0.0062)	0.1667*** (0.0200)	0.1570*** (0.0168)	0.1545*** (0.0167)	0.1598*** (0.0298)	0.2065*** (0.0584)	0.1655** (0.0803)	0.1820** (0.0882)
Housing attributes (15)	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes (11)	No	No	No	Yes	Yes	Yes	Yes	Yes
Local authority×year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Output area×year fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Output area×year×boundary segment fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
Observations	10,210,717	10,210,717	10,210,717	10,210,717	8,361,658	1,952,000	1,223,691	826,152
$R^2$	0.6481	0.8563	0.8914	0.8915	0.9024	0.9244	0.9330	0.9403

*Notes:* We set  $\delta = 3.767$ . Housing attributes include housing type dummies (flat, terraced, semi-detached, detached), the number of rooms and the number of habitable rooms, an indicator for newly built properties, the floor level of the property, the height of the property, the number of stories of the building, whether the property has a fire place, whether the property is freehold and variables capturing the energy efficiency of windows, roof, walls. Location attributes are distance rings to roads and rivers and a linear and squared term of distance to the city centre. In column (5) we only include observations that are either further away than 2.5km of a greenbelt boundary or for which the boundary does not intersect with rivers or major roads. In column (6), (7) and (8) we only include observations within respectively 2.5, 1 and 0.5km of a non-intersecting greenbelt boundary. Standard errors are clustered at the postcode level and in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$



land within the own postcode, as well as the (weighted) share of greenbelt land in vicinity. Column (1) seems to suggest that there is a discrete price jump of  $e^{-0.0212} - 1 = -2.1\%$  at the greenbelt boundary. There is also a large positive external effect: properties fully surrounded by greenbelt land are up to 37% more expensive. However, LA fixed effects are probably not capturing unobserved location characteristics that are correlated to the location of greenbelt land. We therefore include much more detailed OA fixed effects in column (2). We then find again that prices are higher close to greenbelts. For example, a property that is fully surrounded by greenbelt land is almost 17% more expensive. We do not find a discrete jump in prices at the greenbelt boundary, which suggests that internal effects are not important. When we control for housing attributes, the results are essentially unchanged. In column (4), Table 3, we include location attributes, such as the distance to the city centre, the distance to the city centre squared, as well as the share of developed land, parks and agricultural land within the postcode. The results are very similar, except for the internal effect: the share of greenbelt land in the own postcode is now negative and statistically significant.

In column (5) we use spatial differencing based on the greenbelt boundary where we include  $OA \times year \times (straight)$  boundary segment fixed effects and exclude observations close to boundaries that intersect with the course of rivers, roads and LA boundaries, as one may argue that such boundaries may capture (un)attractive features that are not a direct result of greenbelt policy. This reduces the number of observations by about 20%. The coefficient capturing the external effect is still highly statistically significant: a property that is fully surrounded by greenbelt land is about 17% more expensive. We find some evidence for internal effects due to restrictions on housing consumption, but this effect is reduced to zero when we focus on areas close to greenbelt boundaries in Columns (6), (7) and (8). More specifically, we make further restrictions by only selecting observations within 2.5, 1 and 0.5km respectively. The results are essentially unchanged: we do find strong evidence for an external effect of greenbelt policy of about 15-20%, suggesting that households do value open space. We do not find convincing evidence for an internal effect.

We also test the impact of  $g_i$  and  $v_i$  on house size close to the boundary. When house size is restricted, or when households adjust their housing consumption because of an external effect,

TABLE 4 – POISSON REGRESSION RESULTS: SUPPLY EFFECTS  
(Dependent variable: the number of dwellings in a postcode)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	LA fixed effects	+ OA fixed effects	+ Location attributes		Boundary design		
					Obs.<2.5km	Obs.<1.0km	Obs.<0.5km
Share of postcode in greenbelt, $g_i$	-0.6593*** (0.0273)	-0.9397*** (0.0082)	-0.9157*** (0.0078)	-0.8576*** (0.0100)	-1.2077*** (0.0197)	-1.2476*** (0.0272)	-1.3152*** (0.0335)
Area size of postcode ( $\log$ )	0.1113*** (0.0078)	0.3210*** (0.0012)	0.3363*** (0.0011)	0.3354*** (0.0012)	0.4837*** (0.0036)	0.5310*** (0.0050)	0.5699*** (0.0071)
Location attributes (8)	No	No	Yes	Yes	Yes	Yes	Yes
Local authority fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Output area fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes
Output area×boundary segment fixed effects	No	No	No	Yes	Yes	Yes	Yes
Observations	1,310,750	1,310,750	1,310,750	1,094,110	228,940	142,203	96,646
Log-likelihood	-1.040e+07	-6.654e+06	-6.381e+06	-4.600e+06	-583216	-254857	-113741

Notes: These are Poisson models, so the coefficients are interpreted as (semi-)elasticities. In column (4) we only include observations that are either further away than 1.5km of a greenbelt boundary or for which the boundary does not intersect with rivers or major roads. In column (5) we only include observations within 1.5km of a non-intersecting greenbelt boundary. Standard errors are clustered at the postcode level and in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

we would expect to find an effect of greenbelt policy on local house size. In Table B1 in Appendix B we estimate the same set of specifications as in Table 3 but we replace the dependent variable house price by house size. The results do not deliver a pretty clear picture. We find some evidence that houses are larger when there is more greenbelt land in the vicinity, likely because the historic urban fringe host larger detached and semi-detached properties.

**Supply effects.** The next step is to identify supply effects. We show the results where we regress the number of dwellings in a postcode on the share of the postcode in the greenbelt. Table 4 reports the results of Poisson regressions, as the dependent variable is a count variable. Hence, the coefficients are interpreted as semi-elasticities.

In column (1) we only include local authority fixed effects and control for the area size of the postcode. The results seem to suggest that postcodes in greenbelts have  $e^{-0.6594} - 1 = 48\%$  fewer dwellings. This effect becomes even stronger once we include OA fixed effects in column (2). This leads to a much stronger supply effect of about 61%. In column (3) we control for distance to rivers, roads, and the city centre, leading to almost identical results.<sup>13</sup> Column (4), Table 4, focuses on the boundary design, where we exclude observations that are near boundaries that intersect with the course of roads, rivers and LA boundaries, and include OA×boundary segment

<sup>13</sup>We do not control for land use here, because land use patterns are likely the result of greenbelt policy.

fixed effects. The coefficient implies that the number of dwellings is reduced by 57% when the postcode is in the greenbelt. In column (5), (6) and (7) we only include observations within 2.5, 1 and 0.5km from a greenbelt boundary respectively. The effect becomes even somewhat stronger. For example, the coefficient in column (7) suggests that the number of dwellings in greenbelts is reduced by 73%. Hence, the results unequivocally suggest that greenbelt policy strongly restricts the number of residential properties in greenbelts.

## 5.2 Aggregate effects

When the number of dwellings is restricted this may have aggregate implications on the prices of dwellings in the housing market. Furthermore, when overall prices are higher, we may expect that households will consume less housing. To estimate these aggregate effects we use the estimated fixed effects  $\log \bar{p}_{jt} \equiv \hat{\rho}_{jt}$  for prices net of external and internal effects, obtained from column (5) in Table 3 and housing consumption net of external and internal effects  $\log \bar{h}_{jt} \equiv \hat{\sigma}_{jt}$ , obtained from column (5) in Table B1 in Appendix B.<sup>14</sup>

In Table 5 we regress prices net of external and internal effects on the share of greenbelt land within 15km. Column (1) only controls for year trends and the share of counterfactual greenbelt land. The results do not seem to suggest that prices are higher in areas with a higher share of greenbelt land. It seems that areas with more counterfactual greenbelt land are substantially cheaper, although the standard error is too high to draw strong conclusions. In column (2) we improve on this by adding Cubic Splines of the share of counterfactual greenbelt land and the population density within 15km in 1951. We find a sizeable but imprecise effect of having more greenbelt land within 15km of the property. Column (3) adds controls, such as population density, the share of land devoted to Areas of Outstanding Natural Beauty (AONBs) and national parks, as well as government region $\times$ year fixed effect. The point estimate is virtually identical to the estimate in the previous column, but it is now highly statistically significant. The coefficient implies that a 10 percentage points increase in the share of greenbelt land within 15km raises prices by 2.3%. Given an average share of greenbelt land of 0.27, this means that

---

<sup>14</sup>We also estimate the following regressions for other specifications in Tables 3 and B1, leading to similar results. Unsurprisingly, we do not get reliable estimates for  $\hat{\rho}_{jt}$  and  $\hat{\sigma}_{jt}$  for locations that are further away from the greenbelt boundary in the specifications where we focus on observations close to the greenbelt boundary (see columns (6), (7) and (8) in Table 3 and Table B1). We therefore decided to focus on the estimation of  $\hat{\rho}_{jt}$  and  $\hat{\sigma}_{jt}$  based on the whole sample.

TABLE 5 – AGGREGATE SUPPLY EFFECTS: PRICES  
*(Dependent variable: Output area×year fixed effect,  $\hat{\rho}_{jt}$ )*

	(1)	(2)	(3)	(4)	(5)
		+ $f(\cdot), g(\cdot)$	+ Location attributes, f.e.	+ Output area, characteristics,	Outside greenbelts
Share greenbelt land, < 15km	0.0963 (0.1942)	0.2328 (0.1626)	0.2328*** (0.0620)	0.2016*** (0.0501)	0.1560** (0.0621)
Share counterfactual greenbelt land, < 15km	-0.3354 (0.2385)				
Population density, < 15km ( <i>log</i> )			0.0630 (0.0433)	0.0853*** (0.0315)	0.0753** (0.0313)
Standard deviation of altitude, < 15km ( <i>log</i> )			0.0137 (0.0218)	0.0100 (0.0149)	0.0263 (0.0163)
Share AONB, < 15km			0.0547 (0.0718)	0.0174 (0.0481)	0.0329 (0.0501)
Share national park, < 15km			0.4703*** (0.1434)	0.3112*** (0.1045)	0.2810** (0.1095)
Share developed land, < 15km			0.1011 (0.1542)	0.1159 (0.1754)	0.0849 (0.2526)
Share water bodies, < 15km			-1.6984** (0.7428)	-0.4962 (0.4346)	-0.9897* (0.5173)
Share counterfactual greenbelt land, $f(\cdot)$	No	Yes	Yes	Yes	Yes
Population density in 1951 < 15km, $g(\cdot)$	No	Yes	Yes	Yes	Yes
Output area characteristics	No	No	No	Yes	Yes
Government region×year fixed effects	No	No	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	2,439,647	2,439,647	2,439,647	2,439,647	1,391,289
$R^2$	0.4725	0.5925	0.7286	0.8098	0.8059

*Notes:* The dependent variable is obtained from the specification listed in column (5) in Table 3. In column (5) we only include OAs which are outside greenbelts and further than 5km from a greenbelt boundary. Standard errors are clustered at the travel-to-work-area level and in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

without greenbelts, prices would be on average 6.2% lower.

Column (4), Table 5, further includes output area characteristics and in column (5) we only include OAs that are outside greenbelts and further than 5km from an inner or outer greenbelt boundary. The results are very similar, although the coefficient is somewhat lower: a 10 percentage points increase in the share of greenbelt land within 15km implies a price increase of 1.6%. Hence, these results suggest that housing becomes more expensive once a housing market is more restricted by greenbelts.

When looking at the control variables, note that areas that are denser tend to be more expensive (with an elasticity of about 0.07), in line with the literature (Albouy 2016). It seems that the share of national parks within 15km also raises prices substantially. Note that we do not control for any external or internal effects caused by national parks, so that this association may partly

TABLE 6 – AGGREGATE SUPPLY EFFECTS: HOUSING CONSUMPTION  
*(Dependent variable: Output area×year fixed effect,  $\hat{\sigma}_{jt}$ )*

	(1)	(2)	(3)	(4)	(5)
		+ $f(\cdot), g(\cdot)$	+ Location attributes, f.e.	+ Output area, characteristics,	Outside greenbelts
Share greenbelt land, < 15km	-0.2044*** (0.0305)	-0.1519*** (0.0304)	-0.1302*** (0.0223)	-0.1676*** (0.0254)	-0.1689*** (0.0212)
Share counterfactual greenbelt land, < 15km	-0.0456* (0.0263)				
Population density, < 15km ( <i>log</i> )			-0.0812*** (0.0151)	-0.0647*** (0.0133)	-0.0758*** (0.0159)
Standard deviation of altitude, < 15km ( <i>log</i> )			-0.0220*** (0.0073)	-0.0267*** (0.0063)	-0.0283*** (0.0064)
Share AONB, < 15km			0.0259 (0.0475)	0.0164 (0.0532)	-0.0111 (0.0550)
Share national park, < 15km			0.1521** (0.0627)	0.0907 (0.0647)	0.1028 (0.0624)
Share developed land, < 15km			0.0296 (0.0490)	0.0338 (0.0685)	-0.0366 (0.0281)
Share water bodies, < 15km			-0.1991 (0.2874)	0.3247 (0.2095)	0.0774 (0.2020)
Share counterfactual greenbelt land, $f(\cdot)$	No	Yes	Yes	Yes	Yes
Population density in 1951 < 15km, $g(\cdot)$	No	Yes	Yes	Yes	Yes
Output area characteristics	No	No	No	Yes	Yes
Government region×year fixed effects	No	No	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	1,251,323	1,251,323	1,251,323	1,251,323	710,352
$R^2$	0.0712	0.0887	0.1142	0.3119	0.3041

*Notes:* The dependent variable is obtained from the specification listed in column (5) in Table B1 in Appendix B. In column (5) we only include OAs which are outside greenbelts and further than 5km from a greenbelt boundary. Standard errors are clustered at the travel-to-work-area level and in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

capture positive amenity effects of living close to national parks. Furthermore, the average share of national parks within 15km is very low compared to the share of greenbelt land within 15km (1.5% versus 27% respectively). Hence, the supply effects are expected to be quantitatively much less important compared to the effects of greenbelts.

We may expect that more expensive homes may also mean that housing consumption is affected, *i.e.* households may prefer smaller homes once dwellings are more expensive. We test this in Table 6 where we regress house size net of internal and external effect on the share of greenbelt land within 15km. Column (1) suggests that houses are about 2% smaller when the share of greenbelt within 15km land increases by 10 percentage points. The results are similar once we include Cubic Smoothing Splines of the share of counterfactual greenbelt land and population density in 1951 in column (2). We further control for location attributes and government region×year fixed effects in column (3) and for output area characteristics in column (5), leading

TABLE 7 – COUNTERFACTUAL SCENARIOS AND WELFARE

	Scenario 1: 10% reduction in greenbelt land			Scenario 2: No greenbelts			Scenario 3: Counterfactual greenbelts		
	$\Delta$ dwellings	$\Delta$ ext+int	$\Delta$ supply	$\Delta$ dwellings	$\Delta$ ext+int	$\Delta$ supply	$\Delta$ dwellings	$\Delta$ ext+int	$\Delta$ supply
All	0.3406	-£524	£1843	1.3647	-£2178	£9753	-2.5878	£1830	-£13246
Metropolitan greenbelt	0.1440	-£136	£1184	0.6919	-£621	£6925	-0.7357	£392	-£5133
North-West greenbelt	0.1827	-£222	£659	0.6316	-£766	£3018	-0.6722	£489	-£2515
West-Midlands greenbelt	0.0463	-£63	£208	0.1883	-£246	£1200	-0.3036	£324	-£1466

*Notes:* The change in the number of dwellings, as well as the annual monetary changes are in millions. We assume a discount rate of 2.5%. We base these calculations on estimates from column (5) in Table 3 and Table B1, as well as on column (4) in Tables 4, 5 and 6.

to very similar results. When we only focus on observations outside and further than 5km from the greenbelt boundary, the results are essentially unaffected. Hence, it seems that because houses are more expensive due to greenbelt policies, households occupy smaller properties. The results also imply that the elasticity of prices with respect to house size is about 1.

### 5.3 Counterfactual analysis and welfare

We now put the different estimates together to consider the welfare effects of greenbelt policies. We consider these estimates as back-of-the-envelope calculations, as we have to make some additional assumptions. First, we assume that the external effects are the same for different greenbelts, and that the local supply effect we measure in Table 4 extends to the whole greenbelt. Second, we use an implied discount rate of 2.5%, in line with evidence from England from [Bracke et al. \(2017\)](#). Third, we assume that all welfare effects can be measured by changes in house prices. While the latter assumption is unlikely to hold, in the next section we show that greenbelt policy does not seem to relate to other (aggregate) effects that may be omitted (such as longer commutes, recreational visits to the greenbelt, more housing CO<sub>2</sub> emissions). Despite those assumptions, we expect that the numbers will at least give a good indication on the order of magnitude of the different effects.

In Table 7 we consider three scenarios. In the first scenario we consider a 10% reduction in greenbelt land from the inner greenbelt boundaries. That is, we first determine inner greenbelt boundaries and shift the boundary approximately 800m outwards so that the total greenbelt land is reduced by 10%. The results first suggests that this allows for the construction of 341 thousand new dwellings, which is about 2% of the dwelling stock.

Overall, the reduction in the external effect is approximately equal to £0.5 billion a year. However, this effect is considerably lower than the overall gain because greenbelt land can be developed and housing units can be sold. The supply is effect about £1.8 a year, so reducing the amount of greenbelt land by 10% seems to have a strong positive welfare effect.

We also calculate the welfare effects for the main greenbelts separately. We first estimate greenbelt-specific regressions to obtain the external, internal and supply effects, see Section 5.4 for details. In the Metropolitan greenbelt (London) the ratio external/supply effects is only 10%, while it is around 30% in the North-West (Manchester-Liverpool) and West-Midlands (Birmingham) greenbelts. Hence, particularly in and around London, the welfare costs of greenbelt policy are high.<sup>15</sup>

We note that calculating supply effects using equation (5) or (6) lead to essentially the same outcomes and only differ maximally 1%. The reason is that while prices are lower when greenbelts are removed, housing consumption increase at the same time.<sup>16</sup> Although the fact that aggregate changes in prices and housing consumption have little implications for the overall welfare estimates, they have strong distributional consequences. More specifically, households that rent now pay lower rents when greenbelt land is reduced. Hence, alleviating greenbelt policy seems to increase housing affordability for a non-negligible share of the population and allows them to occupy larger properties.

We also consider an alternative counterfactual scenario by investigating the implications when greenbelt policy would be completely abandoned. We show that based on our estimates, this implies the construction of about 1.4 million new dwellings, which is about 6% of the total dwelling stock. Back-of-the-envelop calculations by Cheshire (2014) suggested that housing construction between 1994 and 2012 fell short by about 2 million homes. Hence, abandoning greenbelt policy will not completely solve the issue of a presumed housing shortage, but will substantially alleviate it.

---

<sup>15</sup>Because we find that supply effects are stronger for the main greenbelts, we find that if one adds up the reduction in supply for the three different greenbelts, the change in supply is bigger than when looking at the overall estimate. Hence, our overall estimates are likely to provide underestimates.

<sup>16</sup>Hence, if one thinks that the identification strategy using counterfactual greenbelts to identify aggregate effects is invalid, this will likely have limited repercussions for the general conclusion that greenbelt policies are welfare decreasing.

Abandoning greenbelt policy would also lead to a loss the amenity value of open space, which is valued by households. The total monetary loss in the combined external and internal effects are £2.2 billion a year. However, the supply effect due to an increased number of dwellings amounts to £9.8 billion a year, which implies a welfare loss of greenbelt policies of £7.6 billion a year – which is about 0.5% of the GDP in England.

In Scenario 3 we estimate the welfare costs by considering that counterfactual greenbelts would be actual greenbelts, based on historic population concentrations and density in 1951 (see Figure A2 in Appendix A). At the aggregate, this implies that the amount of greenbelt land is almost doubled and the number of dwellings would be reduced by about 2.5 million. If this would happen, the welfare loss is about £11.5 billion a year.

#### 5.4 Extensions and sensitivity

**Local effects.** We first consider some extensions to the regressions aiming to identify the external and internal effect. These are described in detail in Appendix B.1. First, we test our specification of  $v_i$  by including 250m and 500m distance rings instead of specifying  $v_i$ . We show that the external effect becomes statistically insignificant beyond 500m, suggesting that our choice of  $\delta$  makes sense. We also show that  $\delta = 3.767$  minimises the Mean Squared Error (given the specification in column (5), Table 3), providing additional evidence for the choice of  $\delta$ .

We also make a distinction between accessible and non-accessible greenbelt land, as well as a distinction between agricultural and non-agricultural greenbelt land. The results do not seem to suggest that the effects are statistically significantly different.

Furthermore, we investigate whether the external effect of greenbelt land is constant over time. Indeed, we do not find that an interaction of  $v_i$  with the year of observation is statistically significant.

We also consider heterogeneity in the external effect between different greenbelts. Because our results become imprecise once we have fewer observations, we only estimate the effects for greenbelts for which we have more than 500 thousand sales.<sup>17</sup> The results suggest that the

---

<sup>17</sup>This may seem a lot, but note that because we include OA×year×boundary segment fixed effects and because the number of sales close to greenbelt boundaries is limited one needs many observations to identify the effects of interest.



external effect in the North-West and West-Midlands greenbelts are very similar to the baseline estimate. For London the external effect is about 50% lower. We also find stronger supply effects in all the major greenbelts, likely because the pressure to develop greenbelts is much stronger in those greenbelts. We already showed in the previous subsection that this implies rather substantial welfare losses.

**Greenbelt visits: geocoded pictures.** One may argue that the benefits of greenbelts may extend beyond 1km, for example because people may visit greenbelts to recreate, which would be in line with one of the intended goals of greenbelt policy (“improving access to the open countryside, by providing opportunities for outdoor sport and recreation”). If this is the case, also households further away from the greenbelt may appreciate greenbelt land and the estimate of the external effect may be an underestimate. However, these less localised effects are arguably difficult to capture by looking at (local) house price differentials. We propose another approach to test whether these effects are important. We exploit data on about 10 million geocoded pictures from *FlickrR*, an online image hosting service, between 2000 and 2017. We expect locations that offer aesthetic amenities will have more visits by locals and tourist and therefore a higher picture density. Indeed, [Gaigné et al. \(2018\)](#) show that there is a strong positive correlation between picture density and historic amenities or geographical variables, such as access to open water or open space. We address some issues with the data (see Appendix B.2 for more details). Since we have information on users’ identifiers, we can distinguish between natives’ and tourists’ pictures by keeping users who take pictures for at least 6 consecutive months between 2004 and 2017 in England. It seems unlikely that tourists stay for 6 consecutive months in England.<sup>18</sup>

The results reported in Appendix B.2 show that, when controlling for density, both for tourists and natives, the picture density is lower in greenbelts. The preferred specification indicates that there are 44% fewer pictures made by tourists and 32% fewer pictures made by natives. Having said that, the results become statistically insignificant once we focus on areas closer to the greenbelt boundary. These results indicate that greenbelts are unlikely to be main recreation destinations and hence, this provides once more evidence that external effects of greenbelts are expected to be very local.

---

<sup>18</sup>The correlation between tourists’ and natives’ pictures is 0.748.

**Commuting and energy consumption.** One may argue that there are aggregate effects that we do not take into account when calculating the welfare effects of greenbelt policy. For example, urban development may ‘leapfrog’ over the greenbelts leading to an inefficient urban structure and longer commutes (see [Levkovich et al. 2017](#), [Cheshire et al. 2018](#)). If that is the case, one would expect aggregate housing prices to be lower in areas with more greenbelt land, suggesting that we may underestimate the aggregate price effects of greenbelts. In [Appendix B.3](#) we show, however, that commuting distances are not related to the share of greenbelt land in the vicinity.

Another critique could be that due an inefficient urban structure per capita energy consumption may increase, which leads to more CO<sub>2</sub> emissions. We test this by exploiting data from EPCs on the predicted CO<sub>2</sub> emissions per property. We show in [Appendix B.3](#) that areas with more greenbelt land do not seem to have higher (or lower) CO<sub>2</sub> emissions.

In other words, other aggregate effects that may translate into welfare do not seem to be related to the amount of greenbelt land in vicinity.

**Robustness of aggregate effects.** In [Appendix B.4](#) we test for sensitivity of the aggregate effects. We first include two controls that capture the regulatory restrictiveness: the refusal rate of major and minor construction projects by local planning authorities, which has been used before by [Hilber & Vermeulen \(2016\)](#) and [Cheshire et al. \(2018\)](#). Hence, we evaluate the effects of greenbelt policy, *conditional* on other restrictions. This does not seem to influence our results.

We further show robustness to the calculation of the residuals,  $\hat{\rho}_{jt}$  and  $\hat{\sigma}_{jt}$ , used as dependent variables as these are prices and house size conditional on external and internal effects. More specifically, we use the results reported in column (1) in [Tables 3](#) and [B1](#) to obtain the residuals. This does not change the results.

We also show that our results are insensitive to the specification of the  $f(C_j)$  and  $g(H_j)$ : the nonparametric functions of the share of counterfactual greenbelt land and the population density in 1951. Instead of using splines we use a 5<sup>th</sup>-order polynomial to approximate  $f(C_j)$  and  $g(H_j)$ , leading to the same results.

Furthermore, we test for robustness with respect to the construction of predicted greenbelts (we

use 10km rings around cities instead of 15km) and the definition of the housing market. That is, we repeat results when we take into account the share of greenbelt land in respectively 10km and 50km. For the latter, the effects are somewhat amplified.

## 6 Conclusions

In this paper, we have investigated the economic effects of greenbelt policy that prohibit new construction beyond a predefined boundary. We focus on England, where 13% of the land area is designated as greenbelt land. The greenbelt boundaries were constituted in the 1950s – a time where cities were much smaller – and hardly change since.

We analyse the local external, internal and supply effects of these policies on the housing market. First, we use spatial differencing to show that the external effects of these regulations are substantial (about 15-20%) but very local (within 1km). In contrast to the previous literature, we find no evidence for internal or ‘own-lot’ effects. We further show that supply effects are important: greenbelt policy reduce the the number of housing units in greenbelts by about 80%.

The reduction in the supply of housing likely leads to higher prices throughout the housing market and may impact housing consumption. Using counterfactual greenbelts based on the population in 1931, we estimate that aggregate prices rise by about 7.5%, while the reduction in housing consumption is of a similar magnitude. We do not find evidence for other general equilibrium effects, *e.g.* through longer commutes or more energy consumption.

We show that greenbelt policies imply a negative welfare cost of about £7.5 billion a year (0.5% of England’s GDP). Greenbelt policies also have distributional implications: households that rent have to pay higher house prices, while home owners benefit from the aggregate increase in prices. Hence, greenbelt policies seem to reduce housing affordability for a non-negligible share of the population, essentially forcing them to occupy smaller properties.

## References

- Albouy, D. (2016), ‘What are Cities Worth? Land Rents, Local Productivity, and the Total Value of Amenities’, *Review of Economics and Statistics* .
- Anderson, S. & West, S. (2006), ‘Open Space, Residential Property Values, and Spatial Context’, *Regional Science and Urban Economics* **36**(6), 773–789.
- Bontemps, C., Simioni, M. & Surry, Y. (2008), ‘Semiparametric Hedonic Price Models: Assessing the Effects of Agricultural Nonpoint Source Pollution’, *Journal of Applied Econometrics* **23**(6), 825–842.
- Bracke, P., Pinchbeck, E. & Wyatt, J. (2017), ‘The Time Value of Housing: Historical Evidence on Discount Rates’, *The Economic Journal* (Forthcoming).
- Brander, L. & Koetse, M. (2011), ‘The Value of Urban Open Space: Meta-analyses of Contingent Valuation and Hedonic Pricing Results’, *Journal of Environmental Management* **92**(10), 2763–2773.
- Brueckner, J. (2001), ‘Urban Sprawl: Lessons from Urban Economics’, *Brookings-Wharton Papers on Urban Affairs* **2001**(1), 65–97.
- Cheshire, P. (2014), ‘Turning Houses Into Gold’, *CentrePiece* **Spring**, 14–18.
- Cheshire, P., Hilber, C. & Koster, H. (2018), ‘Empty Homes and Longer Commutes: The Unintended Consequences of More Restrictive Planning’, *Journal of Public Economics* **158**(2), 126–151.
- Economist (2017), ‘Britain’s Delusions About the Green Belt Cause Untold Misery’, *The Economist* .
- Gaigné, C., Koster, H., Moizeau, F. & Thisse, J. (2018), ‘Who Lives Where in the City? Amenities, Commuting and Income Sorting’, *CEPR Discussion Paper 11958* .
- Glaeser, E., Gyourko, J. & Saks, R. (2005), ‘Why have Housing Prices Gone up?’, *American Economic Review* .

- Glaeser, E. & Kahn, M. (2004), ‘Sprawl and Urban Growth’, *Handbook of regional and urban economics* .
- Glaeser, E. & Ward, B. (2009), ‘The Causes and Consequences of Land Use Regulation: Evidence from Greater Boston’, *Journal of Urban Economics* **65**(3), 265–278.
- Green, R., Malpezzi, S. & Mayo, S. (2005), ‘Metropolitan-Specific Estimates of the Price Elasticity of Supply of Housing, and their Sources’, *American Economic Review* **95**(2), 334–339.
- Hilber, C. & Vermeulen, W. (2016), ‘The Impact of Supply Constraints on House Prices in England’, *Economic Journal* **126**(591), 358–405.
- Ihlanfeldt, K. (2007), ‘The Effect of Land Use Regulation on Housing and Land Prices’, *Journal of Urban Economics* **61**(3), 420–435.
- Irwin, E. (2002), ‘The Effects of Open Space on Residential Property Values’, *Land Economics* **78**(4), 465–480.
- Koster, H., Van Ommeren, J. & Rietveld, P. (2012), ‘Bombs, Boundaries and Buildings: a Regression-discontinuity Approach to Measure Costs of Housing Supply Restrictions’, *Regional Science and Urban Economics* **42**(4), 631–641.
- Levkovich, O., Rouwendal, J. & Van Ommeren, J. (2017), ‘The Impact of Highways on Population Redistribution: The Role of Land Development Restrictions’, *Tinbergen Institute Discussion Paper 2017-109/VIII* .
- Mayer, C. & Somerville, C. (2000), ‘Land use regulation and new construction’, *Regional Science and Urban Economics* **30**(6), 639–662.
- Rouwendal, J. & Van der Straaten, J. (2008), ‘The Costs and Benefits of Providing Open Space in Cities’, *Tinbergen Institute Discussion Paper TI 2008-001/3* .
- Saiz, A. (2010), ‘The Geographic Determinants of Housing Supply’, *The Quarterly Journal of Economics* **125**(3), 1253–1296.
- Turner, M., Haughwout, A. & Van der Klaauw, W. (2014), ‘Land Use Regulation and Welfare’, *Econometrica* **82**(4), 1341–1403.

Vermeulen, W. & Rouwendal, J. (2014), 'On the Value of Foregone Open Space in Sprawling Cities', *Journal of Regional Science* **54**(1), 61–69.

## A Data appendix

In this data appendix we first show the descriptives for the matched housing transactions dataset. Table A1 reports the descriptives for all the variables included in the regression. Table A2 reports a similar set of descriptives for the full *EPCs* dataset, which we use to calculate the internal effect in Table B1 and to estimate housing consumption net of internal and external effects.

In Figure A1 we evaluate the implications for different choices of  $\delta$  in the calculation of the external effect of greenbelt land,  $v_i$  (see equation (7)). We display the weight of a hectare greenbelt land for the baseline  $\delta = 3.767$ . In line with the empirical results shown in Appendix B most of the weight is then within 500m. This holds even more so when  $\delta = 5$ . When  $\delta = 2.5$ , a hectare of greenbelt land at 750m has still a weight of about 0.2.

To construct counterfactual greenbelts we exploit data on the population in Parishes from the 1951 census. Parishes are the lowest unit for which the data are available, but are rather small. The median size is 847ha. In line with the suggestion by Duncan Sandy, the Minister of Housing at that time, greenbelts were mainly implemented around larger cities of at least 100 thousand inhabitants. To identify ‘large urban areas’ we first select parishes with a population density of at least 10 people per hectare, about 5% of the Parishes. We then amalgamate all those areas and keep 37 amalgamated urban areas that have a population of more than 100 thousand inhabitants. We then draw circles of 15km around each of these urban areas, in line with the suggestion of Duncan Sandy (*"The Development Plans submitted by the local planning authorities for the Home Counties provide for a Green Belt, some 7 to 10 miles deep, [...]"*). The last step is to erase the areas in the counterfactual greenbelts with a density of at least 10 persons per hectare because that land has already been converted to built-up land, and are therefore not part of the greenbelt.

Figure A2 shows a map with the counterfactual greenbelts. We are able to predict the location of most greenbelts very well. Comparing the outcome with Figure 2, we can see that the actual greenbelts around London, Birmingham and Manchester and Liverpool match the counterfactual greenbelts to a large extent. Only the Cheltenham/Gloucester greenbelt, as well as the greenbelt around Cambridge are not included as counterfactual greenbelt, because those cities had a

TABLE A1 – DESCRIPTIVE STATISTICS FOR LAND REGISTRY DATA

	(1)	(2)	(3)	(4)
	mean	sd	min	max
Price per m <sup>2</sup>	1,768.9130	1,277.2650	100.0000	10,000.0000
Size of the property ( <i>in m</i> <sup>2</sup> )	87.3623	31.5952	25.0000	250.0000
Distance to greenbelt boundary ( <i>in km</i> )	16.4492	29.2659	0.0000	296.8268
Share of postcode in greenbelt	0.0358	0.1631	0.0000	1.0000
Share of greenbelt land in vicinity, $\delta = 3.767$	0.0966	0.1865	0.0000	1.0000
Share of agricultural greenbelt land in vicinity, $\delta = 3.767$	0.0388	0.0853	0.0000	0.9438
Share of accessible greenbelt land in vicinity, $\delta = 3.767$	0.0139	0.0389	0.0000	0.9219
Share greenbelt land 0-250m	0.0472	0.1508	0.0000	1.0000
Share greenbelt land 250-500m	0.0858	0.1945	0.0000	1.0000
Share greenbelt land 500-1000m	0.1317	0.2343	0.0000	1.0000
Share greenbelt land 1000-1500m	0.1697	0.2616	0.0000	1.0000
Housing type – flat	0.0317	0.1753	0.0000	1.0000
Housing type – terraced	0.4118	0.4922	0.0000	1.0000
Housing type – semi-detached	0.1137	0.3174	0.0000	1.0000
Housing type – detached	0.4428	0.4967	0.0000	1.0000
House newly built	0.0690	0.2535	0.0000	1.0000
Number of habitable rooms	5.1040	3.7098	1.0000	25.0000
Number of heated rooms	4.8558	3.8124	1.0000	25.0000
Floor level	0.1006	0.4909	0.0000	21.0000
Height of floor ( <i>in m</i> )	2.4393	0.1941	0.0000	5.0000
Storeys of building	2.0333	0.7103	0.0000	100.0000
Property has fireplace	0.1956	0.3967	0.0000	1.0000
Energy efficiency of windows	0.5481	0.2294	0.0000	1.0000
Energy efficiency of roof	0.4922	0.2610	0.0000	0.7500
Energy efficiency of walls	0.4059	0.3373	0.0000	1.0000
Overall energy efficiency	0.5942	0.1328	0.0000	1.0000
Ownership type – freehold	0.8246	0.3803	0.0000	1.0000
Distance to river ( <i>in km</i> )	0.8640	0.9249	0.0000	47.5119
Distance to street ( <i>in km</i> )	0.0165	0.0071	0.0000	0.8860
Share developed land in postcode	0.8291	0.3093	0.0000	1.0000
Share parks in postcode	0.0331	0.1143	0.0000	1.0000
Share agricultural land in postcode	0.0849	0.2011	0.0000	1.0000
Distance to the city centre ( <i>in km</i> )	35.7809	33.5061	0.0802	313.6746

Notes: The number of observations is 10,210,717.



TABLE A2 – DESCRIPTIVE STATISTICS FOR EPC DATA

	(1)	(2)	(3)	(4)
	mean	sd	min	max
Size of the property ( <i>in m<sup>2</sup></i> )	84.9680	36.3591	25.0000	250.0000
Distance to greenbelt boundary ( <i>in km</i> )	16.7975	30.0692	0.0000	298.6461
Share of postcode in greenbelt	0.0377	0.1693	0.0000	1.0000
Share of greenbelt land in vicinity, $\delta = 3.767$	0.0927	0.1877	0.0000	1.0000
Share of agricultural greenbelt land in vicinity, $\delta = 3.767$	0.0368	0.0850	0.0000	0.9704
Share of accessible greenbelt land in vicinity, $\delta = 3.767$	0.0132	0.0391	0.0000	0.9971
Share greenbelt land 0-250m	0.0476	0.1563	0.0000	1.0000
Share greenbelt land 250-500m	0.0827	0.1954	0.0000	1.0000
Share greenbelt land 500-1000m	0.1246	0.2316	0.0000	1.0000
Share greenbelt land 1000-1500m	0.1598	0.2571	0.0000	1.0000
Housing type – flat	0.2510	0.4336	0.0000	1.0000
Housing type – bungalow	0.0912	0.2879	0.0000	1.0000
Housing type – house	0.6331	0.4820	0.0000	1.0000
Housing type – maisonette	0.0247	0.1553	0.0000	1.0000
Number of habitable rooms	5.9635	5.9055	1.0000	25.0000
Number of heated rooms	5.8043	5.9773	0.0000	25.0000
Floor level	0.3230	1.0469	0.0000	21.0000
Height of floor ( <i>in m</i> )	2.4259	0.1631	0.0000	5.0000
Storeys of building	2.1589	1.0005	0.0000	100.0000
Property has fireplace	0.1951	0.3963	0.0000	1.0000
Energy efficiency of windows	0.5792	0.2488	0.0000	1.0000
Energy efficiency of roof	0.4999	0.2490	0.0000	0.7500
Energy efficiency of walls	0.4766	0.3445	0.0000	1.0000
Everall energy efficiency	0.6217	0.1448	0.0000	1.0000
Distance to river (in km)	0.8757	0.9549	0.0000	49.7727
Distance to street (in km)	0.0178	0.0091	0.0000	2.0704
Share developed land in postcode	0.7967	0.3416	0.0000	1.0000
Share parks in postcode	0.0327	0.1149	0.0000	1.0000
Share agricultural land in postcode	0.0977	0.2181	0.0000	1.0000
Distance to the city centre ( <i>in km</i> )	35.0959	34.4090	0.0000	316.2100

*Notes:* The number of observations is 13,593,673.

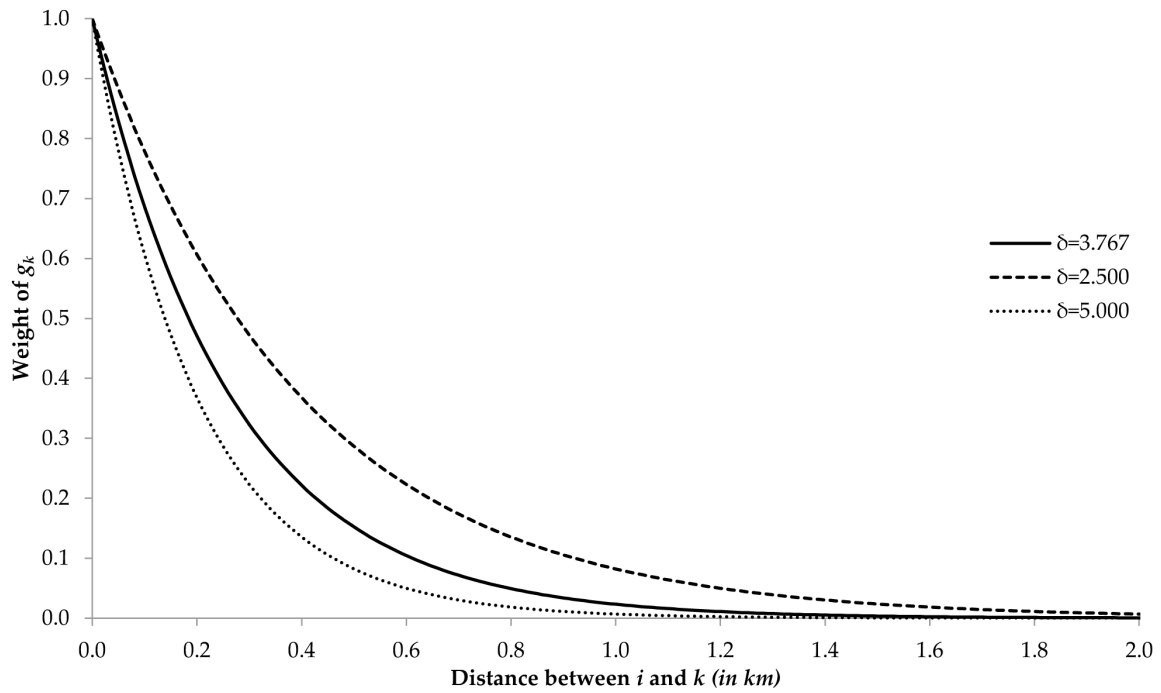


FIGURE A1 – THE DECAY PARAMETER  $\delta$

population lower than 100 thousand in 1951. On the other hand, reasonably large cities such as Leicester, Norwich, Middlesbrough and Plymouth do not have greenbelts, although you would expect a greenbelt around those areas based on the 1951 population.



FIGURE A2 – COUNTERFACTUAL GREENBELTS IN ENGLAND

## B Other results and sensitivity analyses

### B.1 Local effects

We first test the impact of  $g_i$  and  $v_i$  on house size close to the greenbelt boundary, in line with equation (10). When house size is restricted, or when households adjust their housing consumption because of higher prices caused by an external effect, we would expect to find an effect of greenbelt policy on local house size. In Table B1 we replicate the same set of specifications as in Table 3 but we choose the log of house size as dependent variable.

In column (1) we estimate a naïve specification where we regress house size on the share greenbelt land within the own postcode, as well as the (weighted) share of greenbelt land in vicinity, according to equation (7). Column (1) seems to suggest that there is a discrete price jump in house size of  $e^{0.0409} - 1 = 4.2\%$  at the greenbelt boundary. There is also a sizeable external effect: properties fully surrounded by greenbelt land are up to 13.7% more expensive. We therefore include much more detailed OA fixed effects in column (2). Again, houses are larger close to greenbelts. When we control for housing attributes in column (3), we do not find a discrete jump in prices at the greenbelt boundary. In column (4), Table B1, we include location attributes, such as the distance to the city centre, the distance to the city centre squared, as well as the share of developed land, parks and agricultural land within the postcode. The results are very similar, except for the internal effect: the share of greenbelt land in the own postcode is now negative and statistically significant.

In column (5) we use spatial differencing based on the greenbelt boundary where we include  $OA \times year \times (straight)$  boundary segment fixed effects and exclude observations close to boundaries that intersect with the course of rivers, roads and LA boundaries. The coefficient capturing the external effect is still highly statistically significant: a property that is fully surrounded by greenbelt land is about 8.9% larger. We find some evidence that houses are smaller inside greenbelts, but this effect is not robust once we focus on areas close to the greenbelt boundaries in Columns (6), (7) and (8), where we select observations within 2.5, 1 and 0.5km respectively. Still, properties with a high share of greenbelt land around them seem to be somewhat larger (up to 10.5%), but the effect is only marginally significant. The interpretation is that, historically, prices were lower near the urban fringe. Hence, houses were larger once they were further away

TABLE B1 – REGRESSION RESULTS: INTERNAL EFFECTS FOR HOUSE SIZE  
*(Dependent variable: the log of house size (in m<sup>2</sup>))*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	LA fixed effects	+ OA fixed effects	+ Housing attributes	+ Location attributes	Boundary design			
					Obs.<2.5km	Obs.<1.0km	Obs.<0.5km	
Share greenbelt land in postcode, $g_i$	0.0409*** (0.0086)	0.0230** (0.0091)	0.0045 (0.0075)	-0.0416*** (0.0075)	-0.0566*** (0.0124)	-0.0183 (0.0250)	0.0162 (0.0321)	0.0016 (0.0402)
Share greenbelt land in vicinity, $v_i$	0.1284*** (0.0048)	0.1200*** (0.0133)	0.0757*** (0.0107)	0.0779*** (0.0106)	0.0852*** (0.0180)	0.1004*** (0.0346)	0.0820* (0.0451)	0.0886* (0.0520)
Housing attributes (7)	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes (11)	No	No	No	Yes	Yes	Yes	Yes	Yes
Local authority×year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Output area×year fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Output area×year×boundary segment fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
Observations	13,593,673	13,593,673	13,593,673	13,593,673	11,204,496	2,535,248	1,566,877	1,047,097
$R^2$	0.0608	0.4349	0.5888	0.5900	0.6219	0.7167	0.7458	0.7694

*Notes:* We set  $\delta = 2.5$ . Housing attributes include housing type dummies (bungalow, house, maisonette), the floor level of the property, the height of the property, the number of stories of the building, and whether the property has a fire place. Location attributes are distance rings to roads and rivers and a linear and squared term of distance to the city centre. In column (5) we only include observations that are either further away than 2.5km of a greenbelt boundary or for which the boundary does not intersect with rivers or major roads. In column (6), (7) and (8) we only include observations within respectively 2.5, 1 and 0.5km of a non-intersecting greenbelt boundary. Standard errors are clustered at the postcode level and in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

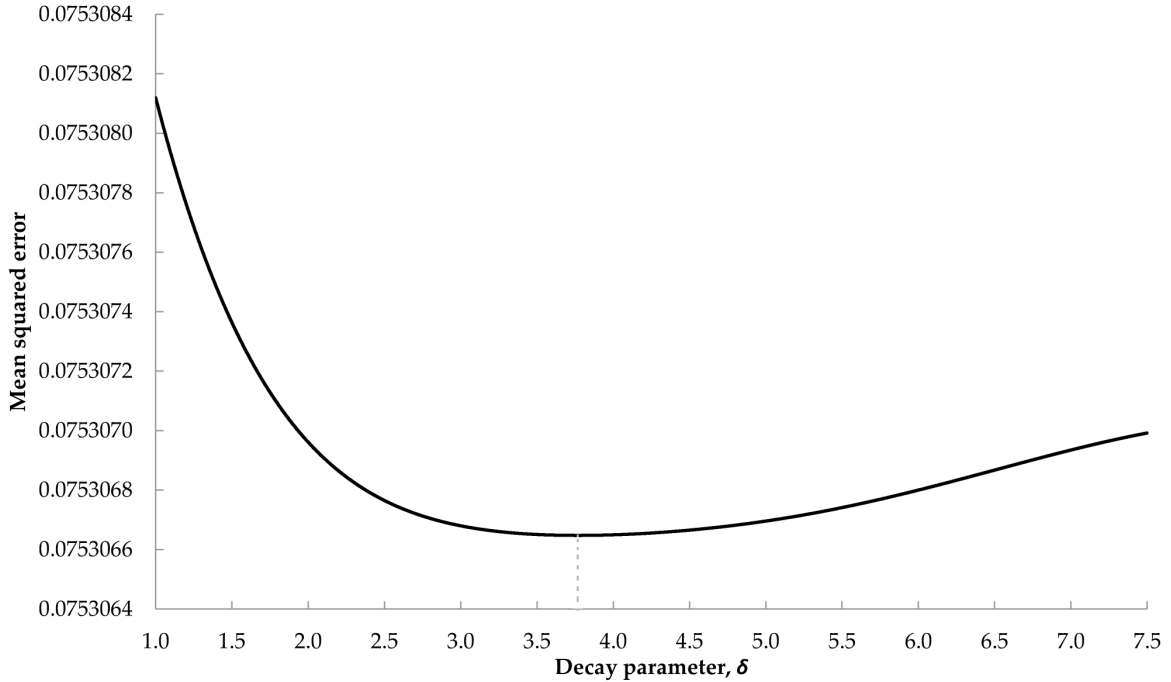


FIGURE B1 – MSE FOR DIFFERENT  $\delta$

from the city centre. Although prices have gone up (partly due to the external effect), adjusting size of the house is often difficult, so that houses are still larger when having more greenbelt land around them.

In Table B2 we explore further the local effects on house prices. First, one may criticise the arbitrary choice of the decay parameter  $\delta$  to determine the weighted share of greenbelt around each property. In column (1) we choose a more general approach by including 250m and 500m distance rings instead of specifying  $v_i$ . The results show that we only find statistically significant coefficients within 500m, suggesting that our choice of  $\delta$  makes sense. We explore the choice further in Figure B1 where we show that for our preferred specification (column (5), Table 3),  $\delta = 3.767$  minimises the Mean Squared Error, providing additional support for our choice of  $\delta$ .

In column (2), Table B2, we make a distinction between ‘accessible’ and ‘non-accessible’ greenbelt land. One may argue that mainly accessible greenbelt will yield amenity benefits for households living close to greenbelts. However, this appears not to be the case. The coefficients seem to be essentially the same, although it should be noted that the standard error regarding accessible greenbelt land is too large to make precise statements.

TABLE B2 – HETEROGENEITY IN THE RESULTS FOR PRICES  
*(Dependent variable: the log of house price per m<sup>2</sup>)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Distance rings	Accessible greenbelt land	Agricultural greenbelt land	Year effects	London	Manchester- Liverpool	West Midlands	Other greenbelts
Share greenbelt land in postcode	-0.0209*** (0.0074)	-0.0227*** (0.0070)	-0.0227*** (0.0070)	-0.0227*** (0.0069)	-0.0165 (0.0116)	-0.0212* (0.0113)	-0.0239 (0.0209)	-0.0202** (0.0100)
Share greenbelt land in vicinity				0.1868*** (0.0380)	0.0987** (0.0447)	0.2150*** (0.0483)	0.1795** (0.0799)	0.2132*** (0.0431)
Share accessible greenbelt land in vicinity		0.1517* (0.0792)						
Share inaccessible greenbelt land in vicinity		0.1609*** (0.0307)						
Share agricultural greenbelt land in vicinity			0.2092*** (0.0493)					
Share non-agricultural greenbelt land in vicinity			0.1260*** (0.0402)					
Share accessible greenbelt land in vicinity × (2017-year)				-0.0024 (0.0027)				
Share greenbelt land 0-250m	0.0382** (0.0156)							
Share greenbelt land 250-500m	0.0510*** (0.0192)							
Share greenbelt land 500-1000m	0.0467 (0.0331)							
Share greenbelt land 1000-1500m	0.0443 (0.0407)							
Share greenbelt land 1500-2000m	0.0322 (0.0483)							
Housing attributes (15)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes (11)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Output area×year×boundary segment fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,361,658	8,361,658	8,361,658	8,361,658	2,342,981	1,481,077	696,994	3,075,631
R <sup>2</sup>	0.9024	0.9024	0.9024	0.9024	0.8960	0.8838	0.8930	0.8962

*Notes:* We set  $\delta = 3.767$ . Housing attributes include housing type dummies (flat, terraced, semi-detached, detached), the number of rooms and the number of habitable rooms, an indicator for newly built properties, the floor level of the property, the height of the property, the number of stories of the building, whether the property has a fire place, whether the property is freehold and variables capturing the energy efficiency of windows, roof, walls. Location attributes are distance rings to roads and rivers and a linear and squared term of distance to the city centre. We only include observations that are either further away than 2.5km of a greenbelt boundary or for which the boundary does not intersect with rivers or major roads. Standard errors are clustered at the postcode level and in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

In column (3) we calculate the share of ‘agricultural’ and ‘non-agricultural’ greenbelt land. The coefficient for the former seems to be somewhat higher (although the difference is only marginally statistically significant). This may seem surprising as households are unlikely to value intensive livestock farming that generates negative externalities (Bontemps et al. 2008). However, it may be that agricultural land is usually ‘open’ land, which may be valued higher by households than *e.g.* forest land. In any case, we refrain from making strong statements given the relatively large standard errors.

In column (4) we investigate whether the external effect of greenbelts is constant over time by including an interaction term of the share greenbelt land in vicinity with a year indicator. As the latter term is small and highly statistically insignificant, it does not seem to be the case that external effects become more (or less) pronounced over time.

We also consider heterogeneity in the external effect between different greenbelts in columns (5)-(8). Because our results become imprecise once we have fewer observations, we only estimate the effects for greenbelts for which we have more than 500 thousand sales. The results suggest that the external effect in the Manchester-Liverpool (column (6)) and West Midland (column (7)) greenbelts are very similar to the baseline estimate. For London the external effect is about 50% lower (column (5)). We find small internal effects for the Manchester-Liverpool greenbelt, as well as for other greenbelts (column (8)).

## *B.2 Greenbelt visits: evidence from pictures*

In this subsection we investigate whether greenbelts are a main destination for natives and tourists. In order to proxy for the attractiveness of a postcode we use data on geocoded pictures from *FlickrR*, an online hosting service for media. Using geocoded pictures involves care. First, to avoid the possibility of inaccurate geocoding, we keep only one geocoded picture per location defined by its geographical coordinates per user per hour of the day. This reduces the number of pictures by about 45%. Second, one may argue that the patterns of pictures taken by tourists and residents may be very different. Since we have information on users’ identifiers, we can distinguish between residents’ and tourists’ pictures by keeping users who take pictures for at least 6 consecutive months between 2000 and 2018 in England. It seems unlikely that tourists



stay for 6 consecutive months in the area. Note that the correlation between natives' and tourists' pictures is equal to 0.748. Third, many recorded pictures may not be related to recreational visits but to ordinary events in daily life occurring inside the house. Hence, in the regressions we control flexibly for the number of dwellings in a postcode, by including a 3<sup>rd</sup>-order polynomial of dwellings we only keep pictures that are taken outside buildings. We then estimate:

$$q_{ij} = \exp(\zeta_1 g_i + \zeta_2 m_i + f(n_{ij}) + \psi_{jb} + \epsilon_{ij}) \quad \text{if } d_{ib} < \bar{d}, \quad (13)$$

where  $q_i$  is the number of pictures in a postcode between 2000 and 2018,  $g_i$  is the share of the postcode in the greenbelt,  $m_i$  are neighbourhood characteristics,  $f(\cdot)$  is a 3<sup>rd</sup>-order polynomial of dwellings of the number of dwellings in a postcode, and  $\psi_{jb}$  are OA×boundary segment fixed effects.<sup>19</sup> We estimate equation (13) by a Poisson model. Table B3 reports the results.

In Panel A we first concentrate on the effects of greenbelts on tourist pictures. In the first specification we do not control for location attributes or detailed fixed effects. We find that pictures are 56% lower in greenbelts. We emphasise that this effect is essentially unaffected when we do not control for the number of dwellings in the postcode. Hence, it seems that the lack of pictures in greenbelts cannot be explained by differences in density. The most likely explanation therefore seems that greenbelts are just not very popular places to go. When we include OA fixed effect in column (2), Panel A, we find that there are 41% fewer tourist pictures in greenbelts. This effect is virtually unchanged when we add location attributes (column (3)), add OA×boundary segment fixed effects (column (4)) and focus on areas within 2.5km of the greenbelt boundary (column (5)). However, when we focus on areas within 1km or even 500m (columns (6) and (7)), the effect becomes statistically insignificant. In other words, greenbelts do not seem to be main destinations for visits of tourists.

We replicate those results in Panel B of Table B3, but replace the dependent variable by the number of pictures made by natives. The results are comparable: we find that the number of pictures is lower in greenbelt areas, but if we focus on postcodes within 1km of the greenbelt the effects become statistically insignificant.

---

<sup>19</sup>Note that, next to the fixed effects, we also control for distance to the nearest city centre, which should absorb any density gradient.

TABLE B3 – VISITS TO GREENBELTS: PICTURES  
(Dependent variable: the number of pictures in a postcode)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Panel A: Tourist pictures</i>	LA fixed effects	+ OA fixed effects	+ Location attributes		Boundary design		
					Obs.<2.5km	Obs.<1.0km	Obs.<0.5km
Share of postcode in greenbelt, $g_i$	-0.8301*** (0.1837)	-0.5337*** (0.1109)	-0.5460*** (0.1079)	-0.5284*** (0.1023)	-0.5841*** (0.1568)	-0.2847 (0.1915)	-0.1501 (0.2513)
Area size of postcode ( $\log$ )	0.6965*** (0.0603)	0.9039*** (0.0227)	0.8976*** (0.0186)	0.9100*** (0.0220)	0.8965*** (0.0302)	0.8873*** (0.0411)	0.8812*** (0.0733)
Dwellings in postcode $f(\cdot)$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes (8)	No	No	Yes	Yes	Yes	Yes	Yes
Local authority fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Output area fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes
Output area $\times$ boundary segment fixed effects	No	No	No	Yes	Yes	Yes	Yes
Observations	1,310,750	1,310,750	1,310,750	1,094,110	228,940	142,203	96,646
Log-likelihood	-6.175e+06	-2.195e+06	-2.181e+06	-1.604e+06	-106078	-49033	-23543
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Panel B: Pictures by natives</i>	LA fixed effects	+ OA fixed effects	+ Location attributes		Boundary design		
					Obs.<2.5km	Obs.<1.0km	Obs.<0.5km
Share of postcode in greenbelt, $g_i$	-0.4748*** (0.1016)	-0.3529*** (0.0829)	-0.3767*** (0.0817)	-0.3869*** (0.0795)	-0.4888*** (0.1551)	-0.3161 (0.2423)	0.0495 (0.2221)
Area size of postcode ( $\log$ )	0.6051*** (0.0251)	0.8442*** (0.0090)	0.8399*** (0.0086)	0.8382*** (0.0100)	0.9100*** (0.0208)	0.9057*** (0.0278)	0.8801*** (0.0369)
Dwellings in postcode $f(\cdot)$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location attributes (8)	No	No	Yes	Yes	Yes	Yes	Yes
Local authority fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Output area fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes
Output area $\times$ boundary segment fixed effects	No	No	No	Yes	Yes	Yes	Yes
Observations	1,310,750	1,310,750	1,310,750	1,094,110	228,940	142,203	96,646
Log-likelihood	-1.300e+07	-4.780e+06	-4.753e+06	-3.353e+06	-246936	-102491	-47061

Notes: These are Poisson models, so the coefficients are interpreted as (semi-)elasticities. In column (4) we only include observations that are either further away than 1.5km of a greenbelt boundary or for which the boundary does not intersect with rivers or major roads. In column (5) we only include observations within 1.5km of a non-intersecting greenbelt boundary. Standard errors are clustered at the postcode level and in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

If greenbelts are not destinations for visits, does this question the positive price effect we found earlier? We do not think so: given that the positive amenity effect of greenbelts is very local, it most likely captures the local view effect of green space.

### *B.3 Commuting and CO<sub>2</sub> emissions*

In this subsection we inspect the potential other aggregate effects of greenbelt policy. The first effect may be that urban development may ‘leapfrog’ over the greenbelts leading to an inefficient urban structure and longer commutes (see [Levkovich et al. 2017](#), [Cheshire et al. 2018](#)). If that is the case, one would expect aggregate housing prices to be lower in areas with more greenbelt land, suggesting that we may underestimate the aggregate price effects of greenbelts. We gather additional data from the 2011 census on commuting distances of households and calculate the average commuting distance in each output area. We might then expect that commutes are longer once they are surrounded by more greenbelt land, so that people have to commute to jobs further afield. In [Table B4](#) we do not find any evidence for this.

In column (1) we initially find a negative effect: commuting distances are lower in areas with more greenbelt land. However, when we control for the share of counterfactual greenbelt land and population density in 1951 in column (2) we do not find any meaningful effect of the share of greenbelt land within 15km on commuting distances. This also holds if we add additional controls and government region×year fixed effects in column (3), and output area characteristics in column (4). In column (5) we only include OAs which are outside greenbelts and further than 5km from a greenbelt boundary, confirming the absence of an effect of greenbelts on commuting distances.

Another concern could be that due an inefficient urban structure per capita energy consumption may increase, which leads to more CO<sub>2</sub> emissions. More specifically, due to lower densities, households may emit more CO<sub>2</sub> emissions ([Glaeser & Kahn 2004](#), see). We test this by exploiting data from EPCs on the predicted CO<sub>2</sub> emissions per property and report the results in [Table B5](#).

In column (1) we find no effect of greenbelt land in the surroundings on housing emissions. When we control for the share of counterfactual greenbelt land and population density in 1951

TABLE B4 – AGGREGATE EFFECTS: COMMUTING  
*(Dependent variable: the log of commuting distance (in km))*

	(1)	(2)	(3)	(4)	(5)
		+ $f(\cdot)$ , $g(\cdot)$	+ Location attributes, f.e.	+ Output area, characteristics,	Outside greenbelts
Share greenbelt land, < 15km	-0.2280*** (0.0549)	-0.0430 (0.0527)	0.0481 (0.0473)	0.0147 (0.0499)	-0.0054 (0.0561)
Share counterfactual greenbelt land, < 15km	0.0119 (0.1321)				
Population density, < 15km ( <i>log</i> )			-0.1477*** (0.0255)	-0.1380*** (0.0238)	-0.1494*** (0.0254)
Standard deviation of altitude, < 15km ( <i>log</i> )			-0.0035 (0.0156)	-0.0065 (0.0150)	-0.0098 (0.0168)
Share AONB, < 15km			0.0515 (0.0467)	0.0471 (0.0452)	0.0649 (0.0514)
Share national park, < 15km			0.0503 (0.0817)	-0.0419 (0.0796)	-0.0250 (0.0864)
Share developed land, < 15km			-0.1512 (0.1753)	-0.1415 (0.1142)	-0.0163 (0.1148)
Share water bodies, < 15km			0.4453 (0.5514)	1.2118** (0.5974)	0.1521 (0.6733)
Share counterfactual greenbelt land, $f(\cdot)$	No	Yes	Yes	Yes	Yes
Population density in 1951 < 15km, $g(\cdot)$	No	Yes	Yes	Yes	Yes
Output area characteristics	No	No	Yes	Yes	Yes
Government region×year fixed effects	No	No	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	3,265,303	3,265,303	3,265,303	3,265,303	1,560,602
$R^2$	0.0223	0.2169	0.2479	0.3649	0.4130

*Notes:* In column (5) we only include OAs which are outside greenbelts and further than 5km from a greenbelt boundary. Standard errors are clustered at the travel-to-work-area level and in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

TABLE B5 – AGGREGATE EFFECTS: CO<sub>2</sub> EMISSIONS  
*(Dependent variable: Output area×year fixed effect,  $\hat{\rho}_{jt}$ )*

	(1)	(2)	(3)	(4)	(5)
		+ $f(\cdot)$ , $g(\cdot)$	+ Location attributes, f.e.	+ Output area, characteristics,	Outside greenbelts
Share greenbelt land, < 15km	-0.0219 (0.0232)	0.0419 (0.0273)	0.0538** (0.0220)	0.0035 (0.0198)	-0.0312 (0.0246)
Share counterfactual greenbelt land < 15km	0.0120 (0.0631)				
Population density, < 15km ( <i>log</i> )			-0.1477*** (0.0170)	-0.1179*** (0.0126)	-0.1214*** (0.0112)
Standard deviation of altitude, < 15km ( <i>log</i> )			-0.0008 (0.0081)	-0.0038 (0.0099)	-0.0149* (0.0076)
Share AONB, < 15km			0.0699** (0.0278)	0.0833*** (0.0249)	0.0985*** (0.0275)
Share national park, < 15km			0.0961** (0.0475)	0.0375 (0.0412)	0.0442 (0.0430)
Share developed land, < 15km			-0.0225 (0.0637)	-0.0578 (0.0482)	-0.0906*** (0.0290)
Share water bodies, < 15km			0.0588 (0.3983)	0.6614** (0.3340)	0.2268 (0.2732)
Share counterfactual greenbelt land, $f(\cdot)$	No	Yes	Yes	Yes	Yes
Population density in 1951 < 15km, $g(\cdot)$	No	Yes	Yes	Yes	Yes
Output area characteristics	No	No	Yes	Yes	Yes
Government region×year fixed effects	No	No	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	1,469,826	1,469,826	1,469,826	1,469,826	709,422
$R^2$	0.0095	0.0420	0.0572	0.2872	0.2911

*Notes:* In column (5) we only include OAs which are outside greenbelts and further than 5km from a greenbelt boundary. Standard errors are clustered at the travel-to-work-area level and in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

TABLE B6 – AGGREGATE EFFECTS: SENSITIVITY FOR PRICES  
*(Dependent variable: Output area×year fixed effect,  $\hat{\rho}_{jt}$ )*

	(1)	(2)	(3)	(4)	(5)	(6)
	+ Refusal rate	$\hat{\rho}_{jt}$ from column (1), Table 3	Counterfactual greenbelts 10km	Polynomials for $f(\cdot)$ , $g(\cdot)$	Share greenbelt land < 10km	Share greenbelt land < 50km
Share greenbelt land, < 15km	0.1116** (0.0471)	0.1336** (0.0640)	0.1369** (0.0592)	0.1732*** (0.0617)		
Share greenbelt land, < 10km					0.1382*** (0.0423)	
Share greenbelt land, < 50km						0.6468*** (0.2112)
Minor refusal rate 1979-2017	0.8735*** (0.2866)					
Major refusal rate 1979-2017	0.2852** (0.1292)					
Share counterfactual greenbelt land, $f(\cdot)$	Yes	Yes	Yes	Yes	Yes	Yes
Population density in 1951, $g(\cdot)$	Yes	Yes	Yes	Yes	Yes	Yes
Output area characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Government region×year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,391,289	1,391,308	1,391,289	1,391,289	1,391,289	1,391,289
$R^2$	0.8119	0.9205	0.8064	0.8052	0.8077	0.8107

*Notes:* The dependent variable is obtained from the specification listed in column (5) in Table 3 and in column (2), Table 3, for the specification listed in column (1). We only include OAs which are outside greenbelts and further than 5km from a greenbelt boundary. Standard errors are clustered at the travel-to-work-area level and in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

the effect becomes positive and just not statistically significant at conventional levels. When we include additional location attributes and government region×year fixed effects, we find a positive effect: the coefficient implies that a 10% increase in the share of greenbelt land within 15km increases CO<sub>2</sub> emissions by 0.5%. Given the average share of greenbelt land of 0.27, this means that without greenbelts, emissions would be on average 1.5% lower, so the effect is not large. Once we control for output area characteristics (column (4)) and focus on areas outside greenbelts (column (5)), the effect is statistically insignificant.

In other words, aggregate effects such as commuting and housing emissions do not seem to be related to the amount of greenbelt land in vicinity. This makes our welfare analysis, which is based on differences in house prices, more credible.

#### B.4 Sensitivity of aggregate effects

Here we report sensitivity with respect to the analysis of aggregate effects. Table B6 reports the results on prices and Table B7 the results with respect to house size. We consider the specification in column (5) of respectively Tables 5 and Table 6 as the preferred specification.

In column (1) of Table B6 we aim to see whether our results are sensitivity to the inclusion of alternative land use regulations. More specifically, Hilber & Vermeulen (2016) show that land use regulation leads to higher house prices using the refusal rate of ‘major’ construction projects as a proxy for regulatory restrictiveness. This measure comes from the DCLGs Planning Statistics. The refusal rate for ‘major’ projects is defined as the share of applications for residential developments of ten or more dwellings that is refused by an LA in any year during the process of ‘development control’, while ‘minor’ projects refer to construction projects of fewer than 10 dwellings. The latter may also refer to renovation projects of existing homes. We calculate the ‘minor’ and ‘major’ refusal rate for each LA using data on all applications and refused applications of developments between 1979 and 2017. We indeed find that both the refusal rate of minor projects as the refusal rate of major construction projects lead to higher prices increase prices. A standard deviation increase in the minor refusal rate is associated with a  $0.296 \times 0.8735 = 26\%$  increase in price, while an increase of one standard deviation in the major refusal rate is associated with a  $0.248 \times 0.285 = 7\%$  increase in price. We should be careful not to interpret this as a causal estimate as there are several endogeneity concerns here (see Hilber & Vermeulen 2016, for more details). What is more important, the estimate of greenbelt land in the vicinity is hardly affected.

In column (2) we test whether the choice of residuals affects the results. In the baseline specifications we choose the residuals obtained from column (5) in Table 3. Instead, here we use the residuals obtained from column (1) in Table 3. This does not affect the results.

In the definition of counterfactual greenbelts we use buffers around large cities of 15km. In column (3), we investigate whether results change if we use smaller buffers. This does not seem to be the case. Column (4) examines the sensitivity of the results once we use polynomials to approximate  $f(\cdot)$  and  $g(\cdot)$  instead of cubic smoothing splines. The results are hardly affected.

In the last two columns of Table 3 we test whether our results are robust to another definition of the housing market. In our preferred specifications, we calculate the share of greenbelt land in the local housing market, which we define as being within 15km of the output area. This is based on the average commuting distance of 15.18km. Column (5), however, shows that the results are very similar once we use the share of greenbelt land within 10km. When we use the

TABLE B7 – AGGREGATE EFFECTS: SENSITIVITY FOR HOUSE SIZE  
*(Dependent variable: Output area×year fixed effect,  $\hat{\sigma}_{jt}$ )*

	(1)	(2)	(3)	(4)	(5)	(6)
	+ Refusal rate	$\hat{\rho}_{jt}$ from column (1), Table 3	Counterfactual greenbelts 10km	Polynomials for $f(\cdot)$ , $g(\cdot)$	Share greenbelt land < 10km	Share greenbelt land < 50km
Share greenbelt land, < 15km	-0.1553*** (0.0211)	-0.0588*** (0.0172)	-0.1679*** (0.0211)	0.1732*** (0.0617)		
Share greenbelt land, < 10km					-0.1328*** (0.0219)	
Share greenbelt land, < 50km						-0.3560*** (0.1365)
Minor refusal rate 1979-2017	-0.5364** (0.2140)					
Major refusal rate 1979-2017	0.1399 (0.1145)					
Share counterfactual greenbelt land, $f(\cdot)$	Yes	Yes	Yes	Yes	Yes	Yes
Population density in 1951, $g(\cdot)$	Yes	Yes	Yes	Yes	Yes	Yes
Output area characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Government region×year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	710,352	710,363	710,352	1,391,289	710,352	710,352
$R^2$	0.3108	0.7056	0.3036	0.8052	0.2962	0.3295

*Notes:* The dependent variable is obtained from the specification listed in column (5) in Table 3 and in column (2), Table 3, for the specification listed in column (1). We only include OAs which are outside greenbelts and further than 5km from a greenbelt boundary. Standard errors are clustered at the travel-to-work-area level and in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

share of greenbelt land within 50km, the price effect becomes much stronger. A 10% increase in the share of greenbelt land within 50km increases prices by 6.5%. Hence, if anything, our estimates of the aggregate price effects are conservative.

In Table B7 we replicate those results but use the residual with respect to house size as dependent variable. It can be quickly seen that the results are very robust. Column (1) shows that the minor refusal rate is associated with smaller houses. This makes sense as the minor refusal rate is associated with higher prices, which should make people to live in smaller houses. Column (2) shows that with the residual obtained from the specification in column (1), Table B1 instead of column (5), the effect of greenbelt land on house size becomes slightly weaker, but the effect is still meaningful and statistically significant. Again, in column (6) of Table B7, we notice that when using the share of greenbelt land within 50km, the effect becomes stronger. The latter also implies that the elasticity of price with respect to size remains similar.