

## Appendix 1 - Data Processing

Having downloaded the UK Price Paid data from HM Land Registry, I must geocode the addresses in order to relate them to the ULEZ. I do this in ArcGIS, using the OS Open Names Locator API from the Ordnance Survey and ESRI<sup>1</sup>. Once this is done, I need feature layers giving the borders of the ULEZ in 2019, 2021 and 2023, to create my distance variables. The London Datastore provides shapefiles for the ULEZ in 2019, 2021 and 2023, from which I extract the borders in ArcGIS<sup>2</sup>. Each of these shapefiles use both a geographical coordinate system (GCS) and a projected coordinate system (PCS), and in each case the projection of the PCS is centred in Britain. In particular, their GCS is some variant of OSGB, and their PCS is some variant of British National Grid, which ensures consistency and accuracy.

I then perform a series of Nears and Spatial Joins in ArcGIS to calculate the Euclidean (crow-flies) distance from each house to each border, as well as which of these zones each house is in. Because of the spatial referencing of the shapefiles from the London Datastore, particularly the fact that they all use a PCS centered in Britain, I calculate planar distances, rather than geodesic ones, as they will be more accurate.

The one small issue with this process is that the Open Names Locator API locates houses at the postcode level. This means all houses with the same postcode has the same spatial reference, which means they all share the same calculated distances. However, this is a very minor issue, since postcodes are small in area and so any errors in distance calculations will be very small relative to the actual calculated distances.

Everything described above is coded in Python and run through the Python window in ArcGIS.

---

<sup>1</sup>For more information, see [here](#)

<sup>2</sup>The 2019 ULEZ shapefile is available from [here](#), the 2021 shapefile [here](#), and the 2023 shapefile [here](#)

## Appendix 2 - Heterogeneity and Robustness

	(1)		(2)
ULEZ*detached	-0.0195 (0.0153)	ULEZ*quartile1	-0.0575*** (0.00676)
ULEZ*semi-detached	0.0344*** (0.00652)	ULEZ*quartile2	-0.00990** (0.00403)
ULEZ*terraced	0.0245*** (0.00537)	ULEZ*quartile3	0.0101** (0.00403)
ULEZ*flat	-0.0918*** (0.00985)	ULEZ*quartile4	0.0208*** (0.00581)
Observations	481,579	Observations	481,579
R-squared	0.694	R-squared	0.845
Treated Group?	2021 & 2023	Treated Group?	2021 & 2023
Postcode Sector FEs?	✓	Postcode Sector FEs?	✓
Quarter FEs?	✓	Quarter FEs?	✓
Controls?	✓	Controls?	✓

Robust standard errors, clustered at the postcode-sector level, in parentheses

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table 1: Results from the static specification, across different house types and price quartiles

In both of these regressions, I take the static specification and replace the  $ULEZ_{pt}$  dummy with a set of its interactions with each house type in (1), and each price quartile in (2). I also include controls for the price quartile in (2) - in (1) the controls for house type are already there. This regression has a valid difference-in-differences interpretation as long as the average price differences between quarters and postcode sector, conditional on house type/price quartile, are constant - this seems reasonable and I assume it to be true. Also, I calculate the price quartile in (2) within quarters - i.e. I compare each house to only houses sold in the same quarter. Quartile 1 represents the lowest quartile, and 4 the highest.

	(1)	(2)	(3)	(4)
ULEZ	-0.0323*** (0.0054)	-0.0353*** (0.0065)	-0.0392*** (0.0048)	-0.0195*** (0.0071)
Treated Group?	2021 & 2023	2021 & 2023	2021 & 2023	2021 & 2023
Parallel Trends Distance	12.5km	7.5km	10km	10km
Spillover Distance	5km	5km	2.5km	7.5km
Postcode Sector FEs?	✓	✓	✓	✓
Quarter FEs?	✓	✓	✓	✓
Controls?	✓	✓	✓	✓
Observations	544,515	424,485	769,700	239,620
R-squared	0.700	0.690	0.687	0.698

Robust standard errors, clustered at the postcode-sector level, in parentheses

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table 2: Results from the static specification, across different assumed cutoff distances

I am unable to vary the distances much more than this - extending the spillover distance far past 5km gives very few observations in the 2021 zone, and contracting the parallel trends distance much shorter than 10km gives very few observations in the control zone. Unfortunately, these are the directions of distance changes I would be most interested in examining. However, careful analysis of the parallel trends assumption and spillover effects using Rambachan and Roth (2023) and Butts (2024) respectively should convince the reader that my results are robust to these assumptions.

## Appendix 3 - Event Study Regression Table

Event time ( $j$ )	(1)	(2)	(3)	(4)	(5)
$j = -5$	-0.0015 (0.0099)	-0.0091 (0.0105)	0.0017 (0.0038)	-0.0052 (0.0077)	0.0073** (0.0035)
$j = -4$	-0.0074 (0.0079)	-0.0131 (0.0087)	-0.0038 (0.0032)	-0.0133** (0.0067)	0.0009 (0.0034)
$j = -3$	-0.0042 (0.0079)	-0.0084 (0.0086)	-0.0002 (0.0032)	-0.0082 (0.0065)	0.0057* (0.0032)
$j = -2$	- (-)	- (-)	- (-)	- (-)	- (-)
$j = -1$	-0.0275*** (0.0091)	-0.0242** (0.0094)	-0.0164** (0.0082)	-0.0166** (0.0075)	0.0065 (0.0040)
$j = 0$	-0.0283*** (0.0099)	-0.0259** (0.0103)	-0.0200** (0.0081)	-0.0290*** (0.0067)	-0.0137** (0.0066)
$j = 1$	-0.0471*** (0.0103)	-0.0420** (0.0102)	-0.0314*** (0.0078)	-0.0388*** (0.0067)	-0.0088 (0.0067)
$j = 2$	-0.0325*** (0.0110)	-0.0330*** (0.0103)	-0.0226*** (0.0078)	-0.0319*** (0.0073)	-0.0223*** (0.0072)
$j = 3$	-0.0335*** (0.0091)	-0.0311*** (0.0092)	-0.0221*** (0.0067)	-0.0194*** (0.0069)	-0.0102 (0.0067)
$j = 4$	-0.0436*** (0.0109)	-0.0342*** (0.0096)	-0.0293*** (0.0080)	-0.0277*** (0.0075)	-0.0098 (0.0066)
$j = 5$	-0.0534*** (0.0134)	-0.0364*** (0.0114)	-0.0387*** (0.0118)	-0.0492*** (0.0102)	-0.0238*** (0.0067)
Observations	481,579	481,579	481,592	978,367	978,367
Estimand	Treatment Effect	Treatment Effect	Treatment Effect	Treatment Effect	Spillover Effect
Estimator	TWFE OLS	Abraham and Sun (2021)	Gardner (2021)	Butts (2024)	Butts (2024)
Controls?	✓	✓	✓	✓	✓

Robust standard errors, clustered at the postcode-sector level, in parentheses

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table 3: Results from the event study regressions, for selected event times

## Appendix 4 - Alternative Visualisation of Treatment Effects

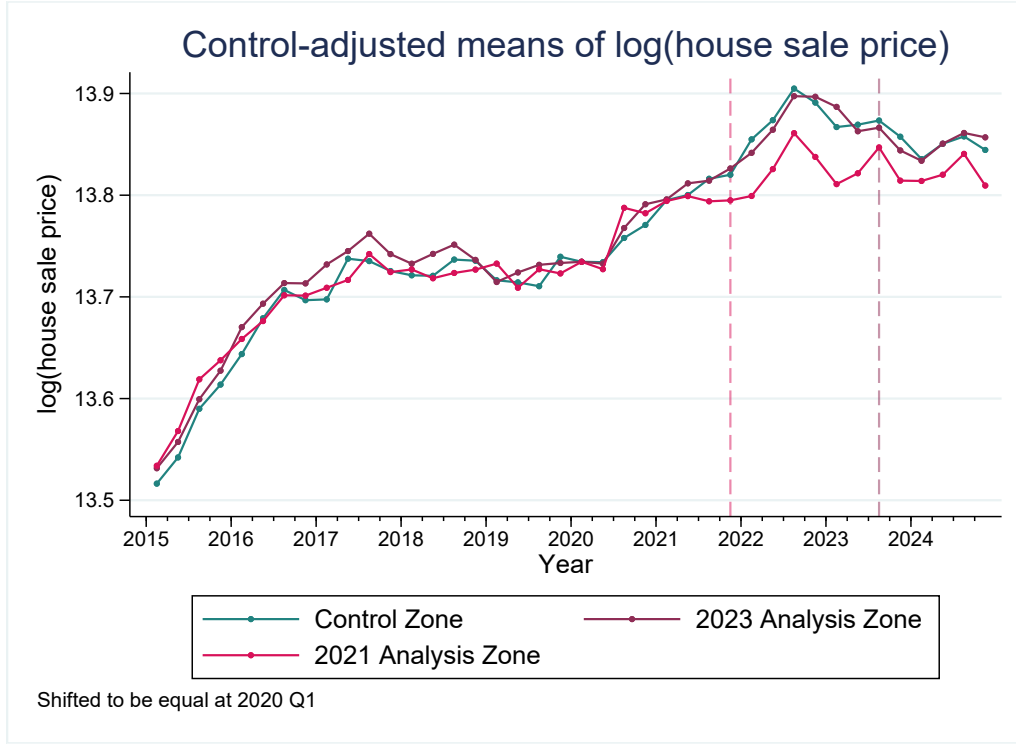


Figure 1: Evolution of mean log(house price) over time, in each region

In the above, I plot the evolution of the quarterly mean of the natural log of house prices over time in each zone, once adjusted for fixed-effects and controls. Specifically, I run the following regression:

$$Y_{ipt} = \delta_t \cdot zone_p + \alpha_p + \rho \mathbf{X}_{ipt} + \varepsilon_{ipt}$$

where  $zone_p$  is just an indicator of the zone  $p$  is in. I then take the fitted values for each quarter-zone combination, setting everything else to 0 (and therefore taking the baseline post-code sectors in each region). Importantly, the mean price values are arbitrary, because I shift them all to be equal at 2020 Q1, but their *differences* are meaningful. The same key features are observed as in Figure 4. Before the first treatment, all paths evolve very similarly, supporting the parallel trends assumption. One period before treatment, in both the 2021 and 2023 zones, the mean price drops relative to the untreated and not-yet treated regions, and in the 2021 case this is sustained until the end of the period. Finally, in the 2021 case, there is a recovery in the difference over time. It is also clear from this graph how much more the 2021 zone contributes to the overall effect than the 2023 zone.

## Appendix 5 - Individual Dynamic DiD Results

The following figures give the estimated dynamic treatment effects from the individual dynamic DiDs of the 2021 zone vs the control zone, and the 2023 zone vs the control zone, respectively.

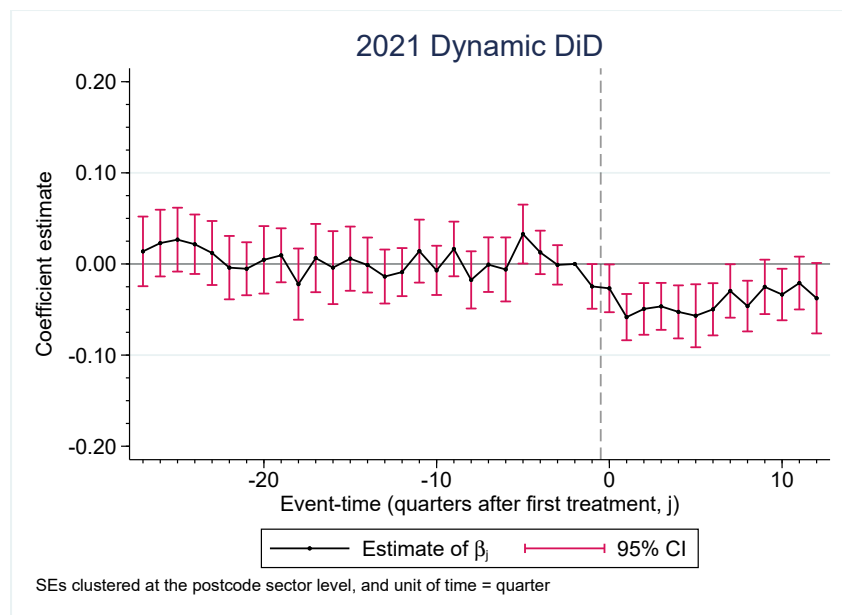


Figure 2: Coefficient plot from the dynamic DiD for the 2021 analysis zone individually

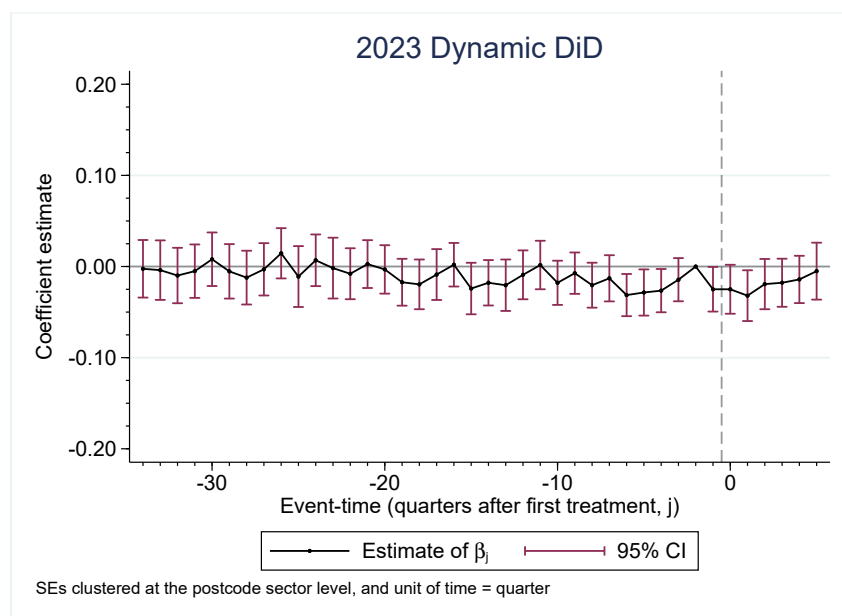


Figure 3: Coefficient plot from the dynamic DiD for the 2023 analysis zone individually

## Appendix 6 - Parallel Trends Analysis

I here expand on the methodology used in the analysis to test the robustness of my event study results to violations of the parallel trends assumption, established in Rambachan and Roth (2023). In showing that they are robust to these violations, I am able to say that while I cannot test the parallel trends assumption, if I am happy to assume that it is violated by no more than a certain amount, my results remain significant. I wanted to include this in the main analysis, but space constraints prevented me from doing so.

The intuitive idea here is that rather than assuming parallel trends holds, I impose bounds on how badly it can fail, thus making more post-treatment trends feasible under the null of no treatment effect. I then seek to show that under reasonable bounds, what I observe in my event studies above is still not feasible under this regime - i.e. the result is significant. In more detail, I replace the parallel trends assumption with the following: any deviation from parallel trends between consecutive periods in the post-period can be no larger than a multiple  $M$  of the maximum such deviation in the pre-period. More formally, using the decomposition from the paper of  $\beta_j$  into the sum of a treatment effect  $\tau_j$  and the difference in trends that would have occurred without treatment,  $\delta_j$ , this assumption states that for all  $t \geq 0$ ,  $|\delta_{t+1} - \delta_t| \leq M \cdot \max_{s < 0} |\delta_{s+1} - \delta_s|$ . Importantly,  $M$  is left up to the researcher, but the paper suggests that if there are an equal number of pre and post periods, then  $M = 1$  is a reasonable and intuitive choice - post-treatment deviations in parallel trends between consecutive periods are bounded by exactly the greatest such deviation before treatment.

I also act on two further important remarks in the paper (on p. 2563, footnote 7). Firstly, to mitigate against seemingly large pre-treatment deviations in parallel trends arising simply from stochastic ‘bouncing’ around a zero trend, I pool time periods together, from the quarter level to the year level. Second, the paper also suggests that pre-treatment periods closer to the treatment point will be more informative than those further back, and so I should allow  $M$  to be higher for these earlier periods. There is no obvious way of doing this within the `HonestDiD` package the authors provide. As an alternative (somewhat crude) solution, and guided by the advised condition when choosing  $M = 1$ , I restrict in my analysis the pre-period back only as far as the post-period extends, so that the number of pre- and post-treatment coefficients are equal.

The below figure gives the confidence intervals for each of the post-treatment coefficient estimates from each of the four estimators above, under this new assumption with  $M = 1$ . It should be first noted that as a result of the method, the confidence intervals get wider for periods further past the treatment point. This is intuitively because the permitted deviation from parallel

trends is allowed to compound over time - by allowing multiple consecutive 'large' deviations from parallel trends, I am implicitly enlarging the permitted failures in periods further on from treatment. However, the key point to make is that in each of the three robust models, the upper bound of the 95% confidence interval for the average treatment effect two years after treatment is below zero, indicating that my results are robust to large, repeated failures of parallel trends. In two of these models, the estimate of the treatment effect a year after treatment is also significantly negative at the 5% level under this more flexible specification. As such, while there is good evidence from the event study plots to argue in the conventional way that parallel trends holds in my analysis, I have shown that it holds 'as much as it needs to' in order for me to claim that the treatment effect is significantly negative. It is also important to note that this analysis does not affect my point estimates of the effects - it just provides further evidence to support their robustness.

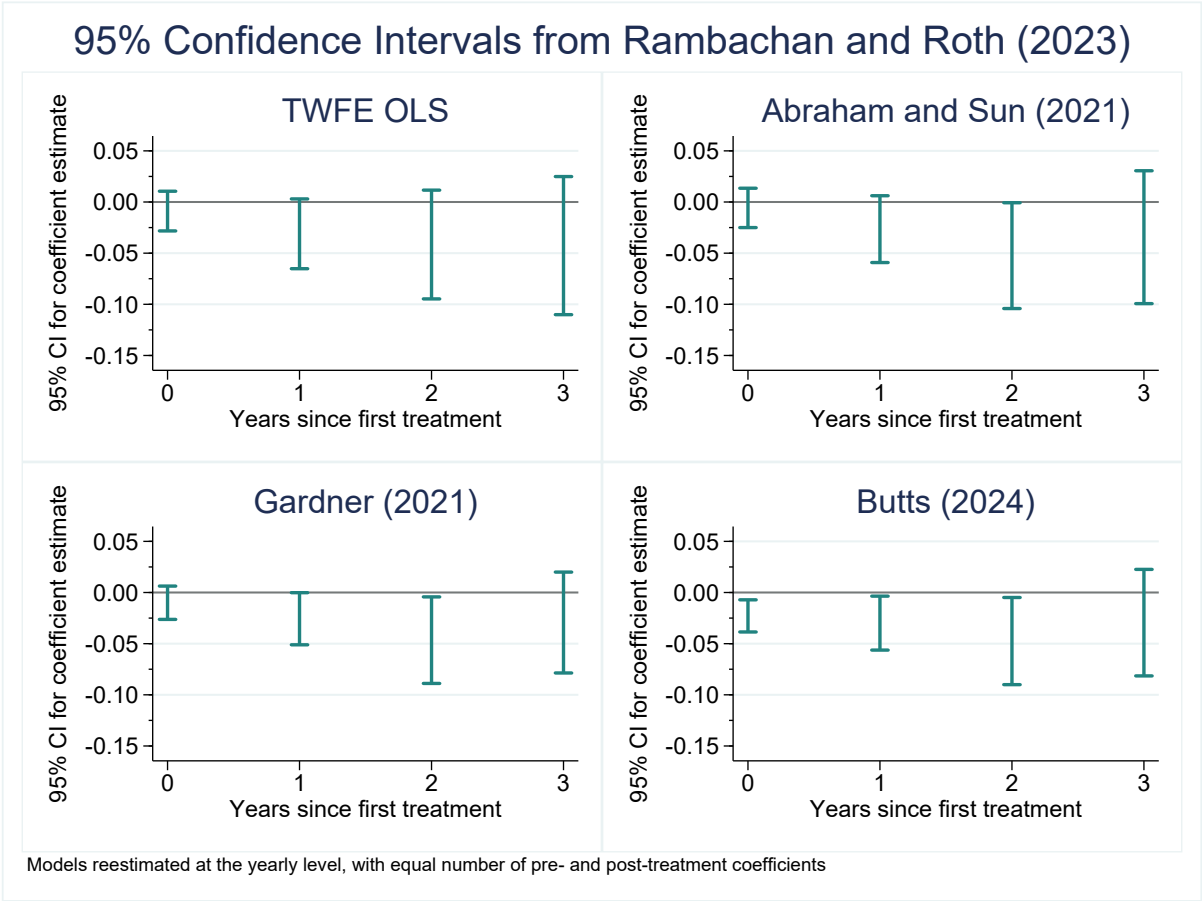


Figure 4: 95% confidence intervals for the treatment effect, from Rambachan and Roth (2023)

The fact that  $M$  is the choice of the researcher also allows me to report the largest  $M$ , denoted



$\bar{M}$ , at which a given coefficient estimate retains its significance. This sensitivity analysis is a key feature of the method that the paper provides. The below table reports these maximal  $\bar{M}$  values, for coefficients corresponding to one and two years after treatment, i.e. event times  $j = 1$  and  $j = 2$ , which are the time periods in which the most pronounced negative effect is observed. It indicates that even if I was to allow these post-treatment consecutive parallel trends deviations to be 15% larger than the maximum in the pre-period, and occur repeatedly over the post-period, I would still find that my results are significant at the 5% level. It should be emphasised how flexible this weakening of the parallel trends assumption is - this robustness analysis indicates that my initial results are strongly robust to possible violations in the parallel trends assumption.

<b>Estimator</b>	$\bar{M}$	
	$j = 1$	$j = 2$
TWFE OLS	0.85	0.70
Abraham and Sun (2021)	0.75	1
Gardner (2021)	1	1.15
Butts (2024)	1.15	1.15

$j$  := number of *years* since ULEZ first implemented in the region

Table 4: Maximal M ( $\bar{M}$ ) from Rambachan and Roth (2023) that retains significance of the relevant coefficient (nearest 0.05, rounded down)

## Appendix 7 - Mechanism Discussion

I briefly expand on a qualitative dynamic framework through which the observed treatment effect may be understood - in particular, why there is such a negative effect, why it persists for a while, and why it then gradually recovers. The ULEZ can be thought of as having both positive effects (improved air quality, reduced congestion, etc) and negative effects (notably the charge, as well as possible relocation of firms), somewhat separately. The hedonic theory of house prices pioneered in Rosen (1974) suggests house prices are derived from not just the properties of the house but also properties of its location. These positive and negative effects, all properties of the region, will respectively have positive and negative effects on house prices as a result. Crucially though, they only affect the price of a house to the extent that they have been understood by the participants in the housing market. As such, the magnitude of the effect of ULEZ on house prices depends on the relative size of these effects as understood and incorporated into the hedonic house valuation process. The dynamics of the effect thus depend on both the relative size of the positive and negative effects, and how quickly the population has been able to transmit and understand information about their size.

Information about the negative effects, most importantly the charge, is very quickly transmitted across the population. As soon as an individual first makes a journey into the zone in a non-compliant vehicle, they are aware of the charge via a payment notice from Transport for London - a sharp and noticeable change from the status quo before the implementation. However, the positive effects are far slower to emerge, and information about them is slower to spread. The change itself is very gradual, particularly air quality and its associated health costs, as people slowly adjust their behaviour to comply with the charge. Air pollution and congestion are also hard to measure, and so systematic changes in their levels over time are generally revealed through academic analysis, which is far harder to understand and communicate among the population.

Clearly, if the negative effects are faster to emerge, and are then quickly understood and incorporated into house valuations, then in the periods soon after implementation, house prices will see a larger fall from the well-understood negative effects than any countering rise from the less well-understood benefits. This explains the initial drop in house prices. However, beyond a point in time the charge is completely realised and understood by everyone, leaving no more room for it to have any further effect on house prices. Since it takes longer for the benefits to emerge and then be understood and factored into house valuations, the upward force they have on house prices takes longer to materialise. Therefore, in the period where negative effects have been fully understood and the positive effects are still being processed, house prices begin

to recover again. As such, the results of the above analysis may not even necessarily be an indication that the ULEZ has a net negative effect on house prices - they could alternatively be rationalised as the product of asymmetric information transmission between its positive and negative effects, which leads to a temporary fall in prices.

From a policy perspective, this framework is a very important way to think about the long-term impact of ULEZ, and emissions charges elsewhere. What matters from a long-term perspective is not the relative speed at which the positive and negative effects are revealed, but their relative valuation once they become clear. The dynamics of the estimated treatment effect, and the discussion above, mean that policymakers must not judge the overall success (or failure) of the policy until they have let the population internalise the full extent of the effects. It may be the case that the net effect by this point is positive, and this cannot be determined until enough time has been given to the local residents. Regardless of this though, the analysis has shown the short-term effects are too large to be simply ignored in favour of longer-term goals.