The London School of Economics and Political Science

Essays in Law and Urban Economics

Di Song Tan

A thesis submitted to the Department Economics of the London School of Economics and Political Science for the degree of Doctor of Philosophy, April 2022

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 39,831 words.

I confirm that Chapter 1 was jointly co-authored with Yi Jie Gwee, and I contributed 50% of this work.

I confirm that Chapter 2 was jointly co-authored with Yi Jie Gwee and Jacob F Field, and I contributed 45% of this work.

Abstract

This thesis consists of three chapters that investigate the impact of the attributes of judges and juries, as well as the structures in which they make their decisions, on major societal outcomes such as urban development and criminal convictions. The first chapter examines how NIMBY-ism ("not in my back yard") affect housing prices. In the UK, developers who have their plans rejected by local authorities can appeal the rejection via the Planning Inspectorate. Our empirical strategy exploits the quasi-random assignment of inspectors to appeals. This allows us to use inspector leniency as an instrument for whether an appeal is successful. We find that overturning the local authority's decision does not lead to a large fall in housing prices. This suggests (i) local homeowners could be misinformed about the price impact of marginal local development projects; (ii) house prices are not the primary motive of NIMBY residents' opposition of local residential developments; or (iii) opposition against development projects being driven by immediate neighbours rather than by the majority of neighbours in the vicinity of a development project.

The second chapter uses the 1666 Great Fire of London as a natural experiment to study whether natural disasters can eliminate development frictions and bring about higher quality structures in the rebuilding process. First, using a difference-in-differences (DiD) strategy, we show evidence that the Fire resulted in higher quality structures. Second, using a DiD and an Instrumental Variables strategy, we find evidence that legal rulings arising from the Fire Court (a court specially set up by the English Parliament to hear rebuilding disputes) were able to anchor expectations and facilitate the development of London.

In the third chapter, I build a structural model to understand how jury composition interacts with peremptory strikes to affect verdicts. I find that strikes are strategic complements, suggesting a tit-for-tat strategy where lawyers use strike to balance

each other out. While the racial composition does favour the prosecutor, increases in the allowed number of strikes do not necessarily benefit her more than the defence.

Acknowledgements

This thesis is the culmination of many hours of fun and insightful conversations with supervisors, PhD colleagues, friends and family.

I would like to thank my supervisors Guy Michaels and Vernon Henderson for their guidance. Guy is a man of details and I benefited immensely from his numerous annotations on my papers and presentation slides, as well as his point-by-point post-presentation debriefs. I only regret that we could not continue our meetings in person, after the onset of the COVID-19 pandemic, and I finished up my thesis away from London. Although I took longer than expected to finish up my thesis, Guy has been nothing but patient and kind. I also benefited immensely from Vernon's sharp questions and breadth of knowledge about the urban economics literature. Due to time differences, he often had to wake up early to provide feedback on my work, which I really appreciate.

I also benefited from the comments at the Labour Work-in-progress seminars. In particular, I could always count on Steve Pischke and Alan Manning to give critical and constructive feedback on my work. My PhD colleagues Thomas Brzustowski, William Matcham, Jamie Coen, Bilal Tabti, Daniel Albuquerque, Sacha Dray, Heidi Thysen, Nicola Fontana, Hugo Vilares and Kohei Takeda also provided valuable feedback and awesome company over meals, coffee breaks and casual chats.

I would also like to thank my co-authors Jacob Field and Yi Jie Gwee. Jacob kindly shared the data he painstakingly compiled for his PhD for our paper. Yi Jie is a dream of a co-author, many ideas in this thesis were born out of the collision of my crazy ideas with his scepticism and encyclopaedic knowledge of the economic literature. Our work benefited from his unwavering commitment to impeccable empirical work and to ground this work in economic theory.

My partner Huishan Wang provided unrelenting support in my PhD journey, especially in moments of stress and self-doubt. Together we built a cosy and and vibrant life in London with our two boys Asher and Anders. I regret that we had to leave London abruptly, but the memories will outlast my ability to recall the key ideas in this thesis.

Contents

Chapter 1

The Effect of Overcoming NIMBY-ism on Housing Prices

1.1 Introduction

The introduction of land use regulations is often predicated on efficiency grounds (e.g., separating different types of land use would limit negative externalities). However, recent theories concerning the determinants of land use regulation have highlighted that this might not always be the case. In their handbook chapter on regulations and housing supply, Gyourko and Molloy (2015) note that land use regulations are often "shaped by the incentives and influence of actors in the local political process". Given that homeowners constitute a sizeable proportion of the voting pool in local elections, Fischel (2001) argues that "concern for home values is the central motivator of local government behaviour". These political economy arguments have drastic implications for the type of developments that end up being built or vetoed in each locality.[1]

For example, if homeowners have NIMBY ("not in my back yard") mindsets, such as the desire to constrain housing supply in their locality (but not other localities) to keep their home prices high, then land use regulations would reflect this NIMBY behaviour by making it hard for new housing units to be built. Taken to the extreme, local governments could end up constraining supply at all cost. This means that new developments that come with, or draw in, attractive amenities could end up not being built simply because they would have increased the housing

[1]The issue of local influence is also a salient issue: Cheshire et al. (2014) note that in recent times, countries such as the United Kingdom (UK) are increasingly moving towards localism.

stock substantially.

Even if local governments do not have the objective to "constrain supply at all cost", we can still expect that at the margin, projects that give rise to attractive amenities could still be blocked. This is because local governments do not know for sure whether the demand effects arising from the new developments (e.g., attractive amenities) dominate the supply effects (e.g., more housing units). To do so they would need an accurate estimate of demand and supply curves for housing in each locality. Consequently, marginal projects might get rejected even though their demand effects dominate.

Our paper (i) models how local decision making, which may be influenced by NIMBY-ism, affects the type of developments that end up being built or vetoed in each locality; and (ii) estimates the causal effect of overcoming NIMBY-ism on housing prices. We do so in the context of land use regulations in the UK, where individuals and developers have to apply to the local planning authority (LPA) to seek development permission (ownership alone does not confer the right to develop land in the UK). Applicants who have their plans rejected can appeal to the Secretary of State, via the Planning Inspectorate. The Planning Inspectorate then assigns an inspector to decide whether to overturn the local authority's decision.

We follow the institutional context closely in our model and show how NIMBY-ism affects the type of developments that end up being built or vetoed in each locality. In particular, we model the application and approval process as a multi-stage sequential game between local planning authorities and inspectors. In our model, the local planning authority accepts or rejects the project based on the level of NIMBY-ism in the area.

Our model yields two key propositions. First, there is negative selection into appeals because the developments developments we observe in appeals (i.e., rejected by local planning authorities) are likely to be the ones with a more negative impact on local amenities. Second, local power exacerbates selection into appeals. This is because an increase local power, which enables more NIMBY-ism, makes marginal developments that reduce amenities less desirable.

This analysis suggests that we can understand the prevalence of NIMBY-ism from the selection of projects into our appeals sample. If only the worst projects, which have a deleterious impact on the neighbourhood, select into appeals, then NIMBY-ism is not so serious; the planning process is working as intended as bad apples are rejected by the planners. On the other hand, if fairly benign projects select into appeals, then NIMBYs may be prevailing.

The next part of the paper embarks on our reduced form analysis as informed by our model. We begin by examining the effect of a successful appeal (i.e., overturning the local authority's decision) on housing prices. To establish causality, we employ two research designs. First, an OLS using a difference-in-differences (DiD) approach; and second, an instrumental variable (IV) DiD approach. The latter exploits the fact that inspectors who are assigned to the appeals have different preferences and are quasi-randomly assigned to the cases. This means that we can use the leniency of the inspectors as an instrument for whether an appeal is successful.[2] We find that overturning the local authority's decision does not lead to a large fall in housing prices. In fact, our IV estimates suggest a positive impact of overturning decisions.

We interpret projects that were successfully appealed as marginal development projects that managed to overcome local NIMBY-ism. Our finding that there is little evidence that such developments reduce housing prices in the vicinity suggests (i) local homeowners could be misinformed about the price impact of marginal local development projects; (ii) house prices are not the primary motive of NIMBY residents' opposition of local residential developments; or (iii) opposition against development projects being driven by **immediate** neighbours (who would be affected by construction noise or hampered views) rather than by the majority of neighbours in the vicinity of a development project (who may benefit from amenities associated with the marginal development project).

Our paper contributes to the burgeoning literature on the effects of land use regulations in two ways. First, the impact of land use regulations on housing

[2]The use of expert leniency as an instrumental variable has been used widely in the economics literature. For example, the leniency of judges has been used to estimate the impact of eviction on poverty (Humphries et al. (2018)), incarceration on economic and family outcomes (Aizer and Doyle Jr (2015); Bhuller et al. (2020)), the effect of pretrial detention on legal and economic outcomes (Dobbie et al. (2018)), the effect of foster care on child outcomes (Doyle Jr (2008)), the effect of disability on labour supply (Dahl et al. (2014); Kostøl et al. (2019)), and even the effect of patents on innovation (Galasso and Schankerman (2015)).

prices have been documented in many empirical studies such as Albouy and Ehrlich (2018), Shertzer et al. (2018), Turner et al. (2014), Libecap and Lueck (2011), Saiz (2010), Glaeser et al. (2005), Quigley and Raphael (2005), Hilber and Vermeulen (2016), Koster (2020) and Cheshire and Sheppard (2002). Our paper is also related to the literature on the effect of land use regulations on welfare. For example, Turner et al. (2014) and Hsieh and Moretti (2019) show that land use regulations lead to large decreases in welfare.

Second, our paper demonstrates why it is not correct to view land use regulations as simply being a supply shock.[3] Before Turner et al. (2014), the literature often took contradictory positions as to whether increases in housing prices indicate welfare increases or decreases. One position was that price increases reflected welfare improvements. This view is predicated on viewing land use regulations as improving the amenities of an area and hence a demand shock. The other position was that price increases reflected welfare lost. In this framework, land use regulations are viewed as a supply shock. Following Turner et al. (2014), we model land use regulations as being both a demand and supply shock. We then show how land use regulations being both a demand and supply shock have political economy implications.

In terms of the political economy literature, we contribute by empirically quantifying how overcoming NIMBY-ism affects housing prices. In recent times, there has been a growing number of papers such as Parkhomenko (2020) and Ortalo-Magné and Prat (2014) that formalize Fischel's home voter hypothesis by modelling land use regulations as a function of political competition between renters and owners. Besides the role of homeowners in the local political process, Calabrese et al. (2007), Epple et al. (1988) and Hamilton (1975) argue that land use regulations are ways to prevent certain types of households from entering a community. Similar to these papers, our paper models land use regulations as a political economy mechanism to protect the interest of particular local groups ("insiders") at the expense of other groups ("outsiders"). In addition, just like Hilber and Robert-Nicoud (2013) and Glaeser et al. (2005), we also show how developers react to the local political

[3] Gyourko and Molloy (2015) note that the predicted effects of regulations become less obvious if households can move freely among cities. This is because under spatial equilibrium, utility must be equal across cities. Therefore, population flows can erode away price differences across locations. The price differences across areas thus reflect the amenity value of growth controls and not the lower elasticity of housing supply.

process. Most of these political economy papers are mainly theoretical papers or calibration exercises which do not quantify the effect of NIMBY-ism. Our paper therefore contributes to this literature by providing an empirical estimate of how NIMBY-ism affects housing prices.

More broadly, our paper is related to the literature which examines the effects of housing policies. Examples of housing policies other than land use regulations include, urban revitalization (Greenstone and Gallagher (2008); Rossi-hansberg et al. (2010)), direct provision of housing (Collins and Shester (2013)), price controls (Autor et al. (2014); Diamond et al. (2019)), and taxes and subsidies (Baum-Snow and Marion (2009); Collinson and Ganong (2018); Diamond and McQuade (2018)). Since land use regulations are a form of place-based policies, our paper relates to the literature on the use of place-based polices to rectify regional disparities. Kline and Moretti (2014) provide a good overview on the economics of place-based policies. In addition, because land use regulations limit the size of cities, our paper is also related to the literature on optimal city size. Au and Henderson (2006) find that migration restrictions have resulted in many undersized cities in China. To the extent that there are huge benefits from urban agglomeration, the costs of being undersized due to land use regulations are potentially high.

Finally, our paper is related to the literature on the economic effects of the decentralisation of policy responsibilities to local governments. A review by Pike et al. (2012) suggests that there is no clear relationship between decentralization and economic outcomes. For example, Zhang and Zou (1998) find that decentralization is associated with lower economic growth at the provincial level in China. However, Akai and Sakata (2002) and Stansel (2005) find a positive relationship between decentralization and economic growth in the US. Yet another group of studies such as Davoodi and Zou (1998), Xie et al. (1999) and Rodríguez-Pose and Bwire (2004) find that there is no link between decentralization and economic outcomes. Much of this literature is somewhat dated and does not establish causality. By examining how the decisions of local governments are exogenously overturned, our paper presents an attempt at reviving this literature.

The rest of the paper proceeds as follows. Section 2 provides the background to land use regulations in the UK and the rise of local power. Section 3 develops the theoretical model which we use to interpret our reduced form regressions. Section 4 describes the data sources which we use for our analysis. Section 5 explains our empirical strategy and Section 6 presents the results. Finally, Section 7 concludes.

1.2 Background to Land Use Regulations in the UK

1.2.1 Overview

Land use regulation in the UK has a long and storied history – Corkindale (1999) notes that as early as 1540, Queen Elizabeth I forbade any new buildings within three miles of the City gates of London. The present day land use regulations in the UK has its roots in the Town and Country Planning Act which came into law in 1947. The purpose of the 1947 act was to contain urban areas and stop them from spilling out into the surrounding countryside as well as preserve amenities of various kinds. The act also separated land uses which might be incompatible (e.g., industry from residential). At the same time, the act also aimed to provide lower density and greener living conditions in the new towns.

In order to do so, the act set out the principle of "development control" which means that planning permission is required for land development – ownership alone no longer conferred the right to develop the land. As a result, owners had to apply and seek permission from their local planning authorities whenever they wanted to develop their land. The local planning authorities would only approve the development if it is consistent with the local development plan.[4] The UK system therefore controls independently of prices, the amount of land available not just for housing but for all urban uses of land including offices, retail or commercial use.

There have been many modifications since 1947.[5] For example, in the 1950s, the restrictions were tightened as "Greenbelt" boundaries were established. This resulted in increasing amount of land around towns and cities being taken out

[4]Today, many parts of England have three tiers of local government – (i) county council, (ii) district, borough or city councils and (iii) parish or town councils. The district councils are the ones that are typically responsible for most planning matters including preparing local development plans and approving planning applications.

[5]Each country within the UK has its own variations in their planning system. Since the English system is the dominant one, the details which we provide here are based on the system in England.

of the effective land supply. Later revisions of the act were legislated in 1962, 1971 and 1990. While the 1990 act is the current legislation, the act has been substantially amended and added to, for example, through the use of legislative orders. Nonetheless, the 1947 act established an approach and framework that has not been superseded.

Today, the main planning acts that are in force in England are the Town and Country Planning Act (1990), Planning and Compulsory Purchase Act (2004), Planning Act (2008) and Localism Act (2011). These acts are also read alongside a series of planning policy documents and guidelines such as the National Planning Policy Framework which was published in March 2012. Other national policies and restrictions such as "Greenbelts", "Sites of Special Scientific Interest" and "Areas of Outstanding Natural Beauty" also form part of the planning regulations. Many of the recent reforms such as the Localism Act (2011) was aimed at giving locals more power in deciding land use in their neighbourhood. For example, neighbourhood plans were introduced by the Localism Act (2011). These gave communities the direct power to develop a shared vision for their neighbourhood and shape the development and growth of their local area. The neighbourhood plans are extremely powerful. This is because when considering whether to grant planning permission, the local authorities have to ensure that the proposed development is consistent with the neighbourhood plans. [6] More details about neighbourhood plans can be found in Appendix A.

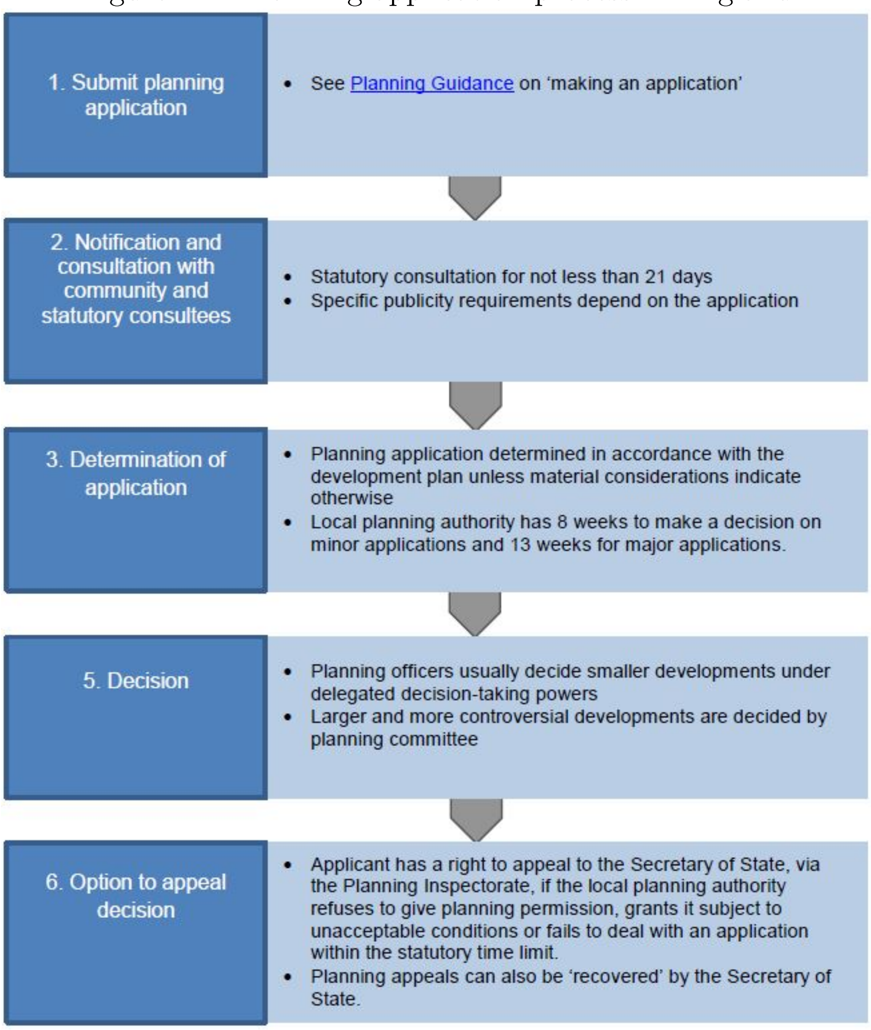
Figure 1.1 shows the stages in the planning application process in England.[7] A planning application needs to be submitted to the local authorities if an individual or developer wants to (i) build something new, (ii) make a major change to the building (e.g., building an extension), or (iii) change the use of the building. In determining whether to grant planning permission, the local authority has to assess whether the proposed development is consistent with (i) national policies, (ii) the local plan and (iii) the neighbourhood plans (if any exists). Since most developments put a strain on existing infrastructure such as roads, schools and open

[6] The UK's land use regulations are typical of the planning systems in OECD countries. Common features include (i) land use planning is predominantly a local task but involves some coordination between different levels of government; (ii) formal stakeholder engagement involve a public consultation process; (iii) land use regulations are well enforced; (iv) taxes levied on the increases in land value from redevelopment; and (iv) the public expropriation of land for the construction of infrastructure is possible.

[7] Based on documentation from GOV.UK

spaces, the local planning authorities can impose a Community Infrastructure Levy (a charge which new developments pay, based on the size and type of development) to mitigate the impact of the proposed development. The Levy collected is then used to fund a wide range of infrastructure needed to support the development of the area. Alternatively, under Section 106 of the Town and Country Planning Act (1990), the local planning authority can also grant permission in return for some specified gain to the community. For example, the local planning authority may require that the developer provides a certain amount of affordable housing. This gives rise to a negotiation process between the local planning authority and the developer.

Figure 1.1: Planning application process in England



Source: Plain English guide to the Planning System (2015)

Applicants who have their plans rejected can appeal to the Secretary of State, via the Planning Inspectorate. The Planning Inspectorate then assigns one inspector to the case who decides whether to overturn the local planning authority's decision.[8]

Data from the UK's Ministry of Housing, Communities & Local Government indicate a rejection rate of around 20% for all planning applications for major residential developments (see Section 4 for our reasons for focusing on major residential developments) in the UK from the fourth quarter of 2012 to the fourth quarter of 2017. Based our reading of a sample of appeals, common reasons cited for the rejection of plans include (i) detrimental impact on neighbouring amenity (e.g., lost of green or communal spaces); (ii) negative effect on character and appearance (e.g., proposed architecture of housing is not consistent with that of other buildings in the area); and (iii) traffic and parking pressures (e.g., additional residents would lead to severe congestion of local roads).

Among these major residential developments that were rejected, 83% appealed the decision; and about 43% of these appeals were successful (i.e., initial rejection was overturned and development could proceed as planned). While the administrative costs of lodging an appeal is fairly low (e.g., the developer just needs to fill out a form and forward the documents that were part of the submission to the local planning authority), the mean turnaround time for appeals was around 26 weeks.[9] Although this represents a significant delay in the development timeline, developers also incur other costs from not appealing (e.g., reworking plan according to local planning authority's requirements, which may result in loss of expected profits from development). The high rate of appeal suggests that the costs of appealing are not sufficiently high, relative to that of not appealing, to deter most developers.

In order to understand how inspectors are assigned to cases, we spoke to a representative from the Planning Inspectorate. The representative shared with us that the Planning Inspectorate uses an algorithm as well as human judgment to decide on the assignment of inspector to cases. First, the algorithm uses the (i) complexity of the case (e.g., size of the development, whether the proposed development is in

[8] The Secretary of State has the power to take over the decision making from the Planning Inspectorate if the case raises particular issues that justify a Ministerial decision. In such cases, a planning inspector will submit a report and recommendation to the Secretary of State. Taking into account the inspector's assessment of the proposals, the Secretary of State will then make a decision.

[9] Analysis from LandTech.

a "Greenbelt", agricultural area, "Area of Outstanding Natural Beauty", "Sites of Special Scientific Interest", etc.), (ii) inspectors qualifications and experience and (iii) how far away each inspector lives from the location of the proposed development to identify a list of suitable inspectors.[10] Second, a case worker then manually goes through the list and excludes inspectors who (i) already have a heavy case load, (ii) recently had a case in the area where the development is being proposed, (iii) live in the area of the proposed development, (iv) previously worked for a consultancy firm that is involved in the appeal, and/or (v) have personal difficulties such as illness. This means that conditional on these variables, the assignment of inspectors to cases is as good as random. This forms the basis of our empirical strategy to identify the causal effects of having a successful appeal.

1.3 Model

1.3.1 Overview

The aim of our model is to help interpret our results from the empirical analysis. We study developments in the UK that went through the appeals process. These projects are self-selected and would not be similar to projects that that did not appeal. Hence, we use the model to analyse this selection effect and how it changes with NIMBY-ism. The model is a sequential game between the local planning authority (L) and inspector (N) played out in three stages:

Table 1.1: Stages of the game

Stage	Player	Action set	Consequence of action
0	Nature	Draws κ, s and η	All players observe κ and s, only N observes η
1	L	Accept	$\{\kappa, s\}$ implemented; move to stage 3
		Reject	Move to stage 2
2	N	Accept	$\{\kappa, s\}$ implemented; move to stage 3
		Reject	No development
3	Housing market clears		

A land development plan, defined as the pair $\{\kappa, s\}$, is submitted to L. If accepted, this development adds an exogenous amenity value of $\kappa \in (-\infty, \infty)$ to the location. The destruction of green spaces can be represented as a negative κ, while the removal of an abandoned building as a positive κ. A value of $\kappa = 0$

[10]The Planning Inspectorate maintains a list of inspectors which they classify based on the inspectors' suitability to assess cases according to the different levels of case complexity.

means there is not change to the amenity value relative to the status quo. s is the intensity of residential development, where $s \in [0, \infty]$. A higher s means more residents in the location which creates congestion and reduces the amenity value of the location.[11] If the development is accepted, $P(\kappa, s)$ is the market clearing price of the development.

We assume that all residents (current and potential) are homeowners, which allows us to study the impact of wealth effects (i.e., property prices affecting the local land use decisions). This abstracts from the influence of renters on the decisions of the local planning authority, which may counteract the influence of homeowners. For instance, while a decline in prices is a detriment to the wealth of homeowners, these declines may translate to lower rents for renters (i.e., a cost to homeowners but a benefit to renters). Thus we think of the proposition derived from this model as valid for the situation when the incentives of homeowners dominate those of renters (e.g., renters are footloose and do not participate in local politics and decision making).

We proceed to analyse the game backwards from stage 3:

1.3.2 Stage 3: Household behaviour and market clearing

Resident homeowners

There is a mass $\underline{s} \in [0, 1]$ of resident home owners who are deciding between residing in the location or selling their property and moving elsewhere (outside option). Staying provides a utility of $W_r + \psi(\kappa, s)$, where W_r is the wealth (and hence composite consumption) of the residents, $\psi_\kappa(.) > 0$ and $\psi_s(.) < 0$ (i.e., amenities). The residents have homogenous utility from residing but heterogeneous outside utility that is increasing in the price from selling their property (P) and moving elsewhere. Specifically, resident i has an outside utility of $P(\kappa, s) + u_i$ and would choose to leave if $u_i \geq W_r + \psi(\kappa, s) - P(\kappa, s)$.[12] The mass of potential residents with an outside utility of u and lower is $F_r(u)$. Therefore, the mass of residents who want to leave ($\dot{s}$) is such that

[11] One way to think about congestion is that of more residents competing for scarce local services like schools and transport.

[12] This is consistent with u_i reflecting the amenity and composite consumption value of residing elsewhere, and $P(\kappa, s)$ reflecting the additional consumption value from the sale of their property (i.e., wealth effect)

$$\dot{s}\,(\kappa,s) = \underline{s} - F_r\,(W_r + \psi(\kappa,s) - P(\kappa,s))$$

And the mass of residents who stay is $\tilde{s}\,(\kappa,s) = F_r\,(W_r + \psi(\kappa,s) - P(\kappa,s))$

Potential homeowners

There is a mass of potential resident who are deciding between purchasing a property in the location and living elsewhere (outside option). Moving to the location provides a utility of $C + \phi(\kappa,s)$ subject to $C + P \leq W_p$, where W_p and C are the wealth and the composite consumption of the potential residents. The potential residents have homogenous utility from residing but heterogeneous outside utility. Specifically, potential resident j has an outside utility of u_j and would choose to reside in the location if $u_j \leq W_p + \phi(\kappa,s) - P(\kappa,s)$. The mass of potential residents with an outside utility of u and lower is $F_p(u)$. Therefore, the mass of potential residents who want to reside in the location ($\dot{m}$) is such that

$$\dot{m}\,(\kappa,s) = F_p\,(W_p + \phi(\kappa,s) - P(\kappa,s))$$

For the the market to clear, the residents who want to leave and the new supply of housing from the development must be taken up by potential residents. Therefore:

$$\begin{aligned} s + \dot{s}\,(\kappa,s) &= \dot{m}\,(\kappa,s) \\ \implies s + \underline{s} - F_r\,(W_r + \psi(\kappa,s) - P(\kappa,s)) &= F_p\,(W_p + \phi(\kappa,s) - P(\kappa,s)) \end{aligned} \tag{market clearing}$$

1.3.3 Stage 2: Inspector decides whether to overturn the local planning authority's decision

The inspector (N) compares the social utility of accepting the project versus rejecting it. She observes η, which is the weight of current residents who stay relative to the weight of new residents. So the social utility of accepting is:

$$\eta * \tilde{s}\,(\kappa,s) * (W_r + \psi(\kappa,s)) + (s + \dot{s}\,(\kappa,s)) * (W_p + \psi(\kappa,s) - P(\kappa,s))$$

Current residents, who choose to stay are weighted η, get their initial level of consumption (W_r) and enjoy the net stock of amenities. New residents, who are weighted 1, enjoy the same stock of amenities.

The social utility of rejecting is simply:

$$\eta * \tilde{s}\,(0,0) * (W_r + \psi(0,0)) + \dot{s}\,(0,0) * (W_p + \psi(0,0) - P(0,0)) \equiv \eta R_r + R_p$$

That is the market clears with the initial level of housing and amenities.

L does not observe the value of η but know its distribution (cdf of $G(\eta)$ and pdf of $g(\eta)$). Given a value of s we can rewrite the condition for accepting as the range of values of η for which N will accept:

$$\eta \leq \frac{(s + \dot{s}\,(\kappa, s)) * (W_p + \psi(\kappa, s) - P(\kappa, s)) - R_p}{R_r - \tilde{s}\,(\kappa, s) * (W_r + \psi\,(\kappa, s))} \equiv I(\kappa, s)$$

Hence, the probability of a successful appeal is $G\,(I(\kappa, s))$.

1.3.4 Stage 2: Local planning authority decides whether to accept or reject the developer's proposed plan

The local planning authority (L) knows that it can accept a development and it will be built. But if it rejects, N may still accept it. The social utility of L is similar to N but instead of η, she weighs current residents with a value of ω relative to new residents. She does not know the exact η but she knows the distribution of η. Assuming L is risk neutral, she will accept a development if and only if

$$\omega * \tilde{s}\,(\kappa, s) * (W_r + \psi(\kappa, s)) + (s + \dot{s}\,(\kappa, s)) * (W_p + \phi(\kappa, s) - P(\kappa, s)) \geq \omega R_r + R_p$$

1.3.5 Selection into appeals

We hope to understand the range of κ that will result in appeals. We hope to understand the range of κ that will result in appeals. To make some progress we make the following assumptions:

1. $\omega \geq \frac{W_p}{W_r} \geq 1$. The assumption $W_p \geq W_r$ is natural because potential homeowners have not spent their wealth on housing. $\omega W_r \geq W_p$ constrains social utility such that L would not strictly prefer all current residents to be replaced by new ones simply because new residents are substantially wealthier.

2. $\psi(\kappa, s) \geq \phi(\kappa, s)$ and $\psi_\kappa(\kappa, s) \geq \phi_\kappa(\kappa, s)$. These assumptions can be thought of as a type of endowment effect, where residents value their amenities, as well as any improvement in amenities, greater than outsiders. They also constrain the social utility such that L would not strictly prefer all current residents to be replaced by new ones simply because new residents value the amenities more.

3. $s = 0$. To simplify the analysis we shut down the supply effects. This constrain the social utility such that L would not accept developments due to the fact that there is more housing and a mechanical higher utility because of more residents in the location. This assumption is also justified because our data suggests the supply shock from the data is quantitatively small.

If we denote the weighted social utility of accepting as $U(\kappa)$ and the weighted social utility of rejecting as the constant R then:

Proposition 1 (Negative selection into appeals) *Given assumptions 1, 2 and 3,* $\exists \kappa^* : \kappa = \kappa^* \Rightarrow U(\kappa) = R$ *and* $\kappa < \kappa^* \Rightarrow U(\kappa) < R$ *and* $\kappa \geq \kappa^* \Rightarrow U(\kappa) > R$

Proof: See Appendix B1. □

Appeals take place only when $\kappa < \kappa^*$. This says that the developments we observe in appeals are likely to be the ones with a more negative impact on amenities, because they are the ones that result in a lower weighted social utility.

Proposition 2 (Local power exacerbates selection into appeals) *If* $\kappa^* < 0$ *then* $\omega' \geq \omega \Rightarrow \kappa^*(\omega') \geq \kappa^*(\omega)$. *If* $\kappa^* > 0$ *then* $\omega' \geq \omega \Rightarrow \kappa^*(\omega') \leq \kappa^*(\omega)$.

Proof: See Appendix B2 □

An increase in ω accentuates the difference in weighted social utility between marginal developments (i.e., developments with $\kappa = \kappa^*$) and maintaining the status quo (i.e., $\kappa = 0$). Therefore, an increase in local power makes marginal developments that reduce amenities less desirable, and makes marginal developments that increase amenities more desirable.

1.3.6 Implications of selection on price regressions

Given propositions 1 and 2, we can now work out what our regressions are estimating.
We consider two regressions: the first is an OLS regression $ln(p_i) = \beta_0 + \beta \mathbf{1}.\text{Successful Appeal}_i + e_i$, where the variation in successful appeals is generated by the random assignment

of inspectors (or the random variable η). Here our coefficient of interest is β. The second is an 2SLS regression, where we run the regressions $ln(p_i) = \gamma_0 + \gamma_1\eta_i + u_i$ and $\mathbf{1}.\text{Successful Appeal}_i = \pi_0 + \pi_1\eta_i + v_i$, and our coefficient of interest is $\frac{\gamma_1}{\pi_1}$.

OLS versus IV regressions

The price change when we randomly assign η is: $\Delta P = E\left[P(\kappa, s) - P(0,0)|\kappa < \kappa^*\right]$.

Hence we know that β_1 in our OLS regression will be:

$$\frac{\Delta P}{P_0} = C\int_{\underline{\kappa}}^{\kappa^*} \left[P(\kappa, s) - P(0, s)\right] dH(\kappa)$$

$$C \equiv \frac{1}{H(\kappa^*) - H(\underline{\kappa})} * \frac{1}{P_0}$$

$\underline{\kappa}$ is lowest κ, for which the market still clears at a price ≥ 0, and $H(\kappa)$ represent nature's distribution of κ. There are two main effects that influence the sign of $P(\kappa, s) - P(0, 0)$:

1. $P(\kappa, s) - P(\kappa, 0)$ is a supply effect, or a move along the demand curve, and is always negative.

2. $P(\kappa, 0) - P(0, 0)$ is a demand shifter, and depends on whether κ is greater or less than 0.

We can also work out that $\frac{\gamma_1}{\pi_1}$ in our IV regression will be:

$$C\int_{\underline{\kappa}}^{\kappa^*} \tilde{F}\Big(I(s, \kappa)\Big) \left[P(\kappa, s) - P(0, 0)\right] dH(\kappa)$$

where $\tilde{F}(I(s, \kappa)) = \frac{g(I(s,\kappa))}{\int_{\underline{\kappa}}^{\kappa^*} g(I(s,\kappa))dH(\kappa)}$The key difference between the OLS and the IV estimates are the weights (1 versus $\tilde{F}(.)$). The OLS estimates the Average Treatment Effect (ATE), while the IV estimates a Local Average Treatment Effect (LATE), which skews towards developments that have a high *increase in the probability of success* from being assigned an inspector with a higher η. Under an assumption of homogeneity, that is κ is a constant, $\tilde{F}\Big(I(\tilde{s}, \kappa)\Big) = 1$ and the OLS and IV estimates are the same. This forms the basis of our regression analysis in Section 5.

1.4 Data

1.4.1 Description of datasets

The data we use are from four sources: (i) the Planning Inspectorate's Appeals Casework Portal (ACP); (ii) the UK Land Registry's Price Paid Database (PPD); (iii) the Domestic Energy Performance of Buildings Registers; and (iv) the Royal Mail Postal Address File (PAF).

Planning Inspectorate's Appeals Casework Portal (ACP). The ACP is the same database that the Planning Inspectorate caseworkers use to assign cases to inspectors and to manage cases. The ACP dataset contains decisions on appeals starting from October 1, 2012 to December, 28, 2018. It contains the following key variables:

Table 1.2: Summary of ACP dataset

Fields	Description
Local Planning Authority (LPA) details	Name of local planning authority (LPA), LPA code, LPA's case ID (the case ID allows us to search for documents related to appeal, e.g., site maps)
Key dates	Date when the appeal was received, started and decided
Inspector	Names of the inspectors
Development details	Residential or commercial, address, floor space, site area, number of residences whether the proposed development is in a "Greenbelt", "Area of Natural Beauty", "Sites of Special Scientific Interest", or agricultural area; whether it involves a conservation area or historical building
Others	Type of appeal, whether a planning consultant was hired, whether a bespoke timeline was agreed between the parties, whether the inspector ordered the local planning authority to foot the cost of the developer's appeal

Since the ACP data is also used by the Planning Inspectorate for casework management, most fields are accurate.[13] There is, however, one key field that is subjective because it is filled in by the developer – development type. Therefore, we do not use this variable to construct our sample.

[13] At the very least, they are consistent with the information the inspectorate had in deciding the appeal.

UK Land Registry's Price Paid Database (PPD). For tax purposes, the PPD contains the registered price of all residential property transactions. The version of the dataset that we are using runs from January, 1, 2010 to January, 31, 2020. It includes the address of the property, the property type (i.e., detached, semi-detached, terrace or flat), whether the property is a new built and whether the property is freehold or leasehold. Crucially, the PPD does not include specific characteristics of the property such as the floor area or the number of rooms. Therefore, to obtain these variables, we use the Domestic Energy Performance of Buildings Registers.

Domestic Energy Performance of Buildings Registers. By law, Energy performance certificates (EPCs) are needed whenever a property is built, sold or rented. Besides the address of the property and its energy rating, this dataset also includes specific characteristics of the property like property type, floor area and number of rooms.

Royal Mail Postal Address File. This database contains all addresses and location of residential units in the UK. We use it to construct estimates of housing stock near the appeal site.

1.4.2 Sample for regression

We are interested in estimating the effect of a successful appeal (i.e., overturning the local authority's decision) on housing prices. To obtain the sample for our regression, we apply a number of sample restrictions.

Selection of appeals sample

Include only appeals involving residential developments that add 10 or more dwellings. There are a number of reasons why we apply this sample restriction. First, as noted in our background to land use regulations in the UK, almost any substantial changes one makes to a property requires planning permission – this includes, for instance, a loft conversion. This means that the ACP database includes appeals for a myriad of issues such as displaying advertisement on a building, house rear extensions and changes to commercial store front. Therefore, for this paper, we focus specifically on major residential developments, defined in the UK legislation as any development that adds 10 or more dwelling units.[14] We focus on housing because the debate on NIMBY-ism is generally about the restrictions in housing supply. Second, appeals involving smaller developments (less than 10) tend to comprise mainly of homeowners who are seeking to divide their house into two apartments to maximize rental yield or adding units to house family members. Such extensions and modifications are likely to be very different from a developer seeking to build residential units for sale. Third, there is a long tail of very small developments. For instance, more than 50 percent of residential developments add only one dwelling unit, and more than 75 percent of developments add two units. Since minor developments tend to have smaller externalities, we would be skewing our results towards finding zero effect if we include them in our sample.

Exclude variation in conditions and caravan parks. Variation in conditions are appeals that seek to amend the original plan approved by the local authority. For instance, a developer who planned for extensive landscaping, but now find it prohibitively costly, may appeal to the inspectorate to remove that part from the agreed plan. These appeals typically involve minor amendments and, if successful, barely change the development.

We also exclude caravan parks because it is not clear whether they are residential in nature. Some may serve as long term lots for families living in caravans but others may serve as winter parking or tourist lots.

Choose earliest appeal if several are near each other. This reduces the need to account for cross-appeal effects in our regressions. We can also think of the number of subsequent appeals in the area as an outcome variable that is influenced by the earliest appeal. This may introduce endogeneity into our regressions.

[14] Town and Country Planning Order (2010) available at http://www.legislation.gov.uk/uksi/2010/2184/made. Note that this definition would include mixed use or primarily commercial developments that also add 10 or more residential units. In addition, there are certain requirements that "major developments" have to adhere to and these requirements make it easier for locals to exercise their power to influence the local planning authority. For example, the developer is required to give notice of the planned development at the development site, serve the notice to adjoining occupiers and publish the notice in a local newspaper. These requirements mean that locals are more likely to be aware of such developments, enabling them to voice their opinions and hold the members of the local planning authority accountable if they approve the development.

Location plans available online. We had to geocode the planned development manually because many were developed on greenfield sites without existing addresses.

Include appeals for which inspector leniency is estimable. Section 1.5.3 details how we construct the instrument for our IV regressions: we need a decent sample of cases per inspector to ensure estimates of inspector leniency converge to the true leniency. Thus we excluded appeals for which the inspector did not have many cases.

The sample counts following the various restrictions are summarized here:

1. Major residential appeals 2012-2018 (i.e., $\geq$ 10 dwellings) $\rightarrow$ **4,324 appeals**
2. Exclude variation in conditions, caravan parks, choose earliest appeal $\rightarrow$ **4,211 appeals**
3. Able to manually geocode $\rightarrow$ **3,513 appeals**
4. Can estimate inspector leniency $\rightarrow$ **3,121 appeals**

Table 1.3 shows us the summary statistics of the appeals data, after applying the above sample restrictions.

Table 1.3: Summary statistics (Appeals data)

	Mean	SD	Min	Max	N
Year appeal started	2015.42	1.65	2012	2018	3121
Year appeal decided	2015.86	1.61	2012	2018	3121
No. of dwellings	56.19	100.93	10	4022	3121
Site area (hectares)	6.88	141.55	0	6780	2987
East Midlands	0.11	0.31	0	1	3121
East of England	0.14	0.35	0	1	3121
London	0.09	0.28	0	1	3121
North East	0.03	0.16	0	1	3121
North West	0.09	0.29	0	1	3121
South East	0.24	0.43	0	1	3121
South West	0.15	0.36	0	1	3121
West Midlands	0.09	0.29	0	1	3121
Yorkshire and The Humber	0.05	0.23	0	1	3121
Appeal successful	0.43	0.50	0	1	3121

Selection of price paid sample

Include only resale transactions of residential properties that are ever within 1km of any appeal. The reason for including only resale transactions is because we want to see how the appeals affect existing properties and not the new properties that are being built. As for the 1km restriction, we apply this because if we were to increase the radius beyond 1km, we end up having many properties that are linked to multiple appeals. In addition, the effect of having an appeal near-by is likely to decay with distance. If we define too big a radius, we would be skewing our results towards finding zero effect. In future work, we will be checking if our results are robust to using different distance thresholds.

Exclude "others" category of residential transactions. The dataset includes residential-related transactions such as garages and parking spaces that we exclude.

Include only resale transactions that took place within 3 years before or after an appeal. Since the appeals dataset only starts on October 1, 2012, we have fewer and fewer resale transactions beyond 3 years of an appeal. Statistical inference might become unreliable if the number of resale transactions in certain years before or after an appeal becomes too small. We will show as a robustness check that our results are stable to the choice of year bandwidths.

Trimmed top and bottom 1% in prices. Regressions may be sensitive to outliers and we trimmed extreme prices to reduce that sensitivity. Our results hold with an untrimmed or a Winsorised sample.

Table 1.4 shows us the summary statistics of the resale transactions data.

Table 1.4: Summary statistics (Resale transactions data)

	Mean	SD	Min	Max	N
Year of transaction	2015.88	2.07	2010	2020	1238407
Transacted price	297463.88	209187.77	54950	1504000	1238407
Terrace	0.29	0.45	0	1	1238407
Flat	0.27	0.45	0	1	1238407
Semi-detached	0.24	0.42	0	1	1238407
Detached	0.20	0.40	0	1	1238407
Freehold	0.70	0.46	0	1	1238407
Distance from first appeal	631.36	246.15	1	1000	1238407

1.4.3 What happens after a successful appeal?

So far we have assumed that successful appeals lead to new residential developments at the appeal site. Concretely, what it does lead to is an option to develop the appeal site. Developers may choose not to develop at all if property markets are poor.[15] Figure 1.2 plots the number of new dwellings, within 200m of the centre of appeal sites, before and after an appeal decision, and indicates that options do translate to new housing. Controlling for location and year fixed-effects, this translate to a 0.42 % increase in housing supply near the appeal site.

Figure 1.2: Comparison of new dwellings in appeal sites, by success of appeal

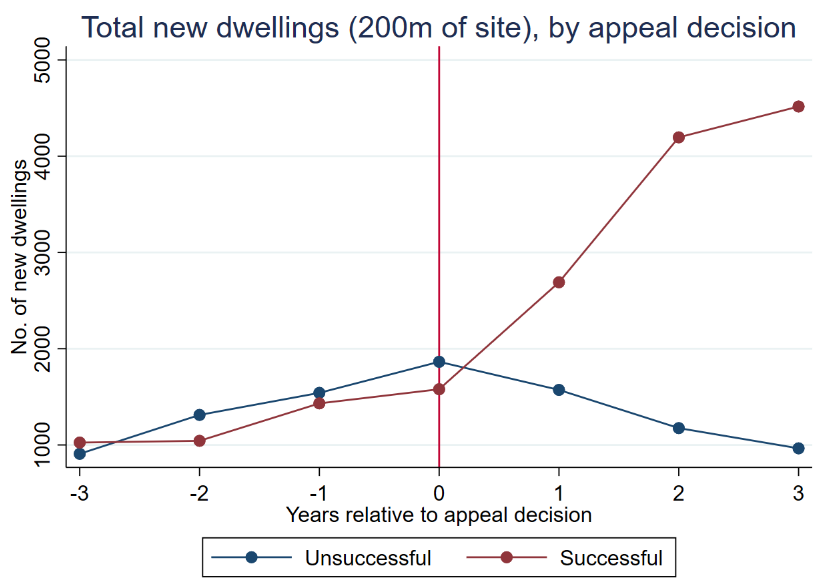
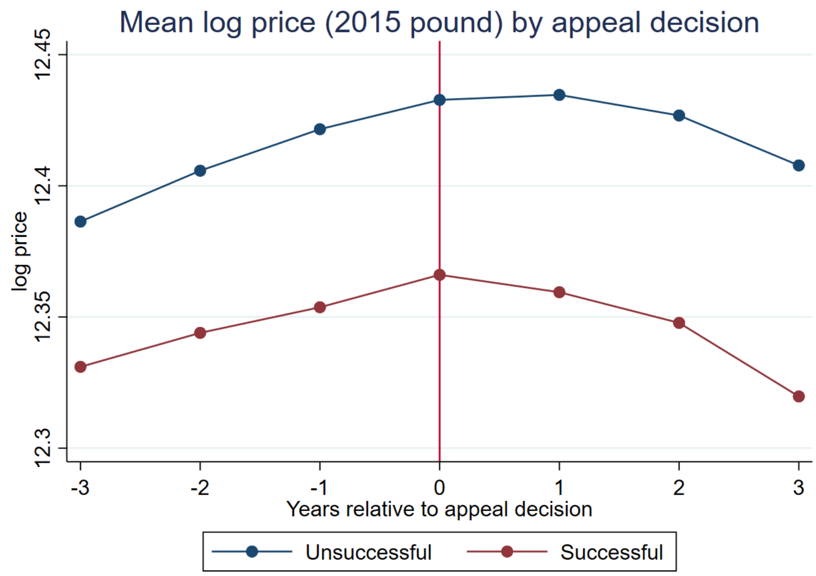


Figure 1.3 plots the average log real price by appeal decision. We see that successful appeals tend to be in neighbourhoods with lower house prices. This suggests neighbourhood characteristics, which determine market prices, may affect appeal decisions. For instance, properties near a green belt are priced higher, due to the supply restrictions; and the bar for an appeal there would also be higher. However, we see that the price trends prior to appeal decisions are similar across both groups.

[15] What they can do, however, is to sell the option to other developers. Regardless of market conditions, developers who have gone through the appeals process would have sunk in a lot of money (e.g. purchasing the land, drafting plans, hiring architects and consultants) and would usually develop the site except in exceptional circumstances.

Figure 1.3: Unconditional mean log price by appeal decision



1.5 Empirical strategy

1.5.1 Overview

To estimate the causal effect of overturning the local authority's decision on housing prices, we can run the following DiD regression:

$$ln(price_{ijt}) = \alpha_j + \delta_t + \beta Success_i \times Postappeal_t + \gamma' X_{ijt} + \epsilon_{ijt} \tag{1.1}$$

$ln\,(price_{ijt})$ is the log price of property i within 1km of appeal j in year-month t. $Success_i$ is an indicator variable that denotes whether property i is **ever** within 1km of a **successful** appeal. $Postappeal_t$ is an indicator variable for the period that the appeal takes place and periods after the appeal. X_{ijt} is a vector of individual property characteristics such as property type, whether the property is leasehold or freehold, the number of rooms, floor area, the property's current and potential energy efficiency, whether the property was built before 2012 and the supply trend in the 1km vicinity of the appeal prior to the appeal decision. α_j are fixed-effects for properties around appeal j. Finally, δ_t are time fixed effects. These consists of calendar year-month fixed-effects and relative year fixed-effects. For example, the relative year fixed-effect takes on the value of 1 if the sale of the property is within 1 year after the appeal. We cluster the standard errors at the outward code level.

Using the example of the postcode SE16 7BB, Table 1.5 shows how the varying levels of postcode granularity are defined in the UK. The outward code thus corresponds to a district in the UK (roughly the size of a town or part of a large town).

Table 1.5: Postcode format in the UK

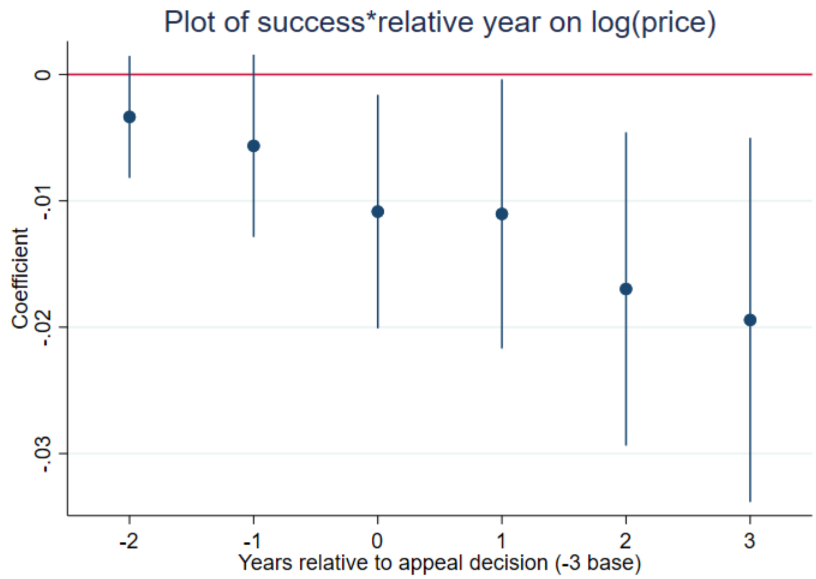
Postcode			
Outward code		Inward code	
Area	District	Sector	Unit
SE	16	7	BB

1.5.2 Limitations to Difference-in-differences

Figure 1.4 graphs the coefficients (with the 95% confidence intervals) from a dynamic DiD regression (with controls) that compares the prices, relative to three years before the appeal decisions, between properties near successful versus unsuccessful appeal sites.[16] There is no significant difference in price trends between the two groups, suggesting that a DiD specification may give a consistent estimate of the Average Treatment Effect (ATE) of having a successful appeal development nearby. The values of the coefficients used to plot Figure 1.4 are reported in Appendix Table C1.

[16] This regression includes the standard controls we use in our main specifications, as outlined in equation 1.1.

Figure 1.4: Difference in prices relative to 3 years before appeal decision



However, even if equation 1.1 passes the parallel pre-trends, we might still worry about unobserved time-varying neighbourhood characteristics that vary **after** the appeal and also affect housing prices. For example, after an appeal (regardless of whether it was successful or not), residents in high NIMBY areas might take the opportunity to quickly come up with a neighbourhood plan so as to make it difficult for other developers to develop new housing in their neighbourhood in the future. Therefore, to address such potential selection bias that happen after an appeal, we adopt an IV DiD approach. We exploit the fact that appeals are randomly assigned to inspectors (conditional on the assignment criteria used by the Planning Inspectorate which we elaborated on in Section 1.2). In addition, some inspectors are systematically more lenient that others. Taken together, these lead to random variations in the probability that an appeal will be successful based on which inspector the appeal is assigned to. Since there is no evidence of pre-trends in our DiD regression, when reporting our empirical results, we report both the DiD and IV DiD results.

1.5.3 Instrumental variable calculation

We measure the average leniency of an inspector based on the appeal success rate for all the other randomly assigned cases that the inspector handled.[17] These cases include both past and future appeals but not the existing appeal. This leave-out measure is important because it avoids introducing mechanical reverse causality. To construct the instrument, we follow the existing literature on expert leniency and regress whether the appeal is successful on the variables that the Planning Inspectorate uses to assign the inspectors to cases. The residual from this leave-one out regression is our leniency measure. The assignment criteria variables include the workload of the inspectors, characteristics of the area around the appeal site and the complexity of the case. Controlling for the assignment criteria is important because it accounts for the fact that randomisation by the Planning Inspectorate occurs within the pool of available and suitable inspectors.

Ideally, we would control for all the assignment criteria based on the exact way the Inspectorate's algorithm and case workers filter out inspectors to reach the final pool of suitable inspectors. However, we were unable to obtain these exact procedures from the Inspectorate. This is not surprising; the Inspectorate needs to keep these procedures opaque to prevent developers and local planning authorities from gaming the allocation process (e.g., try to alter case characteristics to obtain a "favourable" inspector). Instead, we proxy for the broad criteria revealed to us by the Inspectorate using data available (see table 1.6). The detailed regression results of the relationship between these variables and the probability of a successful appeal are included in C9.

[17] Although our regression sample involves only major dwellings, we also use the success rate for appeals involving non-major dwellings to construct the instrument. The purpose of doing this is to improve the power of the instrument.

Table 1.6: Assignment criteria for inspectors

Assignment criteria	Proxies available in data
Case characteristics	
Complexity of the case	Characteristics of appeals cases: key words (e.g., redevelopment, commercial), number of residences, case has bespoke timeline, developer appealing for costs, intermediary involved (e.g., consultant), linked to another appeal, type of appeal procedure (e.g., hearing, written), number of past appeals and success rate of these appeals in LPA
Inspector characteristics	
Qualification and experience	Total days taken to resolve appeals handled, case characteristics of appeals handled (see above)
Distance lived from appeal	LPA of appeal X year FE
Caseload	Number of live appeal cases
Another case near appeal	LPA of appeal X year FE
Lives in areaa appeal	LPA of appeal X year FE
Worked for consultancy firm	Whether intermediary involved
Personal difficulties	Not able to find appropriate proxy in data

While we use all the available data to estimate the leniency of the inspectors, after applying our sample restrictions (see Section 4.2), for our regression sample, we have 412 inspectors who were assigned to appeals involving major dwellings. Each of these inspectors presided over an average of around 65 appeals. The highest number of appeals presided over by an inspector is 313 and the smallest number of appeals presided over by an inspector is 1. Essentially, we estimate the leniency of inspectors using her mean residualised probability of approving an appeal. This estimate would be very noisy for inspectors with only a few appeals. For instance, the leniency would be estimated off 1 appeal if an inspector only judged 2 appeals in total.[18] Therefore, we did not use inspectors, in our regressions, who presided over <30 appeals. This excludes 31.6% of inspectors from our sample.[19]

Figure 1.5 shows the identifying variation in our data. Controlling for the vector of assignment variables used by the Planning Inspectorate, the inspector leniency measure ranges from -0.25 to 0.30 with a standard deviation of 0.09. The histogram

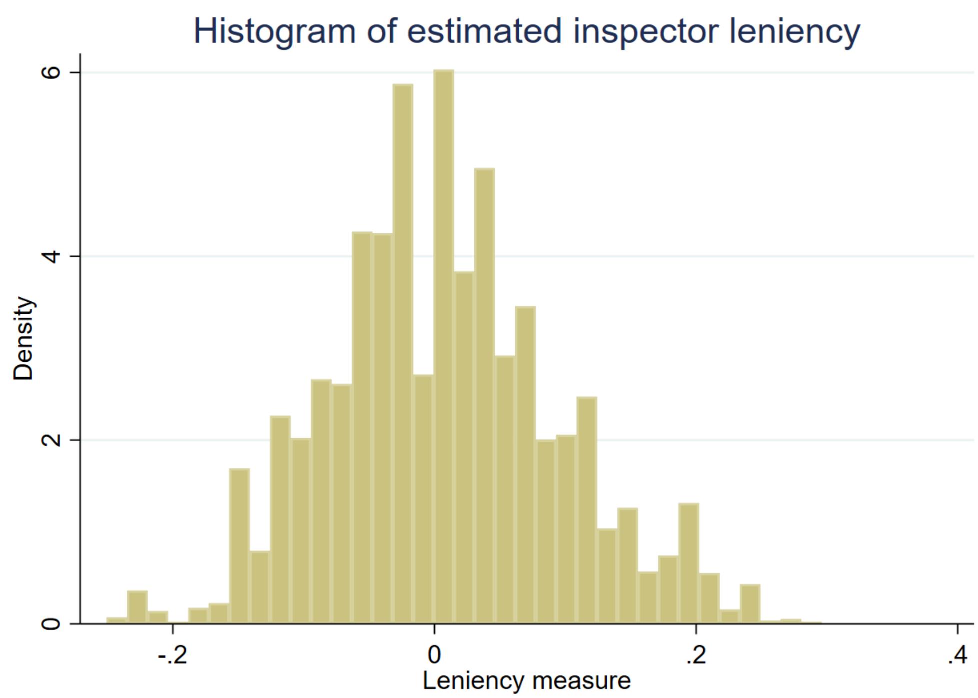
[18] Indeed, the predictive power of leniency on approval, without dropping inspectors, is very poor.

[19] Our results are robust to using various arbitrary cut-offs.

suggests that there is a wide variation in whether an inspector is likely to allow an appeal to be successful.

Finally, we interact the leniency measure with the post appeal dummy. $Leniency_i \times Postappeal_t$ is thus our instrument for $Success_i \times Postappeal_t$.

Figure 1.5: Distribution of Inspector Leniency (Regression Sample)



1.5.4 Validating the IV strategy

Relevance of Instrument

We estimate the following linear probability model to examine the first-stage relationship between inspector leniency and whether an appeal is successful:

$$Success_i \times Postappeal_t = \alpha_j + \delta_t + \rho Leniency_i \times Postappeal_t + \theta' X_{ijt} + \upsilon_{ijt} \quad (1.2)$$

where X_{ijt} is a vector of control variables that is the same as in equation 1.1.

Table 1.7 presents the first-stage results. The estimates are highly significant, suggesting that being assigned to an inspector who is 10 percentage points more lenient increases the probability of the appeal being successful by around 6.6

percentage points. Furthermore, the first-stage has a KP F-statistic value of 24, suggesting that we can reject the null hypothesis of weak instruments.[20] Table 1.8 shows that we can reject the null hypothesis of weak instruments even if the data is aggregated at the appeal-year level.

Table 1.7: First-stage estimates

	(1)	(2)
VARIABLES	Success*Postappeal	
Leniency*Postappeal	0.664***	0.664***
	(0.135)	(0.135)
Observations	1,176,342	1,176,164
Year-by-Month FE	✓	✓
Appeal FE	✓	
Appeal X Postcode Sector FE		✓
Controls	✓	✓
KP F-stat	24.04	24.07

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Table 1.8: First-stage estimates (aggregated at appeals-year level)

	(1)	(2)
VARIABLES	Success*Postappeal	
Leniency*Postappeal	0.549***	0.555***
	(0.112)	(0.115)
Observations	20,683	20,116
Year FE	✓	✓
Appeal FE	✓	✓
Postcode Sector FE		✓
Controls	✓	✓
KP F-stat	23.89	23.23

Notes: Standard errors clustered at outward code level. Variables are aggregated as averages.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Validity of Instrument

Conditional Independence. In order for inspector leniency to be a valid instrument, the assignment of appeals to inspectors must not be correlated with variables which also affect the outcome variable (housing prices). We compiled the characteristics of the local planning authority and of property transactions that were within 1km of **future** appeals and run a balancing test. This test checks if historical

[20]The results presented in Table 1.7 include all controls. In Table C2, we show how the coefficient estimate changes when we move from a specification with just the basic fixed effects to one where controls are added.

characteristics of the local planning authority and the type of properties transacted nearby can predict the leniency of the inspector assigned to an appeal.

Table 1.9 shows that past transaction characteristics of properties within 1km of an appeal are jointly not predictive of inspector leniency. Although the coefficients on two of the potential energy efficiency variables are significant at the 5% level, all the other estimates are close to zero and statistically insignificant at the 5% level. Notably, the characteristics of the local planning authority, such as its acceptance rates of development plans, its mean gross weekly wages and the interaction of these two variables are close to zero and statistically insignificant at the 5% level. More importantly, the variables are not jointly significant with a p-value of 0.176. We also aggregated these variables to the appeals level to check if this conclusion is robust to different levels of aggregation. Table 1.10 shows that although the coefficients on several of the potential energy efficiency variables are significant at the 5% level, the variables are still not jointly significant with a p-value of 0.278. These estimates provides empirical support that conditional on the Planning Inspectorate's assignment rules, inspectors are randomly assigned to cases.

The random assignment of inspectors (conditional on the Planning Inspectorate's assignment criteria) gives us consistent estimates of the reduced form effect of inspector leniency on housing prices. However, interpreting the IV estimates as the causal effects of a successful appeal on housing prices requires two further assumptions.

Table 1.9: Balancing test (transactions within 1km of future appeals)

	(1)
VARIABLES	Inspector Leniency
ln(price)	0.002
	(0.004)
ln(floor area)	-0.001
	(0.003)
3 rooms	0.001
	(0.002)
4 rooms	0.000
	(0.002)
5 rooms	0.001
	(0.002)
6 rooms or more	-0.000
	(0.002)
Current energy B	-0.025
	(0.026)
Current energy C	-0.028
	(0.026)
Current energy D	-0.029
	(0.026)
Current energy E	-0.028
	(0.026)
Current energy F	-0.028
	(0.026)
Current energy G	-0.028
	(0.026)
Potential energy B	-0.005**
	(0.002)
Potential energy C	-0.006**
	(0.002)
Potential energy D	-0.005*
	(0.003)
Potential energy E	-0.004
	(0.003)
Potential energy F	-0.005*
	(0.003)
Potential energy G	-0.003
	(0.004)
Detached house	0.002
	(0.004)
Semi-detached house	-0.002
	(0.002)
Flat	0.007
	(0.006)
Leasehold property	-0.007
	(0.005)
LPA acceptance rate	0.010
	(0.200)
LPA log(mean gross weekly wages)	-0.000
	(0.000)
LPA acceptance rate*log(mean gross weekly wages)	-0.000
	(0.000)
Observations	573,635
Adjusted R-squared	0.003
Year FE	✓
F-stat for joint test	1.259
p-value for joint test	0.176

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Table 1.10: Balancing test (aggregated to appeals level)

VARIABLES	(1) Inspector Leniency
ln(price)	0.002
	(0.006)
ln(floor area)	0.000
	(0.014)
3 rooms	0.006
	(0.024)
4 rooms	0.013
	(0.027)
5 rooms	0.017
	(0.025)
6 rooms or more	0.003
	(0.026)
Current energy B	0.093
	(0.348)
Current energy C	0.098
	(0.345)
Current energy D	0.102
	(0.344)
Current energy E	0.106
	(0.344)
Current energy F	0.128
	(0.346)
Current energy G	0.107
	(0.347)
Potential energy B	-0.091**
	(0.037)
Potential energy C	-0.115***
	(0.037)
Potential energy D	-0.117***
	(0.039)
Potential energy E	-0.136***
	(0.041)
Potential energy F	-0.156***
	(0.055)
Potential energy G	-0.070
	(0.058)
Detached house	0.001
	(0.012)
Semi-detached house	-0.018
	(0.014)
Flat	0.025
	(0.027)
Leasehold property	-0.012
	(0.022)
LPA acceptance rate	0.010
	(0.108)
LPA log(mean gross weekly wages)	-0.033
	(0.084)
LPA acceptance rate*log(mean gross weekly wages)	0.000
	(0.000)
Observations	2,912
Adjusted R-squared	0.000
F-stat for joint test	1.150
p-value for joint test	0.278

Notes: Standard errors clustered at outward code level. Variables are aggregated as averages.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Exclusion Restriction. This restriction requires that inspector leniency affects housing prices (outcome variable) only through its effect on whether an appeal is successful. Although there is no direct way to test the exclusion restriction, the fact that inspectors are randomly assigned to cases lends support to the exclusion restriction.

Monotonicity. If the causal effect of a successful appeal is constant across all cases, then the instrument only needs to satisfy the exclusion assumption. However, with heterogeneous effects, monotonicity must also be assumed. Monotonicity gives the IV estimate a local average treatment effect interpretation – the average causal effect among the subgroup of cases that would have received a different appeal decision had the case been assigned to a different inspector. In our setting, the monotonicity assumption requires that appeals that are ruled successful by a strict inspector would also be ruled successful by a lenient inspector. One testable implication of the monotonicity assumption is that the first-stage estimates should be non-negative for any subsample. To test this, we split the sample into subsamples based on (i) geographic regions; (ii) year of appeal decision; and (iii) number of dwellings in the development. Table 1.11 shows that for all of these subsamples bar one, the first-stage estimates are positive. For the region, West Midlands (region 11 in table 1.11), which accounts for 9% of appeals, the coefficient is negative but not statistically significant. Excluding this region does not change our main results, so we do not think this is strong evidence that the monotonicity assumption is violated.

Table 1.11: Testing monotonicity assumption

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
VARIABLES							Success*Postappeal						
Leniency*Postappeal	1.098***	0.390	1.147***	1.448**	0.966**	0.399	0.804**	-0.660	0.528	1.049***	0.368*	1.001***	0.379**
	(0.418)	(0.416)	(0.291)	(0.673)	(0.492)	(0.273)	(0.340)	(0.446)	(0.447)	(0.191)	(0.189)	(0.201)	(0.184)
Observations	102,782	136,421	236,083	21,726	99,084	268,379	156,215	81,220	74,007	474,035	702,307	578,132	598,210
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Appeal FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
KP F-stat	6.878	0.878	15.49	4.635	3.856	2.135	5.583	2.188	1.394	30.14	3.793	24.75	4.250
Region	1	2	3	4	5	8	9	11	12				
Decision year										2012-2015	2016-2018		
No. of dwellings												10-25	>25

Notes: Standard errors clustered at outward code level.

Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1

1.6 Results

Before presenting the DiD (OLS) and IV DiD estimates formally, we show in Figure 1.6 top panel, the conditional mean of the probability of a successful appeal by inspector leniency, and in the bottom panel the conditional mean log price by inspector leniency. This allows us to visually inspect the underlying variation in our data. The top panel is akin to a dynamic first-stage regression with controls. The only difference is that for presentation purposes, we plot inspector leniency in terms of being above or below the median. Similarly, the bottom panel is akin to a dynamic reduced-form regression with controls. The top panel reveals that inspector leniency has a high predictive power as to whether an appeal will be successful. The bottom panel shows us that prices of resale transactions do indeed diverge based on whether a lenient or strict inspector is assigned to a nearby appeal.

Figure 1.6: Conditional dynamic first-stage and reduced form by inspector leniency

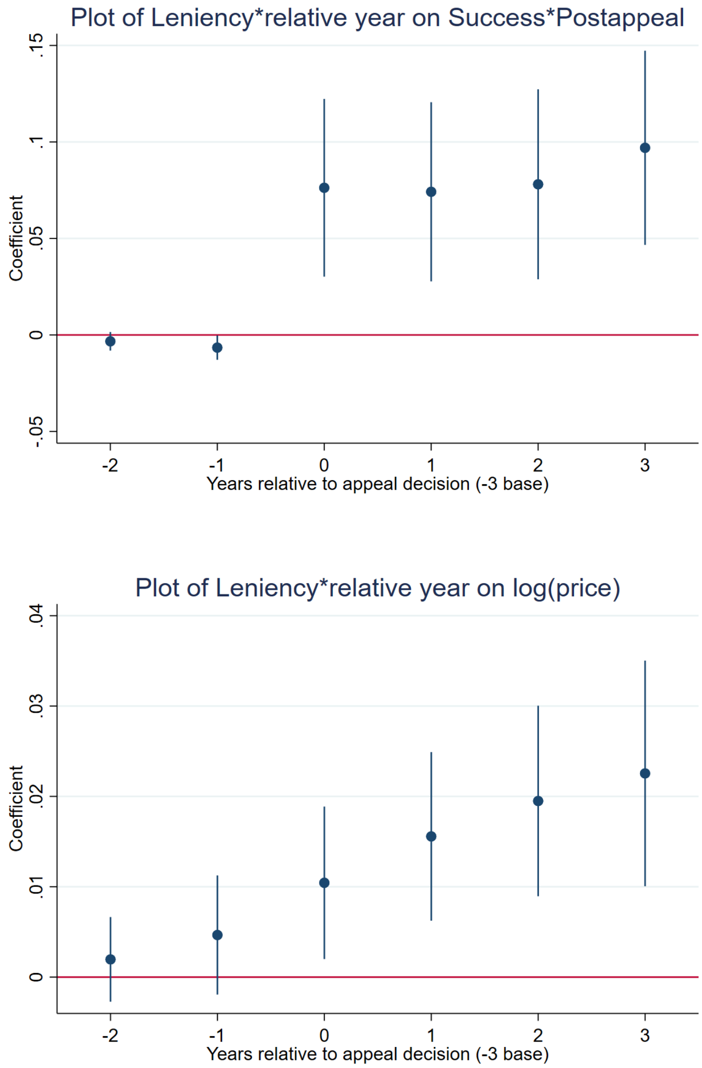


Table 1.12 formally presents the DiD (OLS) and IV DiD estimates of the effect of overturning the local authority's decision on housing prices. Columns 1 to 3 report OLS estimates with all the controls. In the latter column we also include appeal-by-postcode sector FE instead of appeals FE as well as a local planning authority-by-year FE. The OLS estimates suggest that a successful appeal has a small negative effect of around 1% on housing prices (compared to an unsuccessful appeal). The 95% confidence interval is able to rule out an impact more negative than -1.8%. The rejection rate, of major developments, for all local authorities in the UK was about 20% from 2012 to 2017. This suggests that local authorities

were rejecting only the most egregious developments. Hence, it is surprising that we can rule out a large negative impact on prices in the neighbourhood.

The IV estimates are presented in columns 4 to 6. They suggest that, instead of depressing prices, overturning the local authorities' decision actually increased the value of properties in the neighbourhood by around 6%. The coefficient is significant at the 10% level, and its 95% confidence interval is able to reject an impact more negative than -0.5%. Similar to the OLS, we are not finding that developments that are rejected have a large detrimental impact on the neighbourhood.[21]

For both the OLS and IV, inclusion of local planning authority-by-year FE attenuates the coefficients towards 0, which support the conclusion that these developments do not have a large detrimental impact. But we treat these regressions as checks on our preferred specifications (i.e., columns 1, 2, 4 and 5 without local planning authority-by-year FE) because the local planning authority-by-year FE absorbs some of the valid variation from Success*Postappeals (e.g., appeals that are the only ones from their respective local planning authorities in our sample).

Table 1.12: Effect of overturning the Local Authority's decision

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	IV	IV	IV
VARIABLES	ln(price)	ln(price)	ln(price)	ln(price)	ln(price)	ln(price)
Success*Postappeal	-0.012***	-0.012***	-0.001	0.054	0.056	-0.002
	(0.004)	(0.004)	(0.002)	(0.034)	(0.034)	(0.014)
Observations	1,177,861	1,177,683	1,177,647	1,176,342	1,176,164	1,176,128
R-squared	0.842	0.855	0.860	0.623	0.623	0.632
Year-by-Month FE	✓	✓	✓	✓	✓	✓
Appeal FE	✓	✓	✓	✓	✓	✓
Postcode sector FE		✓	✓		✓	✓
LPA X Year FE			✓			✓

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

The 95% confidence intervals of the OLS and IV estimates overlap, so we cannot reject the null that the IV and OLS are different. This is mostly because the IV estimates are not estimated very precisely. However, the direction of the estimates are different and the magnitude of the IV estimate is a substantial positive number. We think this is due to the IV estimating a LATE instead of the ATE. The OLS

[21] The results presented in Table 1.12 include all controls. In Tables C3 (DiD) and C4 (IV DiD), we show how the coefficient estimate changes when we move from a specification with just the basic fixed effects to one where controls are added.

estimates tell us that the ATE of overturning local authorities' decision is a small negative impact on the neighbourhood. But for a subsample of these projects, those that are quite marginal and hence are very susceptible to the leniency of inspectors, the impact may be positive.

We interpret projects that were successfully appealed as marginal development projects that managed to overcome local NIMBY-ism. Our finding that there is little evidence that such developments reduce housing prices in the vicinity suggests (i) local homeowners could be misinformed about the price impact of marginal local development projects; (ii) house prices are not the primary motive of NIMBY residents' opposition of local residential developments; or (iii) opposition against development projects being driven by **immediate** neighbours (who would be affected by construction noise or hampered views) rather than by the majority of neighbours in the vicinity of a development project (who may benefit from amenities associated with the marginal development project).

To further support our interpretation that NIMBYs are driving the initial rejection of the developments in our sample, we look at comments lodged by residents close to these projects. We randomly sampled 68 projects in our sample (rejected projects) from London. For each project, we randomly select a project near to it (within 1 to 2km) as a comparison that was accepted by the local authority.[22] The mean number of comments lodged for rejected projects is 3.07 as compared to 0.72 for accepted projects (p-value of 0.0375). To the extent that comments are used to discourage projects, this is evidence for NIMBY-ism.

1.6.1 Heterogeneous effects

We attempt to decompose the impact by distance from the appeal site. Columns 1 and 3 of Table 1.13 show the estimated impact when we restrict the sample to properties within 500m of the appeal site, while Columns 2 and 4 of Table 1.13 show the estimated impact when we restrict the sample to properties 500m to 1km from the appeal site. Both the OLS and IV results suggest little heterogeneity in the impact by distance from the appeal site.[23]

[22]We did not look at projects within 1km of an appeal site because it may be affected by the appeal.

[23]The results presented in Table 1.13 include all controls. In Tables C5 (500m sample) and C6 (500m to 1km sample), we show how the coefficient estimate changes when we move from a specification with just the basic fixed effects to one where controls are added.

Table 1.13: Effect by distance from appeal site

VARIABLES	(1) OLS ln(price)	(2) OLS ln(price)	(3) IV ln(price)	(4) IV ln(price)
Success*Postappeal	-0.009**	-0.010**	0.051	0.073*
	(0.005)	(0.004)	(0.031)	(0.038)
Observations	360,933	816,622	360,563	815,474
R-squared	0.857	0.861	0.615	0.619
Time FE	✓	✓	✓	✓
Appeal X Postcode sector FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Sample	<500m	500m to 1km	<500m	500m to 1km
KP F-stat			28.96	19.85

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Table 1.14 explores heterogeneity by the supply shock from a successful appeal. We define the supply shock as the number of new dwellings that would be introduced by the appeal, divided by the total number of dwellings within 1km of the appeal site, in the year prior to the appeal decision. Small and large shocks are categorized based on whether they were below or above the median magnitude of supply shocks. For the OLS, there was no significant heterogeneity by the size of the supply shock. For the IV, small supply shocks have a greater positive impact on prices in the neighbourhood than larger supply shocks. This is consistent with a simple demand and supply model, as a smaller supply shock suggests a small movement along the demand curve. However, a positive impact also suggests a shift in the demand curve. This could be because the new developments also bring in new or better amenities to the neighbourhood. In the case of developments with a small supply shock, the demand shifters overwhelmed the supply shock; while for larger developments the two seem to cancel out. This interpretation is for the IV and hence applies only to marginal developments (LATE). Also, the 95% confidence intervals of the two samples overlap, so we cannot reject the null of no differences.[24]

[24] The results presented in Table 1.14 include all controls. In Tables C7 (small shock) and C8 (large shock), we show how the coefficient estimate changes when we move from a specification with just the basic fixed effects to one where controls are added.

Table 1.14: Effect by supply shock

VARIABLES	(1) OLS ln(price)	(2) OLS ln(price)	(3) IV ln(price)	(4) IV ln(price)
Success*Postappeal	-0.009*	-0.007	0.070**	-0.011
	(0.005)	(0.005)	(0.035)	(0.083)
Observations	913,192	264,491	912,619	263,545
R-squared	0.860	0.831	0.618	0.646
Time FE	✓	✓	✓	✓
Appeal X Postcode sector FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Sample	Small supply shock	Large supply shock	Small supply shock	Large supply shock
KP F-stat			24.70	1.838

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1

1.6.2 Mechanisms

One limitation of the data is that we only observe transaction prices. However, the outcome of an appeal may affect the choice to buy or sell a nearby property (selection into transaction). For instance, richer households in the neighbourhood, who have the finances to move quickly, may choose to sell after a successful appeal as they anticipate that the new supply would depress the value of their housing asset. If these properties are also more valuable (e.g., bigger and more luxurious) then our IV result of a price increase from a successful appeal, might be due to this selection into transaction.[25] We check if this is an issue by regressing some characteristics of the transacted properties on the treatment variable ($Success_i * Postappeal_t$). If there is selection into transaction, then we should expect the types of properties being transacted to be different between neighbourhoods with successful and unsuccessful appeals.

Table 1.15 summarises our findings on three property characteristic: (i) floor area; (ii) number of rooms; and (iii) energy rating (a higher rating is more energy efficient). The first two measures size and the third is a proxy for the type of materials used in the property.[26] Both the OLS and the IV regressions cannot reject

[25] We do control for factors related to the characteristics of transacted properties in all our regressions. However, if characteristics are endogenous to appeal decisions then they should be outcome variables instead of controls (bad controls problem).

[26] This is an imperfect measure of "quality" because there may be luxurious building materials that are not energy efficient, and rich households may be willing to pay higher heating costs to maintain them. We recognize this and are simply interested to test if the materials used are different as opposed to better.

the null that there are no differences in transacted properties near successful versus unsuccessful appeals. We interpret this as evidence that selection into transaction is not a major driver of our main results above.

Table 1.15: Possible mechanisms

VARIABLES	(1) OLS ln(Floor area)	(2) OLS No. rooms	(3) OLS Energy rating	(4) IV ln(Floor area)	(5) IV No. rooms	(6) IV Energy rating
Success*Postappeal	0.002	0.004	0.005	0.012	-0.032	0.029
	(0.001)	(0.007)	(0.004)	(0.013)	(0.059)	(0.032)
Observations	1,213,468	1,177,683	1,177,683	1,211,946	1,176,164	1,176,164
R-squared	0.417	0.395	0.479	0.287	0.261	0.400
Time FE	✓	✓	✓	✓	✓	✓
Appeal X Postcode sector FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
KP F-stat				24.19	24.07	24.07

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Is there evidence of demand shifters? To answer this, we collect annual point-of-interest (POI) data by location and test if successful appeals lead to higher counts of these points. We categorize POIs into 3 categories, summarized in Table 1.16. These categories reflect the possible amenities that major developments could introduce: (i) integrated developments often include commercial spaces for local businesses; (ii) population growth in an area may attract more businesses; and (iii) developers may be asked to contribute to local infrastructure by building them (e.g., green spaces) or via a tax.

Table 1.16: POI Categories

Category	POIs included
Local services	Accommodation, Eating and Drinking, Retail
Local economy	Commercial services, Manufacturing and Production
Local infrastructure	Education and Health, Public Infrastructure, Transport

We use the count of POIs in each category as an outcome variable and study if there were more POIs within 1km of a successful appeal site (Table 1.17). Our results are noisy and we cannot make strong conclusions.[27] However, there is suggestive evidence that the count of local services increased after successful appeals (IV regression significant at 10% level).

[27] This is due to the high annual turnover in POIs: businesses open and shut down, bus stops are added etc.,.

Table 1.17: Possible demand shifters

VARIABLES	(1) OLS Local services	(2) OLS Local economy	(3) OLS Local infra	(4) IV Local services	(5) IV Local economy	(6) IV Local infra
Success*Postappeal	-0.399	1.537	0.220	11.419*	6.340	-0.639
	(0.643)	(1.104)	(0.871)	(6.743)	(7.113)	(9.057)
Observations	18,654	18,654	18,654	18,624	18,624	18,624
R-squared	0.995	0.994	0.991	-0.046	-0.002	-0.000
Time FE	✓	✓	✓	✓	✓	✓
Appeal FE	✓	✓	✓	✓	✓	✓
KP F-stat				27.76	27.76	27.76

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

If we disaggregate local services (Table 1.18), we see that the increase is coming from retail shops. This is evidence that the positive impact from marginal appeals (LATE), may be driven by an improvement in local amenities which increased demand for the neighbourhood.

Table 1.18: Demand shifters within local services

VARIABLES	(1) IV Accommodation, Eating and Drinking	(2) IV Retail
Success*Postappeal	1.162	10.258*
	(2.594)	(5.605)
Observations	18,624	18,624
R-squared	-0.001	-0.065
Time FE	✓	✓
Appeal FE	✓	✓
KP F-stat	27.76	27.76

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Robustness Checks

Next, we subject our results to a series of robustness checks. First, we show that our results are robust to controlling for location specific linear year trends as well as clustering our standard errors at the appeal and distance level. Table 1.19 shows the results for three levels of clustering: (i) appeal-by-postcode sector neighbourhood; (ii) postcode sector; (i) 1km grids. The 1km grids are arbitrarily created grids of 1km by 1km across the whole of the UK. Inference on the OLS does not change much across these different clusters. The standard errors for the IV are noisier under the appeals neighbourhood cluster (column 4), and we can no longer reject a null of zero effect at the 10% level. However, the 95% confidence interval still allows us to reject an impact more negative than -1.7%. Therefore, our original inference of no big negative impact still stands. On the other hand, other methods of clustering help improve precision and allow us to reject a null of zero impact.

Table 1.19: Robustness to different clustering

VARIABLES	(1) OLS ln(price)	(2) OLS ln(price)	(3) OLS ln(price)	(4) IV ln(price)	(5) IV ln(price)	(6) IV ln(price)
Success*Postappeal	-0.010**	-0.010***	-0.010***	0.062	0.062**	0.062***
	(0.004)	(0.003)	(0.003)	(0.039)	(0.025)	(0.022)
Observations	1,177,683	1,177,683	1,177,683	1,176,164	1,176,164	1,176,164
R-squared	0.856	0.856	0.856	0.623	0.623	0.623
Time FE	✓	✓	✓	✓	✓	✓
Appeal X Postcode sector FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Clusters	Appeal-Postcode sector	Postcode sector	1km grids	Appeal-Postcode sector	Postcode sector	1km grids
KP F-stat				18.65	42.77	60.39

Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

In Table 1.20 we check whether our results are robust to different appeal neighbourhood and appeal-by-postcode sector specific time trends. The trends we include are estimated off a similar variation as our treatment variable $Success_i * Postappeal_t$ and we should hence expect our coefficients of interests to attenuate towards zero. This is indeed what happens and none of the estimates can reject a null of zero impact. However, the 95% confidence intervals allow us to reject a null of an impact more negative than -1% (OLS) and -1.9% (IV).

Table 1.20: Robustness to different trends

VARIABLES	(1) OLS ln(price)	(2) OLS ln(price)	(3) IV ln(price)	(4) IV ln(price)
Success*Postappeal	-0.004	-0.004	0.026	0.028
	(0.003)	(0.003)	(0.023)	(0.023)
Observations	1,177,861	1,177,683	1,176,342	1,176,164
R-squared	0.846	0.860	0.629	0.631
Time FE	✓	✓	✓	✓
Appeal FE	✓		✓	
Appeal trends	✓		✓	
Appeal X Postcode sector FE		✓		✓
Appeal X Postcode sector trends		✓		✓
Controls	✓	✓	✓	✓
KP F-stat			26.66	26.58

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Finally, we trim off the top and bottom values in our instrument (inspector leniency), to check if our IV results are sensitive to outlier values (Table 1.21). After trimming 1% of extreme values (column 1), the IV estimates remain positive, at 7.2%, and are now significantly different from zero at the 5% level. Trimming 5% of extreme values causes the estimate to remain positive but attenuate towards zero.

Table 1.21: IV results after trimming

VARIABLES	(1) ln(price)	(2) ln(price)
Success*Postappeal	0.072**	0.040
	(0.033)	(0.034)
Observations	1,155,254	1,111,294
R-squared	0.622	0.626
Time FE	✓	✓
Appeal X Sector code FE	✓	✓
Trim	1%	5%
KP F-stat	25.21	17.93

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

1.7 Conclusion

In this paper, we study the impact of overturning local authorities' rejection of major residential developments in the UK. We interpret projects that were successfully appealed as marginal development projects that managed to overcome local NIMBY-ism. Our finding that there is little evidence that such developments reduce housing prices in the vicinity suggests (i) local homeowners could be misinformed about the price impact of marginal local development projects; (ii) house prices are not the primary motive of NIMBY residents' opposition of local residential developments; or (iii) opposition against development projects being driven by **immediate** neighbours (who would be affected by construction noise or hampered views) rather than by the majority of neighbours in the vicinity of a development project (who may benefit from amenities associated with the marginal development project).

Chapter 2

The Great Fire of London and Urban Development in 17th Century London

2.1 Introduction

The development of cities often involves the rejuvenation or replacement of existing structures (see for instance Henderson et al. (2021)). However, history, in the form of the sunk cost of existing durable structures, often serves as an impediment to urban development. In every period, property owners face a trade-off between receiving rent from the existing building or incurring a cost to tear down the building and rebuilding it. As a result, they often wait long periods of time for their building to depreciate before embarking on upgrading. Furthermore, without some gain to being the first to upgrade their property, property owners may rationally wait for others to upgrade first. In theory, by reducing the opportunity cost of waiting to rebuild to zero, disasters (such as fires (Hornbeck and Keniston (2017)) or wars (Dericks and Koster (2021))) can eliminate these frictions and bring about higher quality structures. In addition, the simultaneous rebuilding after a disaster would allow property owners to experience stronger cross-building spillovers. As described by Hornbeck and Keniston (2017), this "virtuous circle" of cross-plot externalities result in building upgrades encouraging further upgrades of nearby buildings.

Nevertheless, the opportunity cost of waiting to rebuild falling to zero coupled with the prospects of stronger cross-building spillovers, are not sufficient to guarantee higher quality buildings. This is because individuals' investment decisions also depend on their expectations of what others will do. For example, if a city (or more generally, an area) is growing, then individuals will expect other individuals to build higher quality buildings. By contrast, if the expectations are that the area is in decline, then individuals may not even rebuild or may invest at a lower quality since they expect other individuals to do the same.

In this paper, we examine both of these issues using the 1666 Great Fire of London as a natural experiment. Our research questions are as follows. First, we examine whether the Fire was able to free parishes[1] within London from the constraints of their existing durable structures and move them to a new equilibrium involving higher quality structures. In line with the historical context, we define the quality of structures based on the number of hearths in the property (i.e., number of hearths per housing unit and not per building). While the first research question that we examine is similar to the papers on the 1872 Boston fire by Hornbeck and Keniston (2017) and the 1906 San Francisco fire by Siodla (2015), our second research question departs from these paper. In particular, we study what anchors individuals' expectations of what others will do and how this can consequently facilitate the development of cities. We find evidence that legal rulings arising from the Fire Court – a court specially set up by the English Parliament to hear rebuilding disputes – were able to anchor expectations and in so doing, helped to facilitate the development of London.

For the first part of the paper, to examine whether the removal of development frictions through the Fire resulted in higher quality structures being rebuilt, we employ a difference-in-differences (DiD) strategy. The DiD strategy exploits both the cross-sectional and time-series variations arising from the Fire. The time-series variation comes from the timing before and after the Fire which was exogenous. The cross-sectional variation arises because different parishes in London were affected differently by the Fire. For example, some parishes were burned whereas some parishes did not experience any damage from the Fire at all. A null effect from our regression would suggest that there were no frictions to upgrading before the Fire – the quality of properties was optimal. By contrast, a positive effect suggests the presence of upgrading frictions which the Fire effectively removed.

[1]Parishes were administrative units within a city that played a role in both civil and ecclesiastical matters.

Using our DiD strategy, we find that a few years after the Fire, burned parishes experienced a highly statistically significant increase in the number of hearths per property compared to unburned parishes. In addition, the effect varied with the level of damage. Parishes which were more badly damaged saw a highly statistically significant increase in the number of hearths per property compared to parishes which were less damaged. Finally, the effect was biggest for parishes whose neighbouring parishes were all burned compared to parishes whose neighbouring parishes were not all burned.

The result from the first part of the paper suggests that individuals had positive expectations that others will be rebuilding at a high quality. Nevertheless, it does not tell us what is driving these expectations. Therefore, in the second part of the paper, we examine the role of legal rulings in driving expectations. In 17th century England, tenants were legally obliged to rebuild in the event of any disasters which damaged the property, even if it was not their fault. However, the Fire took place amidst a plague and war – an unprecedented joint occurrence of events. To expedite the rebuilding of London, the English Parliament established the Fire Court.

The second part of the paper begins with a model that shows that legal rulings affect expectations because they affect the bargaining between landlords and tenants who do not go to Court. This is because their outside options are based on the Fire Court's initial rulings. For our empirical strategy, we turn once again to a DiD strategy. Just as before, the time-series variation comes from the timing before and after the Fire. However, the cross-sectional variation now arises because different parishes experienced different Fire Court rulings. For example, some parishes saw a disproportionate number of initial cases where the Fire Court voided the existing contracts between the landlord and tenant and consequently assigned the rebuilding to the landlord. This is what we refer to as pragmatic rulings. Voiding the contract means that both the landlord and tenant surrender their contracts. This allows both parties to negotiate a new contract with each other or other parties.

Our regression results show that parishes with a greater share of pragmatic rulings had more hearths per property compared to parishes where there was a lower share of cases with pragmatic rulings. In addition, because only a very small proportion of properties in each parish went to the Fire Court, our results suggest that the rulings of these few cases had an outsized effect on the quality of other buildings in the parish.[2]

While we have included a number of time varying parish-level controls in our regression, a threat to identification in the DiD strategy is that we might not have controlled for all possible confounders. As a result, the change in the number of hearths may be related to changes in parish level characteristics that are not due to the Fire Court rulings – a violation of the parallel trend assumption. Therefore, we augment our DiD strategy with an instrumental variable (IV) strategy.

Our IV DiD strategy exploits the fact that at the parish level, Fire Court judging panels that have different political alignments (i.e., whether they were predominantly Royalists or Parliamentarians) were assigned to the cases. The 1666 Great Fire took place in the midst of the Second Dutch War (1665-1667) and the Great Plague which began in 1665. King Charles II was relying on loans from London and its wealthiest citizens to finance the war. The destruction of the customs house, wharves and more than 13,000 buildings caused a significant drop in royal revenue. The King had a vested interest for London to be rebuilt quickly. Therefore, judging panels that consisted predominantly of Royalists (i.e., more aligned with the King) were more likely to decree pragmatic rulings so as to facilitate the rebuilding of London. As a result, we can use the composition of the judging panels as an instrument for the share of cases in the parish that had pragmatic rulings. This gives us exogenous variations in legal rulings for each parish.

We find that the results from our IV analysis re-affirm our DiD results – legal rulings can indeed anchor expectations and help to facilitate the rebuilding process. To the best of our knowledge, while there are theoretical papers such as Cooter (1998), Basu (2000), McAdams (2000), McAdams (2005), Myerson (2004) and Hadfield and Weingast (2012) that examine how legal institutions can affect expectations and hence the behaviour of individuals, there are relatively fewer empirical papers that provide causal evidence of this.

In examining how expectations affect the behaviour of economic agents, our paper

[2] Based on the initial cases, the average proportion of properties in each parish that went to the Fire Court was 6%.

is related to Krugman (1991) and Rauch (1993). In addition, our paper is related to how cities recover from major shocks and whether they move to a new equilibrium. Beginning with Davis and Weinstein (2002), there has been an extensive literature that examines whether long-run city size is robust to temporary shocks. These shocks include wars and bombing (Davis and Weinstein (2002) and Miguel and Roland (2011)), natural or man-made disasters (Siodla (2015) and Hornbeck and Keniston (2017)), political events (Redding et al. (2011) and Michaels and Rauch (2018)), technology (Bleakley and Lin (2012)) and even diseases (Jedwab et al. (2019)). Our paper provides evidence of how the Great Fire of London freed London from the constraints of history and enabled it to move to a new equilibrium with more hearths per property.

By addressing how legal rulings contribute to the development of cities, our paper is related to the literature on the economic consequences of legal origins. This literature shows how legal origins affect particular legal rules and these in turn affect economic outcomes such as growth, financial development, property rights and contract enforcement. Examples of these studies include Acemoglu et al. (2001), Djankov et al. (2003), La Porta et al. (2004), La Porta et al. (2008) and Dell (2010). In using judging panels that consisted predominantly of Royalist as our instrument in our IV analysis, our paper is also related to North and Weingast (1989), Acemoglu et al. (2005), Jha (2015) and Angelucci et al. (2017). These papers examine the tensions between Parliamentarians and Royalists during various times in English history (e.g., the English Civil War (1642-1651) and the Glorious Revolution (1688)) and show how these affected the development of institutions that facilitated growth in England.

Our paper also contributes to the literation on incomplete contracts and relationship-specific investment. Theoretical papers such as Grout (1984), Tirole (1986) and Hart and Moore (1988) show that parties will invest at less than the socially optimal levels, leading to what is referred to as the hold-up problem. At the same time, Che and Hausch (1999) note that with cooperative investment, if parties have difficulty committing not to renegotiate, then contracting has no value and parties may do better by abandoning contracting altogether in favour of ex post negotiation.

Despite this vast theoretical literature on contracts, there have been very few causal empirical studies that examine the effect of contracts under such settings. One reason behind the dearth of such studies is because contracts are endogenous. Two papers that are related to our paper are Jacoby and Mansuri (2008) which examines the effect of contracts in the context of farming, as well as Card et al. (2014) which looks at this question in the context of wage bargaining. Nonetheless, both these papers do not look at urban development or consider the asymmetrical nature of tenant-landlord relationships.

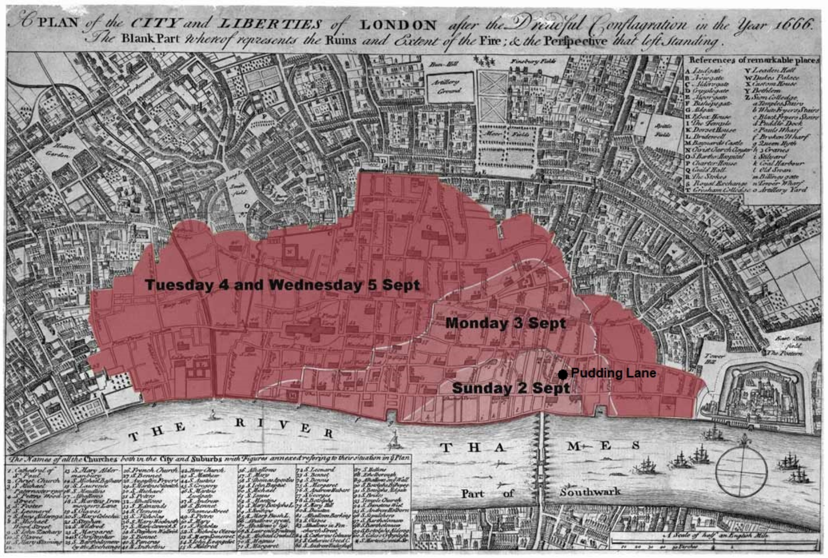
Finally, our paper is also related to the historical literature on the impact of the Great Fire of London. Field (2008) notes that the 1666 Great Fire of London is such an iconic moment in the history of London that the contemporary media frequently used the phrase "The Second Great Fire" to describe the London Blitz during World War II. While the 1666 Fire has been extensively studied by historians (e.g., Reddaway (1940), Porter (1996) and Field (2017)) and even legal scholars (e.g., Tidmarsh (2016)), our paper contributes to this largely qualitative literature by providing a quantitative analysis on the impact of the Great Fire of London.

The rest of the paper proceeds as follows. Section 2 presents the historical background of the 1666 Great Fire of London. Section 3 discusses the novel data sources that we use for our analysis. Section 4 examines the effect that the Fire had on the quality of properties that were rebuilt. Section 5 presents our main contribution which is that legal rulings anchored individuals' expectations of what others will do and this consequently facilitated the development of parishes within London. We conclude in Section 6.

2.2 Historical Background: The 1666 Great Fire of London

This section draws extensively from Reddaway (1940), Porter (1996), Field (2008), Tidmarsh (2016) and Field (2017). The Great Fire of London began on September 2, 1666, in a bakery on Pudding Lane in the City of London (see figure 2.1). The City of London covers an area of 2.8 km^2 or 1.1 $miles^2$ within London and was home to about one sixth of London's inhabitants. The structure of the city made it easy for the Fire to spread. Streets, lanes and alleys were narrow and buildings were made from timber. In addition, the upper floors of houses often cantilevered over the pathways below. This meant that the top floors on one side of the street nearly touched those on the other side, making it easy for the Fire to spread. The Fire lasted for three and a half days and destroyed approximately 13,200 buildings in the City of London. An estimated 70,000 out of 80,000 inhabitants living in the City of London lost their homes. Figure 2.1 shows the geographical spread of the Fire over the three and a half days.

Figure 2.1: Spread of the Fire



Source: of London (2011); with authors' edits to include the location of Pudding Lane

Tidmarsh (2016) notes that despite the urgency to rebuild London, there were significant challenges. At the time of the Fire, the institution of fire insurance had not yet developed. Instead, the common practice was that leases had a covenant that obligated the tenant, regardless of whether the tenant was at fault, to repair or rebuild the premises in the event of disasters or wars. This created substantial challenges for both the tenants and landlords. For the tenants, there was the issue of fairness in whether they should bear the full cost of rebuilding. Many tenants could not afford to rebuild. Moreover, tenants who had a short time left on their lease had little incentive to rebuild. As for the landlords, there were long delays and huge cost in bringing disputes to the common-law courts. Even if the case was brought before the common-law courts, the powers of these courts were constrained by the existing tenancy agreements. As a result, the judges could not calibrate or void the existing contracts to achieve the best incentives for the parties to rebuild. Furthermore, due to the existing tenancy agreements, landlords could not prematurely re-enter the leased premises in order to facilitate reconstruction.

In order to expedite the rebuilding of London, the English Parliament established the Fire Court to adjudicate between landlords and tenants as to who would bear the burden of rebuilding (Fire of London Disputes Act 1666). The bill was passed in the House of Lords on January 23, 1667. A few days later, on January 31, 1667, the House of Commons assented to the bill.[3] Tidmarsh (2016) notes that the Fire Court heard a total of 1,585 cases. Some cases involved more than one property so the 1,585 cases understate the extent of the Court's work. As mandated by the Fire Court legislation, each case was heard by a panel of at least three judges. The judges were given the power to void existing contracts and decide the details of the new contracts (e.g., who rebuilds, new rent and length of the tenancy agreement, etc.). The typical process when a case is brought to the Fire Court is that the judges would first try to mediate and get the tenant and landlord to come to an agreement. In the event that the parties are unable to come to an agreement, the Court will then make a ruling which is legally binding.[4]

In concluding our discussion about the historical background of the Great Fire of 1666, we would like to highlight that there were previously other fires in London that also resulted in substantial damage. For example, Richardson (2001) notes that the Great Fire of 1133 damaged St Paul's, St Bride's, London Bridge and properties as far east as Aldgate. Another example was the Great Fire of 1212 which began at Southwark, destroyed the church and spread to London Bridge. Legal issues surrounding the responsibility of the tenant to rebuild would have also existed back then. Why then was the Fire Court only set up after the 1666 Fire

[3]The year of the enactment of the statute was listed as 1666 even though the bill was passed in the House of Lords on January 23, 1667. This is because based on the calendar that was used during that era, the new year began on Lady Day (March 25).

[4]Reddaway (1940), Porter (1996) and Tidmarsh (2016) note that besides setting up the Fire Court, the Parliament of England also put in place other legislation and measures to facilitate the rebuilding of the city. There were new building regulations to limit damage from subsequent fires (buildings must be made of brick, be of a minimum size, not exceed a certain height and must not cantilever over the streets). To determine boundaries and settle disputes among neighbours, a survey system was put in place. Since properties were taken for public purposes (e.g., widened streets), there was a formal channel to value property. To finance the reconstruction of public buildings, a tax on coal was introduced. To ensure that property owners rebuilt within a reasonable time, sanctions were meted out if this was not done. There were provisions requiring owners to share rebuilding costs that benefited multiple properties (e.g., party walls). Regulations on the price and quality of raw materials used for the rebuilding were implemented. Incentives were given to encourage skilled craftsmen to come to London to help with the rebuilding.

and not earlier? The existing literature is surprisingly silent on this.

One reason could be that while previous Great Fires caused substantial damage, the damage to property from the 1666 Fire was arguably the greatest (see for example Garrioch (2016)). London had grown substantially since the 12th and 13th century. Therefore, even if the entire city was almost destroyed due to the 1133 Fire, by 1666 the size of the city would have been far larger. Nevertheless, due to the lack of data (most of the evidence is qualitative), it remains debatable whether the damage from the 1666 Fire was the greatest. For example, Garrioch (2016) notes that about 3,000 people died in the 1212 Fire, far more than the eight people that was estimated to have died due to the 1666 Fire.[5]

Therefore, we think that the main reason was due to the joint occurrence of war, plague and Fire – a combination of events that was absent in the previous Great Fires. The Great Plague which began in 1665 resulted in the death of almost a quarter of London's population within 18 months. This means that there was now a huge excess supply of vacant properties which vastly increased the bargaining power of tenants. The King could wait for the landlords and tenants to reach a bargained outcome. For example, whether the landlord contributes to the rebuilding or changes the terms of the tenancy contract even though by law the tenant has to rebuild. However, given the ongoing Second Dutch War (1665 to 1667), King Charles II simply could not wait for this to play out. Tidmarsh (2016) argues that the King was relying on loans and taxes from London and its wealthiest citizens to finance the war. The destruction of the customs house, wharves and buildings caused a significant drop in royal revenue from custom and hearth taxes. The Fire Court was therefore a way to expedite reaching a somewhat equitable outcome. It gives the landlord class some portion of what prior precedent would suggest but it also tilts things sufficiently toward tenants to mirror the shift in bargaining power owing to the plague.[6]

[5] Despite the destruction, the largest estimate of deaths directly due to the Fire was eight. This is a shockingly small number and historians such as Field (2017) have offered a number of explanations. First, the incineration of bodies in the Fire meant that corpses could not be recovered and so the death records are underestimates. Second, the Fire took place over three and a half days. This gave sufficient time for people to evacuate. Third, historians postulate that the relatively tight-knit nature of the neighbourhoods meant that there was help and assistance for the vulnerable.

[6] We would like to thank Don Davis for helping us to sharpen this argument.

2.3 Data

Urban investment. In line with how the value of a property was assessed in 17th century London, we measure quality by the number of hearths that are in each property before and after the Fire. The property we refer to is the housing unit within buildings and not the entire building, this is because the hearth tax was accessed at the housing unit level and not the building level. This information is available from the historical manuscripts of the hearth tax assessment records that are held at The National Archives, United Kingdom.

According to the University of Roehampton, Centre for Hearth Tax Research,[7] the hearth tax was introduced in England and Wales in 1662 to provide a regular source of income for King Charles II who was the newly restored monarch. Parliament had estimated that the King required an annual income of £1.2 million. However, by 1661, there was a shortfall of £300,000 and it was hoped that the hearth tax would make up for this. The hearth tax was essentially a property tax on dwellings graded according to the number of fireplaces in the property. The tax was paid in two equal instalments at Michaelmas (September, 29) and Lady Day (March, 25) by the occupier. If the property was vacant, the landlord paid the tax. In order to administer the tax, a list of householders was compiled and this formed the hearth tax assessment records.

Our pre-Fire hearth data comes from two sources. First, we use the full records from the 1666 London and Middlesex hearth tax, along with portions of the 1663 and 1664 documents that have been cleaned and digitized by the London Hearth Tax project.[8] Since the hearth tax was collected twice a year in March and September, the 1666 records are based on the March collection which took place before the Fire in September. Second, we supplement this with the 1664/1665 Southwark hearth tax records that come from the assessment for Surrey. The Southwark data was manually transcribed by Field (2008) for his history PhD

[7] https://www.roehampton.ac.uk/research-centres/centre-for-hearth-tax-research/

[8] In June 2007, the London Hearth Tax project was formed to systematically analyse and digitize the hearth tax records. The project united the expertise of the British Academy Hearth Tax Project, the Centre for Hearth Tax Research (University of Roehampton), Birkbeck College (University of London), and the Centre for Metropolitan History (Institute of Historical Research). In 2011, the full records from the 1666 London and Middlesex hearth tax, along with portions of the 1663 and 1664 documents, were published electronically via British History Online at https://www.british-history.ac.uk/london-hearth-tax/london-mddx/1666.

thesis.[9]

As for the post-Fire hearth data, we rely on the records from the 1675 London and Middlesex hearth tax records as well as the 1673 Surrey (Southwark) hearth tax records. These data were also manually transcribed by Field (2008).[10]

One issue of the data is that most properties may not have been rebuilt by 1675 and hence our sample of the tax records of properties may be biased. For instance, if higher quality buildings were rebuilt first, then any differences in quality we observe pre- and post-fire could be driven by the possibility that higher quality properties were rebuilt first, and lower quality ones later, but we were not able to observe properties rebuilt later (i.e., after 1675). However, this is not likely to be a major issue, because the Act for Rebuilding the City of London 1667 encouraged the rapid rebuilding of the city by requiring properties to be rebuilt within 3 years after the Fire. Indeed, maps of London by John Ogliby and William Morgan show that the city had largely been reconstructed by 1667 [11]. While the possible of bias remains, the rapid rebuilding of London suggests it is not likely to be a major issue.

The unit of geography for our analysis is at the parish level. Due to the differences in the scope and range of the hearth tax assessments, some parishes only appear in the pre-Fire records while others only appear in the post-Fire records. In our regressions, we only use data from the parishes that appear in both the pre- and post-Fire records. Table 2.1 shows the summary statistics of the hearth tax data which we use in our regressions.

Table 2.1: Summary statistics (Hearth Tax data)

	Mean	SD	Min	Max	N
Number of hearths (pre-Fire)	3.83	3.79	0	193	44,724
Number of hearths (post-Fire)	4.33	3.36	0	135	35,006
Number of parishes	.	.	.	.	70

[9] While the data from Southwark is undated, Field (2008) notes that they are most certainly from the period between 1664 and 1665.

[10] Although the London and parts of the Middlesex hearth tax records were presented to Parliament sessions on February 1, 1675, Field (2008) states that a faded note on the manuscript linked it to a collection on 1674. Other parts of the Middlesex records were based on an assessment between 1674 and 1675. The data for Southwark come from an assessment for Surrey that was not dated. However, Field (2008) notes that it is probably associated with a collection in 1673.

[11] The map is available online from the British Library website

Some might question whether the number of hearths is a reasonable way to measure the quality of the building. We believe that it is reasonable for a few reasons. First, unlike assessed values or market values, the number of hearths is an objective measure and is not based on a valuation. Second, Field (2008) documents that research has shown that there is some correlation between the number of hearths and wealth, as well as occupation. To the extent that the wealthier and those with higher social standing live in higher quality buildings, then we should expect the number of hearths to be a reasonable proxy for the quality of the building.

Details of Fire Court judges. The Fire Court was composed of England's twelve common-law judges. There were three common-law courts (Common Pleas, Kings Bench, and the Exchequer) with four justices appointed to each court. In the years after the Fire, some judges retired or passed away and hence our sample contains fourteen judges and not twelve.

In order to get details about the Fire Court judges, we referred to various books such as the dictionary of national biography (2014) and Sainty (1993). From these sources, we obtained information on the judges. Many seismic political events took place in 17th century England. For example, the English Civil War (1642-1651), the restoration of the monarchy (1660), as well as the Puritans' (English Protestants) continuous attempts to get the Church of England (established church) to abandon its Roman Catholic practices. Therefore, from these books, we also obtained information on the judges' religious views and their views on the 1660 restoration of the monarchy (i.e., whether they were Royalists or Parliamentarians). We define binary variables for whether the judges were supportive of the restoration of the monarch (Royalists) and whether they were supportive of the established church. We assign the value of 0.5 if the judges had moderate views. In our IV analysis, we use the composition of the judging panels an instrument for the share of initial cases in the parish that had pragmatic rulings.

Table 2.2 shows us the summary statistics of the Fire Court judges. On average, the judges tend to be slightly pro-restoration of the monarchy and pro-established church. Around 36% of the judges attended Oxford University with the rest attending Cambridge University. The majority of the judges trained at the Inner Temple. Finally, 43% of the Fire Court judges were from the court of the Common

Pleas and 29% of them were the respective heads of their common-law courts (i.e., Lord Chief Justice or Lord Chief Baron).

Table 2.2: Summary statistics (Judges)

	Mean	SD	Min	Max	N
Year of birth	1603.14	6.79	1587	1611	14
Year called to bar	1629.14	6.51	1614	1637	14
Year knighted	1658.71	7.02	1643	1668	14
Pro-restoration of monarchy	.57	.43	0	1	14
Pro-established church	.57	.43	0	1	14
Studied at Oxford University	.36	.5	0	1	14
Served in Grays Inn	.07	.27	0	1	14
Served in Lincolns Inn	.36	.5	0	1	14
Served in Inner Temple	.5	.52	0	1	14
Served in Middle Temple	.07	.27	0	1	14
From Common Pleas	.43	.51	0	1	14
From Kings Bench	.29	.47	0	1	14
From Exchequer	.29	.47	0	1	14
Head of common-law Court	.29	.47	0	1	14
Number of judges	.	.	.	.	14

Details of Fire Court cases. The transcripts of the cases that were heard by the Fire Court were compiled into nine volumes. These records survive up to today and are housed at the London Metropolitan Archives. To commemorate 300 years since the Fire, in 1966, four volumes (volumes A, B, C and D) were calendared (summarized) and converted to modern English by Philip E. Jones. These were subsequently published as two books – Jones (1966) and Jones (1970). The summaries contain extremely detailed information. For example, they give us details on who the landlords and tenants are, the location of the property, the rent and tenure of the tenancy contract before the Fire, the day that the case was heard by the Fire Court, the judges who heard the case, as well as the new rent and tenure that were decreed by the panel of judges. Figure 2.2 shows an example of a case summary from the books while figure F1 shows how some of the case characteristics evolved over time (within the first 716 days).

Figure 2.2: Example of Fire Court case summary

G. A–19; B.M. 5063–13

28 March 1667. Lord Chief Baron(s), Baron Turnor(s), Justice Archer(s).

John Beale v. Sir George Moore Bt.

On the previous day it had appeared from the petition that Sir George Moore in consideration of a fine of £200 and a rent of £90 p.a. had leased the back house and greater shop, part of the Black Mule in Fleet Street, parish of St. Dunstan in the West, ward of Farringdon Without, to the petitioner for 20 years from Mich. 1659, and the other house next the street and the lesser shop to Henry Somner for 20 years from 24 June 1659 for a fine of £140 and a rent of £45 p.a., and that Somner, for a fine of £135 had sublet his house and shop to Peter Pinder for 13 years from 24 June 1666 at the said rent of £45 p.a. Beale and Pinder were in possession at the time of the Fire and Moore had entered upon the premises for non-payment of rent at Christmas 1666, prior to the petition exhibited here in Court.[1]

After a very long debate upon terms for rebuilding Moore offered to increase the petitioner's term to 41 years at a reduced rent of £80 p.a., the first payment at Midsummer 1668, and to contribute £260 towards the cost of rebuilding. The petitioner requested that the rent be reduced by £20 p.a., whereupon Moore asked that his offer be accepted or the lease surrendered. The Court "being unwilling to have the petitioner surprised" gave him until today to consider the offer.

Now, the petitioner still refusing, the Court persuaded Moore to increase his contribution to £360, but the petitioner still refused. The Court thereupon ordered that the petitioner should surrender the lease and the term cease and that Moore should re-enter upon the ground now lying waste with full power to rebuild.

Source: Jones (1966)

As part of his history PhD thesis, Field (2008) transcribed some of the information associated with the cases in these four volumes into a data set. We augment this data set by transcribing additional information that was not captured by Field (2008). Table 2.3 shows us the summary statistics of the Fire Court data based on the cases which we have sufficient information. In 13% of the cases in our sample, the Fire Court voided the existing contracts (i.e., both landlord and tenant surrendered the existing contract) and assigned the cost of rebuilding to the landlord. In 1.3% of the cases, the judges altered the existing contracts (i.e., no surrendering) and assigned the rebuilding to the landlord. In 71% of the cases, the judges altered the existing contracts and assigned the rebuilding to the tenant. In 10% of the cases, the the Fire Court voided the existing contracts but decreed the sharing of cost in the rebuilding. Finally, in 5.2% of the cases, the judges altered the existing contracts but decreed the sharing of cost in the rebuilding. The fine paid is the lump-sum payment made on execution of the lease. For each judging panel, we calculate the share of judges who were supportive of the established church and the share of judges who were supportive of the 1660 restoration of the monarchy (Royalists). On average, in each judging panel, 48% of the judges tend to be supportive of the restoration of the monarchy and 46% tend to be supportive of the established church. This suggests that the judging panels were on average quite moderate in their views.

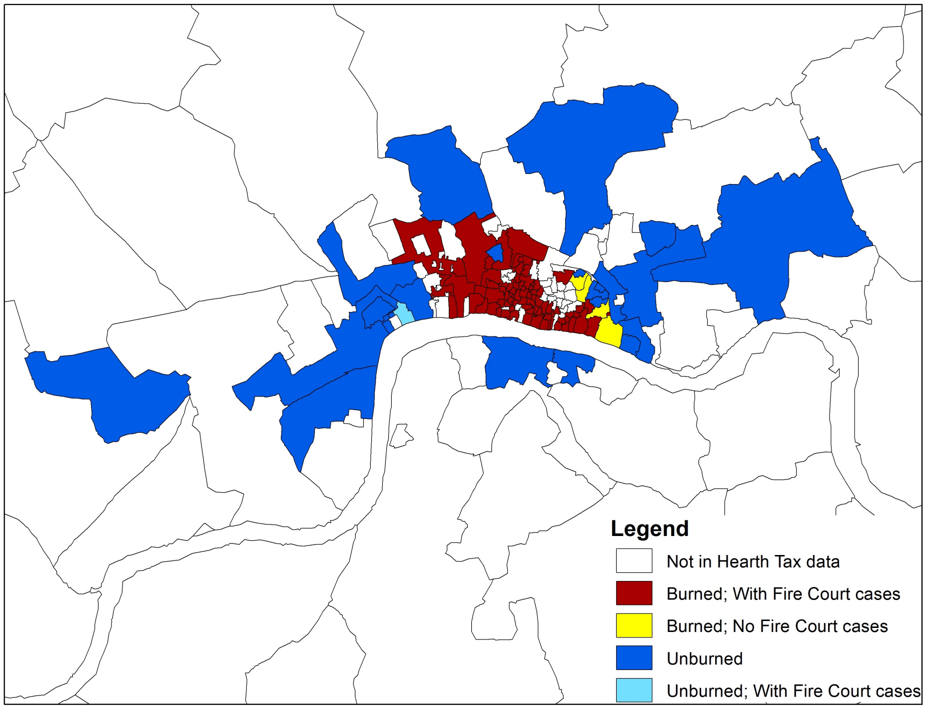
Finally, since we are interested in examining whether legal rulings of the initial cases in each parish can anchor expectations, we do not actually need to observe the rulings of all the cases that went to the Fire Court. Therefore, the four out of nine volumes which have been calendared would suffice for our purposes as these four volumes cover the earlier cases. We refer to these cases from the first four volumes as the "initial" cases.

Table 2.3: Summary statistics (Fire Court data)

	Mean	SD	Min	Max	N
Both parties surrender: Owner rebuilds	.13	.33	0	1	696
No surrender: Owner rebuilds	.01	.11	0	1	696
No surrender: Tenant rebuilds	.71	.46	0	1	696
Both parties surrender: Cost sharing	.10	.30	0	1	696
No surrender: Cost sharing	.05	.22	0	1	696
Degree of separation from owner	1.18	.52	1	6	696
Start year of tenancy	1655.74	10.23	1591	1666	679
Years left in tenancy	34.75	387.83	0	9996.92	663
Fine paid	69.08	202.61	0	4000	696
Rent per annum	31.29	36.91	0	474	692
Amount spent on improvements	48.46	233.07	0	3000	696
Average pro-monarchy of panel	.48	.22	0	1	696
Average pro-church of panel	.46	.21	0	1	696
Head of common-law Court	.57	.6	0	3	696
Number parishes	.	.	.	.	67
Number cases	.	.	.	.	696

Regression sample. Putting all our data sources together, figure 2.3 shows the parishes that are included in our regressions. In the diagram, we label a parish as "burned" as long as any part of it was damaged by the Fire.

Figure 2.3: Regression sample



Source: Satchell et al. (2018); with authors' edits

2.4 The Effect of the Fire

The 1666 Great Fire of London had both quantity and quality effects in the development of London. On the quantity side, the Fire affected the total number of properties and hearths in each parish. As for quality, the Fire affected the number of hearths per property in each parish. In this paper, we focus on the effect that the Fire had on quality. This is because the plague wiped out about a quarter of London's population. Therefore, we should expect fewer properties to be rebuilt in the immediate aftermath since there are now fewer people to house. However, the effect on quality is not clear. In addition, the reduction in the number of properties is consistent with post-Fire regulations that stipulated that properties needed to be of a certain minimum size. Finally, our data end in 1675 (nine years after the Fire) so it could be the case that London had not reached a new stationary state – i.e., it is too early to tell if the number of properties converged to a new steady state. For these reasons, the main focus of our analysis is on quality as opposed to quantity. Nevertheless, in Appendix D, we examine the effect that the Fire had on the total number of properties and hearths in each parish.

2.4.1 Empirical strategy

To examine the effect of the Fire on the number of hearths per property, we use a DiD empirical strategy:

$$ln\left(Hearths_{ijt}\right) = \alpha_j + \delta PostFire_t + \beta Burned_j \times PostFire_t + \gamma' X_{jt} + \epsilon_{ijt} \quad (2.1)$$

$ln\left(Hearths_{ijt}\right)$ is the log number of hearths in property i in parish j in period t. The two periods are before the Fire and after the Fire. $Burned_j$ is an indicator variable that denotes whether property i was in a parish that experienced damage from the Fire. $PostFire_t$ is an indicator variable for the period after the Fire. X_{jt} is a vector of controls. Finally, α_j are parish fixed effects. We cluster the standard errors at the parish level. A null effect would suggest that there were no frictions to upgrading before the Fire – the quality of properties was optimal. By contrast, a positive effect suggests the presence of upgrading frictions which the Fire effectively removed.

While we use the log number of hearths in our main specifications, there could be worries that taking logs results in dropping observations which are recorded as having zero hearths or be susceptible to Jensen's inequality (see E). To account for the zeros in the outcome variables, we adopt two approaches: (i) applying the inverse hyperbolic sine transform to hearths; and (ii) using a Poisson pseudo-likelihood (PPML) regression (see F). We chose the log number of hearths because the DiD of the number of hearths failed the pre-trends test. Switching to a hyperbolic sine transform and a PPML model did not change our conclusion to the impact of legal rulings on the number of hearths per property. However, the impact of the fire on the number of hearths per property became insignificant (see section 2.4.3).

For those interested in the cross-sectional regressions in each time period, the results are reported in table F2. In the pre-Fire period, the number of hearths per property in burned versus unburned parishes was statistically indistinguishable.

2.4.2 Results and discussion

Higher quality structures. Table 2.4 reports the impact that the Fire had on the number of hearths per property in the burned parishes relative to the unburned parishes. The estimate in column 1 shows that controlling for parish and time fixed effects, burned parishes saw a highly statistically significant increase of around 26.3% more hearths compared to unburned parishes. While in percentage terms this magnitude might seem large, given that the average number of hearths before the Fire was 3.83, this translates to an increase of 1.01 hearths.

There could be concerns that there are other time varying parish-level variables that are driving the results. For example, larger or richer parishes may recover faster from the Fire as they are able to bring together more resources. To address these concerns, in column 2, we include a series of parish controls interacted with $PostFire_t$. These include the number of properties in the parish before the Fire, the share of peers,[12] high-ranking military personnel (i.e., Colonel or Captain) and doctors living in the parish. The estimated effect remains robust to the inclusion of these time varying parish-level controls.

Next, to control for geographical characteristics, we classified the parishes into broader locations (i.e., abutting the City of London walls, within the walls and outside the walls). In column 3, we show that the results are stable to the inclusion of these broader locations-by-post fixed effects. Finally, we grouped parishes into terciles based on the number of hearths in the parish before the Fire. This is to control for the possibility that there may be persistence in the number of hearths – properties with more hearths will rebuild with more hearths and those with fewer hearths will rebuild with fewer hearths. In column 4, we show that the results are relatively stable even when we include these pre-Fire hearth terciles-by-post fixed effects. Figure F2 shows the binned scatter plot of the results in column 4.

[12] These are Duke, Duchess, Marquess, Marchioness, Earl, Countess, Viscount, Viscountess, Baron, Baroness, Lord, Lady, Sir, Dame and Ambassador.

Table 2.4: Effect of Fire on the number of hearths per property

	(1)	(2)	(3)	(4)
VARIABLES		ln(No. hearths)		
Parish Burned X Post Fire	0.263***	0.239**	0.219*	0.283**
	(0.092)	(0.098)	(0.127)	(0.116)
Observations	77,093	77,093	77,093	77,093
R-squared	0.009	0.010	0.010	0.014
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	70	70	70	70

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Our results show that after the Fire, inhabitants of the parishes constructed more hearths per property. This suggests that there was indeed the presence of substantial frictions that was impeding development. By reducing the opportunity cost of waiting to rebuild to zero and forcing everyone to build at the same time, the Fire freed the parishes from the constraints imposed by their existing durable structures. This consequently spurred development through stronger cross-building spillovers and led to a new equilibrium which involved more hearths per property.

Finally, as our dependent variable has been log transformed, there could be issues of Jensen's inequality. In particular, running the regression with the log transformed dependent variable could result in an opposite treatment effect as compared to if we were to run the regression without taking logs. In Appendix E we provide a discussion about this potential issue and show that we get a positive treatment effect in both the regression without logs and the regression in logs.

Effect varied with the level of damage. A priori, we should expect the effect of the Fire to vary with the level of damage. For example, in the extreme, if the Fire was so small that it only damaged one building, then the Fire would not have been effective in removing rebuilding frictions and there would be no widespread reconstruction.

We use two different approaches to examine such heterogeneous effects. First, we split the $Burned_j \times PostFire_t$ variable into two dummy variables – $SlightlyBurned_j \times PostFire_t$ and $CompletelyBurned_j \times PostFire_t$. As the names suggest, $SlightlyBurned_j$ refers to parishes where less than half of the parish (in terms of geographical area) was burned while $CompletelyBurned_j$ refers to parishes where more than half of the parish was burned. Table 2.5 reports the results of this heterogeneous treatment effect regression. Across all columns, we see that the effect of the Fire was greater in parishes that were completely burned.

Table 2.5: Effect of the extent of Fire on the number of hearths per property

VARIABLES	(1)	(2)	(3)	(4)
	ln(No. hearths)			
Parish Completely Burned X Post Fire	0.403***	0.445***	0.471***	0.607***
	(0.064)	(0.089)	(0.073)	(0.102)
Parish Slightly Burned X Post Fire	0.173	0.136	0.144	0.192
	(0.117)	(0.133)	(0.146)	(0.130)
Observations	77,093	77,093	77,093	77,093
R-squared	0.012	0.013	0.013	0.018
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	70	70	70	70

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

The second approach which we adopt is to use whether the church in the parish was damaged as a proxy for the level of destruction in the parish due to the Fire. We think that this is reasonable given that the church was often the centre of economic and social life during this period of time. To do this, we run equation (2.1) comparing burned parishes where the church was damaged to unburned parishes. In the same regression, we also compare burned parishes where the church was not damaged to unburned parishes. Table 2.6 reports the results. Across all columns, we see that the effect of the Fire was greater in parishes where the church was also damaged by the Fire.

Table 2.6: Effect of the church being damaged on the number of hearths per property

VARIABLES	(1)	(2) ln(No. hearths)	(3)	(4)
Parish burned and church damaged X Post Fire	0.393***	0.401***	0.400***	0.485***
	(0.073)	(0.100)	(0.115)	(0.122)
Parish burned but church not damaged X Post Fire	0.210*	0.189	0.190	0.252**
	(0.108)	(0.115)	(0.131)	(0.118)
Observations	77,093	77,093	77,093	77,093
R-squared	0.011	0.011	0.011	0.015
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	70	70	70	70

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Both approaches suggest that the effect of the Fire was greater in parishes where the destruction was more widespread. Nonetheless, some of the positive effect that we find in table 2.4, table 2.5 and table 2.6 could be mechanical. This is because as noted by Field (2008), new houses had to be built according to strict regulations that specified the size and materials used. In addition, given the excess supply of land due to the plague, land was probably cheaper. This could lead to people wanting larger houses with more hearths per house. We discuss how we can rule out such mechanical effects in the next paragraph.

Effect varied with the level of damage in surrounding parishes. To rule out the mechanical effect of larger houses having more hearths, we split the $Burned_j \times PostFire_t$ variable into two dummy variables – $AllNeighboursBurned_j \times PostFire_t$ and $NotAllNeighboursBurned_j \times PostFire_t$ and re-run equation (2.1). If the increase in the number of hearths per property is purely due to larger houses, then it should not vary with the level of damage in the surrounding parishes.

Table 2.7 shows that spatial spillovers matter. Burned parishes that were completely surrounded by other burned parishes experienced building investments that were two to three times higher than burned parishes that were not completely surrounded by burned parishes. This regression shows us strong evidence that the increase in hearths per building is driven by cross-building spillovers and not the mechanical effect of larger properties.

Table 2.7: Effect of spatial spillovers on the number of hearths per property

VARIABLES	(1)	(2) ln(No. hearths)	(3)	(4)
Not All Neighbors Burned X Post Fire	0.174	0.143	0.145	0.178
	(0.113)	(0.123)	(0.144)	(0.131)
All Neighbors Burned X Post Fire	0.395***	0.410***	0.409***	0.474***
	(0.069)	(0.098)	(0.102)	(0.124)
Observations	77,093	77,093	77,093	77,093
R-squared	0.010	0.011	0.011	0.014
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	70	70	70	70

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

2.4.3 Robustness checks

Dropping parishes which merged after the Fire. One concern could be that the Fire led to the merging of some parishes and so our results could be driven by these enlarged parishes which might have more resources. In table F3 we re-run equation (2.1) using only parishes that did not merge with other parishes after the Fire. Reassuringly, the coefficient estimates remain very similar to our baseline results.

Using different control groups. There could be concerns that some of the unburned parishes in our control group may not be appropriate for our analysis. To see this, consider a hypothetical unburned parish (parish U) that is surrounded by many burned parishes. Given the destructive nature of the Fire, it is somewhat surprising that parish U did not suffer any damage from the Fire. This could suggest that parish U is fundamentally different from its neighbouring parishes that were burned. For example, parish U could have been more wealthy and hence more able to quickly mobilize fire-fighting efforts. It could also be the case that more buildings in parish U were made of bricks as opposed to wood. In addition, parish U could have also pre-empted the spread of the Fire by tearing down buildings that were near to the parishes that were burning. The Fire Court records suggest that this indeed happened in parishes such as St Botolph Bishopsgate and St Mary-le-Strand. Therefore, it might not be suitable to use such parishes as the control group to the burned parishes.

To address this concern, we run separate regressions based on two samples. First, a "nearby" sample which consists of all burned parishes plus unburned parishes that share a boundary with any burned parishes. The results for this are reported in table F4. Second, a "further away" sample which consists of all burned parishes and unburned parishes that do not share any boundaries with any burned parishes. The results from this sample are reported in table F5. The results using both samples are very similar, suggesting that our analysis is not sensitive to the choice of control groups.

Accounting for zeros in the outcome variable. There are 2,637 observations which are recorded as having zero hearths in the property. Taking logs results in these observations dropping out of the regression. Therefore, to account for the zeros in the outcome variables, we adopt two approaches. First, applying the inverse hyperbolic sine transform to hearths. Second, using a Poisson pseudo-likelihood (PPML) regression. The results are reported in table F6 and table F7 respectively. While the estimated effects remain positive, the magnitudes are now halved and are imprecisely estimated.

Trimming extreme values of the outcome variable. As another robustness check, we drop the top and bottom one percentile of $ln\left(Hearths_{ijt}\right)$. While the estimated effects are now halved, they remain positive and are mostly statistically significant. The results of this robustness check are reported in table F8.

Checking for parallel trends. A key assumption of a DiD strategy is that of parallel trends. There are two methods to argue that the number of hearths per property in the burned versus unburned parishes was experiencing the same trends before the Fire. The first method is to rely on the historical context. The historical setting suggests that the Fire spread based on where the wind blew and not due to the economic or social characteristics of the parish. Referring to figure 2.1, this seems to indeed have been the case. While the Fire started in Pudding Lane which was in the eastern part of the City of London, contemporaneous reporting by the Gazette (1666) notes that due to the "violent Easterly wind", the Fire spread mostly to the west. As a result, almost all the parishes that were damaged were to the west of Pudding Lane.

The wind blowing from the east to the west during the Fire is an important point. This is because Heblich et al. (2021) show that in England, the wind usually blows from the west or south-west. Therefore, we can make an argument that whether a parish ended up being burned was unexpected, random and independent of pre-trends.

The second method would be to run a placebo DiD regression to compare the number of hearths per property in the burned versus unburned parishes in the periods before the Fire. We should find no effect if there are parallel trends. Unfortunately, due to data limitations, we are not able to do so in the most robust manner. This is because the hearth tax was introduced in 1662 and for the pre-Fire period, we only have data from the 1662, 1664, 1665 and 1666 hearth tax records. Due to the differences in the scope and range of the pre-Fire hearth tax assessments, almost all parishes (65 out of 70) appear only in one year. In addition, table 2.8 shows that in some years, the data we have were either all for burned or unburned parishes. Therefore, we had to pool all the 1662, 1664, 1665 and 1666 data to form the pre-Fire period in our regressions.

Table 2.8: Pre-Fire data

Year	Burned parishes in the data	Unburned parishes in the data	Total
1662	16	0	16
1664	0	2	2
1665	0	2	2
1666	38	17	55

To run placebo regressions to test for pre-trends, we would need data for both burned and unburned parishes in at least two pre-Fire years. However, with the exception of 1666, the other three pre-Fire years only consist of data from either burned or unburned parishes. In order to try our best to provide statistical evidence to rule out pre-trends, what we can do is to classify the pre-Fire period into two categories. First, the year 1666 would be period $t-1$ and the years 1662 and 1664 would be $t-2$.[13] This method has some limitations such as assuming that the data in 1664 are very similar to 1662 and the two unburned parishes in 1664 are representative of all the other unburned parishes. Nevertheless, accepting these

[13] We do not include year 1665 because the two parishes that appear in 1665 are south of the river and are hence very different from the other parishes in the data for the placebo regression.

limitations allows us to run the following placebo regression to test for pre-trends:

$$ln\,(Hearths\,per\,property_{jt}) = \alpha_j + \delta PostPlaceboFire_t + \beta Burned_j \times PostPlaceboFire_t + \gamma' X_{jt} + \epsilon_{jt}$$

$PostPlaceboFire_t$ is an indicator variable for the period after a placebo fire. We set this as the period after 1665. Finding a large and statistically significant effect from this phantom event would cast serious doubts on the validity of our identification strategy. If our regression passes the parallel trends test, then we should expect β to be small and statistically insignificant. Table 2.9 shows that this is indeed what we get when we run this placebo regression, suggesting the absence of pre-trends.

Table 2.9: Effect of placebo Fire on the number of hearths per property

VARIABLES	(1) ln(No. Hearths per Property)
Parish Burned X Post Placebo Fire	-0.009
	(0.023)
Observations	40,517
R-squared	0.004
Parish FE	✓
Post FE	✓

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2.5 The Effect of Pragmatic Legal Rulings

2.5.1 Overview

In the previous section, we found that the 1666 Great London Fire resulted in a higher number of hearths per property in burned parishes relative to unburned parishes. While our results suggest that individuals had positive expectations about how much other individuals in their parish would be investing, it does not tell us what is driving these expectations. Therefore, in this section, we examine if legal rulings could be a driver of these expectations.

2.5.2 Defining pragmatic legal rulings

The Fire Court judges were given the power to completely void existing contracts or alter the terms of these contracts. Voiding the existing contract means that both the landlord and tenant surrender their contracts. This allows both parties to negotiate a new contract with each other or other parties. In addition, voiding the contract does away with the judging panels arbitrarily setting a new rent and lease. By contrast, altering the terms of an existing contract means that the tenant remains the same but the Fire Court decrees a new rent and/or length of lease. In addition, the Fire Court would also decree that either the landlord or tenant rebuilds, or that both parties are to contribute towards the rebuilding. In our paper, we define the voiding of existing contracts and the assigning of the cost of rebuilding to the landlord as pragmatic legal rulings. 12.7% of cases fall into this category.

We define such rulings as pragmatic because they help to facilitate a higher number of hearths per property for the following reasons. First, tenants are likely to be more credit-constrained compared to landlords and are hence more likely to rebuild at a lower quality (i.e., fewer hearths per property). Second, assigning the rebuilding to the landlord represents a clear assignment and alignment of property-rights. Third, since the occupant is responsible for paying the hearth tax, if the tenant was assigned the rebuilding, she is likely to rebuild with fewer hearths to reduce her tax burden. For these reasons, assigning the rebuilding to the landlord rather than the tenant facilitates the rebuilding of London.

In theory, we could have expanded our definition of pragmatic rulings to also include cases where the judges altered existing contracts (i.e., did not void the contract) but assigned the rebuilding responsibility to the landlord. However, in such cases the Fire Court's rulings were often multi-dimensional. For example, it could be the case that although the rebuilding responsibility was assigned to the landlord, the judges could have decreed a lower rent. In this instance, the landlord may then choose to rebuild at a lower quality since the rent she is receiving is now lower. In order to circumvent the issue of multi-dimensional rulings, we focus on the most extreme of case outcomes – cases where the Fire Court voided existing contracts and assigned the rebuilding responsibility to the landlord.[14]

[14] Finally, we could have also expanded our definition of pragmatic rulings to include cases where the judges voided the existing contracts but decreed cost sharing in the rebuilding. However, the reason why we do not do this is because we wanted our definition to reflect the complete burden of rebuilding falling on the landlord. This will be clearer when we discuss the model in the next section.

2.5.3 Model: Legal rulings, expectations and investment

How exactly did legal rulings drive expectations and hence help to coordinate investment (i.e., the number of hearths in each property)? We show this using a Nash Jr (1950) bargaining game where tenants and landlords bargained over the terms of rebuilding. The bargaining game consists of two stages. In Stage 1, in each parish j, the landlord and tenant of each property i bargain over a contract given their respective outside options. The outside options are based on the rulings established by the Fire Court in its initial cases for each parish. Therefore, the outside options vary across parishes. For simplicity, we suppress the subscripts j and i. We define the contract $\{r, t, I^l\}$ in terms of the annual rent (r), the tenancy length (t), and the amount of contributions (investment) that the landlord makes towards the rebuilding (I^l).[15] If the tenant and landlord reach an agreement, they move to the second stage where the tenant decides on her amount of contributions (investment) towards the rebuilding. The total amount of building investment (measured in terms of the number of hearths in the property) is given by the sum of the landlord's investment (determined in the first stage) and the tenant's investment (determined in the second stage). If they fail to reach an agreement, they bring their case to the Fire Court. In this framework, the Fire Court's rulings affect the outside options and hence the bargaining dynamics between the landlords and tenants. The model is solved by backward induction.

Solving the Nash bargaining game: Stage 2

The tenant's problem in Stage 2 is to choose I^t to maximize utility given the contract $\{r, t, I^l\}$ that was determined in Stage 1:

$$\max_{I^t} \ U\left(r, t, I^l, I^t\right) = \frac{1-\beta^t}{1-\beta}\left[h(I^l, I^t) - r\right] + \frac{\beta^t}{1-\beta}u^0 - pI^t \tag{2.2}$$

β is the discount factor. p is cost per unit of investment. u^0 is the tenant's utility after the tenancy ends. We assume that $h(I^l, I^t)$ is concave and that the amount of investments that the landlord and tenant make towards the building are complements ($\frac{\partial h}{\partial I^l \partial I^t} > 0$).

[15] More accurately, the parties bargain over the split of the total surplus. In doing so, the parties are implicitly choosing $\{r, t, I^l\}$.

The first order condition is:

$$\frac{1-\beta^t}{1-\beta}h'(I^l, I^t) = p \tag{2.3}$$

equation (2.3) suggests that the tenant's investment does not depend on the rent.

Proposition 1. *The tenant's investment is increasing in tenancy length.*
Proof: From equation (2.3), let $\psi\left(t, I^l, I^t\right) \equiv \frac{1-\beta^t}{1-\beta}h'(I^l, I^t) - p$. By the implicit function theorem and since $h(I^l, I^t)$ is concave:

$$\begin{aligned}\frac{\partial I^t}{\partial t} &= -\frac{\psi_t}{\psi_{I^t}} \\ &= \frac{-\beta^t \beta h'(I^l, I^t)}{-\frac{1-\beta^t}{1-\beta}h''(I^l, I^t)} > 0 \; \square\end{aligned} \tag{2.4}$$

Let the optimal tenant investment be denoted as: $I^{t*} \equiv g\left(t, I^l\right)$. Therefore, total investment is: $I\left(t\right) \equiv g\left(t, I^l\right) + I^l$

Solving the Nash bargaining game: Stage 1

In this stage, the tenant and landlord bargain over the surplus given their respective outside options π^c and u^c. Their outside options are based on the Fire Court's rulings in the initial cases. We assume that π^c and u^c vary across parishes. We represent the distribution of Fire Court decisions in the initial cases as $F(r, t, I^l)$. λ is the bargaining weight which we assume to be exogenous. The Nash bargaining game solution can be characterized as:

$$\underset{\left\{r,t,I^l\right\}}{max} \left[\Pi\left(r, t, I^l\right) - \pi^c\right]^\lambda \left[U\left(r, t, I^l\right) - u^c\right]^{(1-\lambda)} \tag{2.5}$$

where the landlord's utility is $\Pi\left(r, t, I^l\right) = \frac{1-\beta^t}{1-\beta}r + \frac{\beta^t}{1-\beta}r^0\left(I(t)\right) - pI^l$. $r^0\left(I(t)\right)$ is the rent that the landlord receives from the next tenant after the tenancy agreement with the current tenant expires. In addition,

$$\pi^c = \int\int\int \frac{1-\beta^y}{1-\beta}x + \frac{\beta^y}{1-\beta}r^0\left(I(y)\right) - pz \; dF(x, y, z)$$

and

$$u^c = \int\int\int \frac{1-\beta^y}{1-\beta}\left[h(I(y)) - x\right] + \frac{\beta^y}{1-\beta}u^0 - pg(y,z)\; dF(x,y,z)$$

The Nash bargaining solution for the landlord is:

$$\Pi\left(r,t,I^l\right) = \lambda[\Pi\left(r,t,I^l\right) + U\left(r,t,I^l\right) - \pi^c - u^c] + \pi^c \tag{2.6}$$

and that for the tenant is:

$$U\left(r,t,I^l\right) = (1-\lambda)[\Pi\left(r,t,I^l\right) + U\left(r,t,I^l\right) - \pi^c - u^c] + u^c$$

Rearranging equation (2.6), we get that:

$$\begin{aligned}(1-\lambda)\left[\frac{1-\beta^t}{1-\beta}r + \frac{\beta^t}{1-\beta}r^0\left(I(t)\right) - pI^l\right] - \pi^c = \\ \lambda\left[\frac{1-\beta^t}{1-\beta}\left[h(I(t)) - r\right] + \frac{\beta^t}{1-\beta}u^0 - pg(I^l,t) - Q^c\right]\end{aligned} \tag{2.7}$$

where $Q^c \equiv \pi^c + u^c = \int\int\int \frac{1-\beta^y}{1-\beta}h(I(y)) + \frac{\beta^y}{1-\beta}\left[r^0\left(I(y)\right) + u^0\right] - pI(y)\; dF(x,y,z)$

Next, we assume that the judging panel's preferences for the landlord's contribution to the rebuilding is orthogonal to r and t. In other words, $F(x,y,z) \equiv F_{XY}(x,y)F_Z(z)$. Assuming that $F(x,y,z) \equiv F_{XY}(x,y)F_Z(z)$ has two implications. First, this assumption implies that the sum of the outside options is not affected by the transfer of the burden to rebuild:

$$Q^c = \int\int \frac{1-\beta^y}{1-\beta}h(I(y)) + \frac{\beta^y}{1-\beta}\left[r^0\left(I(y)\right) + u^0\right] - pI(y)\; dF(x,y)$$

Second, while the sum of the outside options is not affected by the transfer of the burden to rebuild, the outside option of the landlord still depends on $F_Z(z)$. Given these two implications, equation (2.7) can be expressed as:

$$\begin{aligned}(1-\lambda)\left[\frac{1-\beta^t}{1-\beta}r + \frac{\beta^t}{1-\beta}r^0\left(I(t)\right) - pI^l\right] - \pi^c\left(F_Z(z)\right) = \\ \lambda\left[\frac{1-\beta^t}{1-\beta}\left[h(I(t)) - r\right] + \frac{\beta^t}{1-\beta}u^0 - pg(I^l,t) - Q^c\right]\end{aligned}$$

Now suppose that there is a contract $\left\{\bar{r}, \bar{t}, \bar{I}^l\right\}$ that satisfies the Nash bargaining game solution. However, we switch from $F_Z(z)$ to $F'_Z(z)$, where $F'_Z(z)$ first order stochastically dominates $F_Z(z)$. Recall that $F(r, t, I^l)$ represents the distribution of Fire Court rulings in the initial cases. In our empirical context, moving from $F_Z(z)$ to $F'_Z(z)$ corresponds to the initial cases getting assigned judging panels that have a greater probability of voiding the existing contracts and assigning the cost of rebuilding to the landlord. As explained in the previous section, this is what we define as pragmatic legal rulings.

Proposition 2. *The landlord's outside option* (π^c) *falls when the initial cases are assigned judging panels that have a greater preference for the landlord to contribute more to the rebuilding.*

Proof: Since $F'_Z(z)$ first order stochastically dominates $F_Z(z)$, this implies that the landlord's outside option under $F'_Z(z)$ is now smaller:

$$\begin{aligned} & F'_Z(z) \leq F_Z(z) \; \forall z \quad \text{and} \quad F'_Z(z) < F_Z(z) \text{ for some } z \qquad (2.8) \\ \Rightarrow\; & \pi^c\left(F'_Z(z)\right) < \pi^c\left(F_Z(z)\right) \end{aligned}$$

The last inequality is because $\pi^c = \int\int \frac{1-\beta^y}{1-\beta}x + \frac{\beta^y}{1-\beta}r^0\left(I(y)\right)\; dF_{XY}(x,y) - p\int z\; dF_Z(z)$. Since $F'_Z(z) < F_Z(z)$, $p\int z\; dF'_Z(z) > p\int z\; dF_Z(z)$ and so $\pi^c\left(F'_Z(z)\right) < \pi^c\left(F_Z(z)\right)$.□

Since the landlord now has a smaller outside option, the contract $\left\{\bar{r}, \bar{t}, \bar{I}^l\right\}$ no longer satisfies the Nash bargaining game solution and equation (2.7) no longer holds with equality. Instead, the landlord now has too much of the surplus and so:

$$\begin{aligned} & (1-\lambda)\left[\frac{1-\beta^{\bar{t}}}{1-\beta}\bar{r} + \frac{\beta^{\bar{t}}}{1-\beta}r^0\left(I(\bar{t})\right) - p\bar{I}^l\right] - \pi^c\left(F'_Z(z)\right) > \\ & \lambda\left[\frac{1-\beta^{\bar{t}}}{1-\beta}\left[h(I(\bar{t})) - \bar{r}\right] + \frac{\beta^t}{1-\beta}u^0 - pg(\bar{I}^l, \bar{t}) - \bar{Q}^c\right] \qquad (2.9) \end{aligned}$$

In order to achieve equality, changes in the Nash bargained contract should (1) lower the left-hand side of the inequality and increase the right-hand side or (2) increase the right-hand side more than the left-hand side.

Proposition 3. *The Nash bargained rent* ($\bar{r}$) *decreases when the judging panels have a greater preference for the landlord to contribute more to the rebuilding.*

Proof: Referring to inequality 2.9, since $\frac{\partial LHS}{\partial \bar{r}} = \frac{1-\beta^{\bar{t}}}{1-\beta}(1-\lambda) > 0$ and $\frac{\partial RHS}{\partial \bar{r}} =$

$-\frac{1-\beta^{\bar{t}}}{1-\beta}\lambda < 0$, in order for the left-hand side to equal to the right-hand side, $\bar{r}$ has to decrease. A decrease in $\bar{r}$ decreases the left-hand side and increases the right-hand side.□

Proposition 4. *The Nash bargained landlord investment to the rebuilding ($\bar{I}^l$) increases when the judging panels have a greater preference for the landlord to contribute more to the rebuilding.*

Proof: Referring to inequality 2.9, since $\frac{\partial LHS}{\partial \bar{I}^l} = -p(1-\lambda) < 0$ and $\frac{\partial RHS}{\partial \bar{I}^l} = -p\lambda g_{\bar{I}^l}\left(\bar{I}^l, \bar{t}\right) = p\lambda > 0$, in order for the left-hand side to equal to the right-hand side, $\bar{I}^l$ has to increase. An increase in $\bar{I}^l$ decreases the left-hand side and increases the right-hand side.□

Proposition 5. *If the relative bargaining weight of the landlord is more than the relative marginal benefit of increasing the tenancy length, then the Nash bargained tenancy length ($\bar{t}$) increases when the judging panels have a greater preference for the landlord to contribute more to the rebuilding.*

Proof: Referring to inequality 2.9,

$$\frac{\partial LHS}{\partial \bar{t}} = (1-\lambda)\left\{(-ln\beta)\beta^{\bar{t}}\left[\bar{r} - r^0(I(\bar{t}))\right] + \frac{\beta^{\bar{t}}}{1-\beta}r_1^0(I(\bar{t}))g_{\bar{t}}\left(\bar{I}^l, \bar{t}\right)\right\} = (1-\lambda)\Pi_{\bar{t}}\left(\bar{r}, \bar{t}, \bar{I}^l\right)$$

and

$$\frac{\partial RHS}{\partial \bar{t}} = \lambda\left\{(-ln\beta)\beta^t\left[h(I(\bar{t})) - \bar{r} - u^0\right] + \left[\frac{\beta^{\bar{t}}}{1-\beta}h'(I(\bar{t})) - p\right]g_{\bar{t}}\left(\bar{I}^l, \bar{t}\right)\right\} = \lambda U_{\bar{t}}\left(\bar{r}, \bar{t}, \bar{I}^l\right).$$

To achieve equality, we need:

$$\begin{aligned} \frac{\partial LHS}{\partial \bar{t}} &< \frac{\partial RHS}{\partial \bar{t}} \\ \Rightarrow (1-\lambda)\Pi_{\bar{t}}\left(\bar{r}, \bar{t}, \bar{I}^l\right) &< \lambda U_{\bar{t}}\left(\bar{r}, \bar{t}, \bar{I}^l\right) \\ \Rightarrow \frac{\lambda}{1-\lambda} &> \frac{\Pi_{\bar{t}}\left(\bar{r}, \bar{t}, \bar{I}^l\right)}{U_{\bar{t}}\left(\bar{r}, \bar{t}, \bar{I}^l\right)} \quad \square \end{aligned}$$

This gives us a necessary and sufficient condition for $\bar{t}$ to increase. In other words, if the relative bargaining weight of the landlord is more than the relative marginal benefit of increasing the tenancy length, then $\bar{t}$ increases. Given the historical context that tenants are obliged to repair or rebuild the premises in the event of disasters or wars, this condition is likely to hold. In addition, in the Fire Court data, we see that the judging panels decreed that the tenant had to rebuild 70.9% of the time.

Putting everything together, our model suggests that if the initial cases are assigned judging panels that have a greater preference for the landlord to contribute more to the rebuilding, then this lowers the landlord's outside option (proposition 2). As a result of the lowering of the landlord's outside option, the Nash bargained annual rent (r) decreases (proposition 3), the amount of investment that the landlord makes towards the rebuilding (I^l) increase (proposition 4), and the effect on the tenancy length (t) increases (proposition 5). Crucially, our model shows us that by changing outside options, the rulings of the Fire Court affected **all** tenants and landlords even if they did not bring their case to the Fire Court. This is how legal rulings affect expectations.

Empirical implication

In our empirical analysis, we estimate the change in the average number of hearths per property in parish j as a result of the initial cases in the parish getting assigned judging panels that have a greater propensity to void existing contracts (i.e., pragmatic rulings). This corresponds loosely to:

$$\begin{aligned}
\frac{\partial E_j(I_i)}{\partial E_j(F_Z(z))} &= \frac{\partial E_j(I_i^t)}{\partial E_j(F_Z(z))} + \frac{\partial E_j(I_i^l)}{\partial E_j(F_Z(z))} \\
&= \underbrace{\frac{\partial E_j(I_i^t)}{\partial E_j(r_i)}}_{\text{=0 by eq. 2}} \times \underbrace{\frac{\partial E_j(r_i)}{\partial E_j(F_{Z_j}(z))}}_{\text{<0 by prop. 3}} + \underbrace{\frac{\partial E_j(I_i^t)}{\partial E_j(I_i^l)}}_{\text{>0 by assumption}} \times \underbrace{\frac{\partial E_j(I_i^l)}{\partial E_j(F_{Z_j}(z))}}_{\text{>0 by prop. 4}} \\
&+ \underbrace{\frac{\partial E_j(I_i^t)}{\partial E_j(t_i)}}_{\text{>0 by prop. 1}} \times \underbrace{\frac{\partial E_j(t_i)}{\partial E_j(F_{Z_j}(z))}}_{\text{>0 by prop. 5}} + \underbrace{\frac{\partial E_j(I_i^l)}{\partial E_j(r_i)}}_{\text{ambiguous}} \times \underbrace{\frac{\partial E_j(r_i)}{\partial E_j(F_{Z_j}(z))}}_{\text{<0 by prop. 3}} \\
&+ \underbrace{\frac{\partial E_j(I_i^l)}{\partial E_j(I_i^l)}}_{\text{=1}} \times \underbrace{\frac{\partial E_j(I_i^l)}{\partial E_j(F_{Z_j}(z))}}_{\text{>0 by prop. 4}} + \underbrace{\frac{\partial E_j(I_i^l)}{\partial E_j(t_i)}}_{\text{ambiguous}} \times \underbrace{\frac{\partial E_j(t_i)}{\partial E_j(F_{Z_j}(z))}}_{\text{>0 by prop. 5}} \qquad (2.10)
\end{aligned}$$

The second line gives us the effect on the tenant's investment and this effect is unambiguously positive. However, the effect on the landlord's investment (third line) is ambiguous. The signs of $\frac{\partial E_j(I_i^l)}{\partial E_j(r_i)}$ and $\frac{\partial E_j(I_i^l)}{\partial E_j(t_i)}$ are ambiguous because the landlord can in principle trade off a higher amount of investment to the building process with a lower rent or longer tenancy length. This happens because there are three variables $\{r, t, I^l\}$ that are governed by a single Nash bargaining equation (see equation (2.6)). If the positive effect on the tenant's investment (second line) dominates the ambiguous effect on the landlord's investment (third line) then pragmatic legal rulings can result in a higher number of hearths per property.

To conclude this section, our model shows that even though landlords and tenants of different properties are bargaining separately and do not bring their case to the Fire Court, they end up choosing similar levels of hearths per property. This is because they have the same focal point and hence expectations of what others will do. This focal point is how the Fire Court ruled in the initial cases in their parish.

2.5.4 Empirical strategy

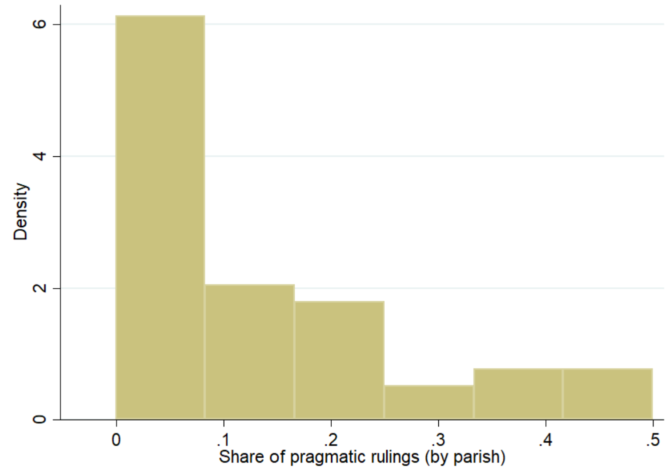
To examine the effect of legal rulings, we continue to use a DiD empirical strategy:

$$ln\left(Hearths_{ijt}\right) = \alpha_j + \delta PostFire_t + \beta PragmaticRulings_j \times PostFire_t + \gamma' X_{jt} + \epsilon_{ijt} \tag{2.11}$$

$ln\left(Hearths_{ijt}\right)$ is the log number of hearths in property i in parish j in period t. The two periods are before the Fire and after the Fire. $PragmaticRulings_j$ denotes the share of initial cases in parish j where the Fire Court judging panels' rulings were pragmatic.[16] Specifically, this is the share of cases in parish j where the Fire Court judging panels voided the existing contracts and assigned the rebuilding to the landlord. Figure 2.4 shows the distribution of the share of pragmatic rulings across the parishes. $PostFire_t$ is an indicator variable for the period after the Fire. X_{jt} is a vector of controls. Finally, α_j are parish fixed effects. We cluster the standard errors at the parish level.

[16] As we only have data from the first four out of nine volumes of the Fire Court cases, these figures are calculated based on the first four volumes.

Figure 2.4: Distribution of the share of pragmatic rulings across parishes



It is important to note that we are not able to distinguish in the data properties that went to the Fire Court and those that did not. The regression therefore includes all properties in the parish – those that went to the Fire Court and those that did not. However, this should not affect our results substantially since the proportion of properties in each parish that went to the Fire Court is a relatively small number. Based on the initial cases, the average proportion of properties in each parish that went to the Fire Court was 6%, the median was 4% and the maximum was 30%. In table F1 we report the proportion of properties in each parish where the landlord and the tenant went to the Fire Court (based on the data that we have). Therefore, in the regression, β also tells us whether the rulings in a small share of properties in the parish that went to the Fire Court affected the quality of other properties in the parish.

For those interested in the cross-sectional regressions in each time period, the results are reported in table F9. In the pre-Fire period, the number of hearths per property in parishes where all the legal rulings were pragmatic versus parishes with no pragmatic legal rulings was statistically indistinguishable.

Recall that in the previous section, our DiD regression compares burned parishes to unburned parishes. As a result, there could be concerns that any positive effect is purely mechanical since rebuilt properties had to be built according to strict regulations that specified the size and materials used. However, in this section, since our sample consists only of burned parishes, the DiD strategy helps to net off these mechanical effects. This allows us to more cleanly attribute the effect that we estimate to legal rulings.

2.5.5 Results and discussion

Higher quality structures. Table 2.10 reports whether the rulings in a small share of properties in the parish that went to the Fire Court affected the quality of other properties in the parish (i.e., number of hearths per property). The estimate in column 1 shows that controlling for parish and year fixed effects, parishes where all the initial cases saw pragmatic Fire Court rulings experienced a highly statistically significant increase of around 144% more hearths compared to parishes where all the initial cases saw unpragmatic rulings.

Importantly, the share of pragmatic rulings differ across parishes because of both exogenous and endogenous reasons. Therefore, in our regression, we control for as many endogenous reasons as we can. In column 2, we include a series of parish controls interacted with $PostFire_t$. These include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. This helps to address concerns that our results could be driven by the politics and resources of the parishes. Reassuring, our results remain extremely stable to the inclusion of these controls.

In column 3, we show that the results are stable to the inclusion of broader locations-by-post fixed effects. Finally, in column 4, we include pre-Fire hearth terciles-by-post fixed effects. The estimated effect continues to be robust to the inclusion of these controls. In particular, parishes where all the initial cases saw pragmatic Fire Court rulings experienced a highly statistically significant increase of around 98.1% more hearths compared to parishes where all the initial cases resulted in unpragmatic rulings. Given that the average number of hearths before the Fire was 3.83, this translates to an increase of 3.76 hearths. Expressed in a different way, what our result suggests is that in terms of the share of pragmatic rulings, going from the 25th percentile parish (0% pragmatic rulings) to the 75th percentile parish (20% pragmatic rulings) resulted in a 19.6% increase in the number of hearths. In absolute terms, this corresponds to an increase of 0.75 hearths. Figure F3 shows the binned scatter plot of the results in column 4.

Table 2.10: Effect of legal rulings on the number of hearths per property

	(1)	(2)	(3)	(4)
VARIABLES		ln(No. hearths)		
Pragmatic X Post Fire	1.444***	1.253***	1.103***	0.981***
	(0.449)	(0.393)	(0.286)	(0.246)
Observations	31,582	31,582	31,582	31,582
R-squared	0.014	0.024	0.026	0.031
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	46	46	46	46

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The results show us that while only a small share of properties in each parish went to the Fire Court, the rulings of these few cases had an outsized effect on the quality of other buildings in the parish. Why would this be the case? We argue that this is because the small share of cases was enough to anchor expectations. This positive result is even more remarkable considering that the hearth tax was introduced in 1662. Occupants of properties would have been incentivized to rebuild with fewer hearths to avoid the tax. Despite this, we still see a positive effect due to legal rulings. This suggests the possibility that even before the Fire, there was latent demand for structures with more hearths. Consequently, the simultaneous building after the Fire led to greater cross-building spillovers and helped to address this demand. Our results hence provide us with evidence that pragmatic legal rulings can indeed anchor expectations of what others will do.

Finally, our regression uses parish-level variations in Fire Court rulings. A natural question to ask is why do the Court rulings in your own parish matter? We take two approaches to address this. First, using the historical context. In early modern London, most interactions took place at the parish-level. Individuals often worked, lived and worshipped in the same parish. Moreover, parishes were given quite a bit of autonomy in civil matters. For example, the Highways Act 1555 made road maintenance the responsibility of the parish and Poor Relief Act 1601 (Poor Law) outlined the responsibility of the parish to look after its own poor. Therefore, because of the context, we argue that what is most salient to inhabitants of the parish is what happens within their own parish.

Second, we show statistical evidence that the rulings of previous cases in your own parish predicts future rulings in your parish. To show this, we run the following regression using the Fire Court cases of parishes that appear in the hearth tax data:

$$PragmaticRuling_{ijp} = \theta_p + \beta PragmaticRulingFirstFewCases_j + \lambda' X_{ijp} + \epsilon_{ijp}$$

$PragmaticRuling_{ijp}$ is a dummy variable that indicates whether the judging panel p for case i in parish j decreed a pragmatic ruling. $PragmaticRulingFirstFewCases_j$ is the share of pragmatic rulings in the first few cases preceding the current case in parish j. When running the regressions, we try different definitions of "first few cases". For example, the first two cases, the first three cases, etc. Taking the average across the first few cases accounts for the fact that the first case may not be precedential and precedents may take some time to be firmly established. θ_p are judging panel fixed effects. X_{ijp} is a vector of controls. These include pre-Fire case characteristics such as the degree of subletting in the property, the number of years left in the tenancy, the rent, the fine paid to secure the contract and whether the tenant spent any money to improve the property. Importantly, the vector of controls also includes the share of pragmatic rulings in other parishes before case i in parish j. The standard errors are clustered at the parish level.

Table 2.11 reports the results. In column 1, the definition of first few cases is the first case, in column 2, the definition of the first few cases is the first two cases, and so on in the other columns. Across all columns, the coefficient estimate of β is positive. This suggests that past rulings in your own parish predicts future rulings. In addition, the coefficient estimates increase as we move across the columns. This reflects the fact that the first case may not be precedential and precedents may take some time to be firmly established. Column 5 suggests that if the first five cases in your parish had pragmatic rulings, the probability that the current case receives a pragmatic ruling increases by 40.5%-points. Therefore, the Court rulings in your own parish matter because they predict future rulings in your parish.

Table 2.11: Effect of past rulings in your own parish on current ruling

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Whether current case ruling is pragmatic				
Share of pragmatic rulings (first few cases)	0.024	0.022	0.202*	0.327**	0.405*
	(0.061)	(0.075)	(0.117)	(0.144)	(0.199)
Observations	303	246	195	163	139
R-squared	0.350	0.417	0.443	0.510	0.488
Judging panel FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
First few cases	1	2	3	4	5

Notes: Case controls include pre-Fire case characteristics such as the degree of subletting in the property, the number of years left in the tenancy, the rent, the fine paid to secure the contract and whether the tenant spent any money to improve the property. It also includes the share of pragmatic rulings in other parishes before the current case. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Effect holds when controlling for spillovers from neighboring parishes. Next, we check if our results are sensitive to potential spatial spillovers. In particular, burned parishes with Fire Court cases tend to be located near each other. Therefore, a burned parish with Fire Court cases not only generated spillovers to other burned parishes but also received inward spillovers from these other parishes. If these spillovers are large, our estimated effects of legal rulings within each parish could be overstated. Therefore, we include as a control the weighted share of cases in neighbouring burned parishes where the Fire Court judging panels decreed pragmatic rulings. Table 2.12 shows that our results are robust to controlling for the legal outcomes in neighbouring parishes.

Table 2.12: Effect of legal rulings on the number of hearths per property (controlling for spillovers)

	(1)	(2)	(3)	(4)
VARIABLES		ln(No. hearths)		
Pragmatic X Post Fire	1.485***	1.149***	1.119***	0.963***
	(0.478)	(0.326)	(0.278)	(0.253)
Pragmatic Spillover X Post Fire	-0.513	0.230	0.013	0.349
	(0.635)	(0.358)	(0.315)	(0.286)
Observations	31,582	31,582	31,582	31,582
R-squared	0.015	0.025	0.026	0.031
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	46	46	46	46

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. We also include the log number of cases in neighbouring parishes. Standard errors are clustered at the parish level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Adding in other dimensions of the rulings as controls. The Fire Court judges were given the power to decree who rebuilds, the rent, as well as length of the new contract. Such multi-dimensional rulings make it difficult to define what constitutes pragmatic rulings that helped to facilitate the rebuilding of London. In order to circumvent the issue of multi-dimensional rulings, in our analysis, we had focused on the most extreme of pragmatic case outcomes – cases where the Fire Court voided existing contracts and assigned the rebuilding to the landlord. Nonetheless, it could well be the case that the newly decreed rent or tenancy length could be playing a role in anchoring expectations and consequently, the number of hearths per property in each parish.

To address this concern, we include as controls the average change in rent and tenancy length in each parish arising from the Fire Court rulings. For example, if there were five cases in parish A that went to the Fire Court, and the judging panel increased the tenancy length for all five of the cases by 10 years, then the average change in tenancy length arising from the Fire Court rulings for parish A would be 10 years. Referring to table F10, the coefficient estimate on our treatment variable $PragmaticRulings_j \times PostFire_t$ remains stable to the inclusion of the other dimensions of the Fire Court's rulings. The coefficient estimates on the average change in the tenancy length interacted with post are extremely close to zero and statistically insignificant. The coefficient estimates on the average change in the rent interacted with post while significant, is relatively small in magnitude. Taken together, the results across the columns provide evidence that it is indeed pragmatic Fire Court rulings (as defined by the share of initial cases where the judging panels voided the contracts and assigned the rebuilding to the landlord) that are affecting individuals' expectations and not the other dimensions of the Fire Court's decisions.

2.5.6 Competing hypotheses/mechanisms

The results in table 2.10 show us that while only a small share of properties in each parish went to the Fire Court, the rulings of these few cases had an outsized effect on the quality of other buildings in the parish. We argue that this is because the small share of cases was enough to anchor expectations. Are we able to rule out competing hypotheses/mechanisms?

One competing hypothesis is that our results have nothing to do with the small share of cases anchoring expectations. Instead, it is simply picking up the direct effect of the Fire Court rulings. This is because we are not able to distinguish in the data properties that went to the Fire Court and those that did not. The regression therefore includes all properties in the parish – those that went to the Fire Court and those that did not. To see why this might be a problem, consider the following example of a parish where there are 100 properties (see table 2.13). Before the Fire, the average number of hearths per property was 10. Now assume that of the 100 properties, 60 did not go to Court but 40 went to Court. Let us further assume that the average number of hearths in the 60 properties was the same before and after the Fire (i.e., 10 hearths). However, in the 40 cases that went to the court, the average number of hearths increased by three per property to 13 hearths. Consequently, the overall average number of hearths per property increased by 1.2 to 11.2. This stylized example shows us how the average number of hearths can increase even without any anchoring of expectations of those that did not go to Court.

Table 2.13: Stylized example

	Total Hearths	Number of properties	Avg. hearths per property
Before Fire	100*10=1,000	100	10
After Fire	(60*10)+(40*13)=1,120	100	11.2

However, we think that this should not affect our results substantially since the proportion of properties in each parish that went to the Fire Court is a relatively small number. Based on the initial cases, the average percentage of properties in each parish that went to the Fire Court was 6%, the median was 4% and the maximum was 30%. In addition, table 2.14 shows the results when we drop all parishes where more than 14.6% (column 2), 7.4% (column 3), 4.1% (column 4) and 2.3% (column 5) of the properties went to the Fire Court. 14.6%, 7.4%, 4.1% and 2.3% correspond to the 95th, 75th, 50th and 25th percentiles respectively.

Table 2.14: Effect of legal rulings on the number of hearths per property (dropping parishes where a "large" proportion of properties went to Court)

	(1)	(2)	(3)	(4)	(5)
VARIABLES	ln(No. Hearths per Property)				
Pragmatic X Post Fire	0.981***	0.893***	0.897***	1.051***	1.121*
	(0.246)	(0.241)	(0.245)	(0.232)	(0.527)
Observations	31,582	30,993	29,210	26,289	21,953
R-squared	0.031	0.032	0.033	0.035	0.042
Parish FE	✓	✓	✓	✓	✓
Post FE	✓	✓	✓	✓	✓
Parish controls X Post FE	✓	✓	✓	✓	✓
Broader location X Post FE	✓	✓	✓	✓	✓
Pre-fire hearth tercile X Post FE	✓	✓	✓	✓	✓
Number of clusters	46	43	34	24	12
Sample	All	¡ 95 pct	¡ 75 pct	¡ 50 pct	¡ 25 pct

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Column 1 shows the results using the full sample. Our results remain robust to dropping parishes where a "large" proportion of properties went to the Fire Court. If anything, our results seem to get bigger when we drop more parishes which is against what we should see if our results are purely picking up the direct effect of the Fire Court rulings.

Another competing hypothesis is that of the share of owner-occupied properties in the parish. In parishes where there is a large share of owner-occupied properties, the share of properties that go to Court must by definition be small since owners cannot sue themselves. Therefore, our results may have nothing to do with the small share of cases anchoring expectations. Instead, they could simply be reflecting the fact that these parishes have a greater share of owner-occupied properties. Landowners are less likely to be credit constrained and can thus allocate more resources towards building more hearths per property. In addition, if the owner occupies the property, then there is a clear assignment/alignment of property-rights. Consequently, the owner builds at a higher quality because the owner is able to accrue the full benefits of living in a higher quality property.

Unfortunately, the data do not tell us whether a property is owner occupied – they only tell us who the main occupant of the property is. To overcome this limitation, we count the share of peers, high-ranking military personnel and doctors living in the parish. To the extent that this group of individuals are more likely to own their own homes, this variable gives us a proxy for the share of owner-occupied properties in the parish. We then include these three variables as controls in our regression. Referring to table 2.10 column 2, we can see that the coefficient estimate on $PragmaticRulings_j \times PostFire_t$ remains stable to the inclusion of these variables as controls. This suggests that there are aspects of the rulings in the small share of properties that went to Court that cannot be attributed to the share of owner-occupied properties in the parish. Therefore, our results lend credence to our proposed mechanism that the small share of cases was enough to anchor expectations for everyone in the parish.

2.5.7 Robustness checks

Dropping parishes which merged after the Fire. One concern could be that the Fire led to the merging of some parishes and so our results could be driven by these enlarged parishes which might have more resources. In table F11 we re-run equation (2.11) using only parishes that did not merge after the Fire. Reassuringly, the coefficient estimates remain similar and even larger than our baseline results, suggesting that our baseline results are conservative.

Accounting for zeros in the outcome variable. For burned parishes with Fire Court cases, there are 801 observations which are recorded as having zero hearths in the property. Taking logs results in these observations dropping out of the regression. Therefore, to account for the zeros in the outcome variables, we adopt two approaches. First, applying the inverse hyperbolic sine transform to hearths. Second, using a Poisson pseudo-likelihood (PPML) regression. The results are reported in table F12 and table F13 respectively. The estimated effects are very similar to our baseline results.

Trimming extreme values of the outcome variable. Another robustness check that we run is to drop the top and bottom 1 percentile of $ln\left(Hearths_{ijt}\right)$. The results of this robustness check is reported in Table table F14 and are similar to our baseline results.

2.5.8 Using an IV estimation strategy

As there could be concerns that there are time-varying parish-level omitted variables which we have not controlled for, we augment our DiD strategy with an instrumental variable (IV) strategy. Our IV DiD strategy exploits the fact that Fire Court judging panels with different political alignments (i.e., whether they were predominantly Royalists or Parliamentarians) were assigned to the cases in the parishes. The 1666 Great Fire took place in the midst of the Second Dutch War (1665-1667) and the Great Plague which began in 1665. King Charles II was relying on taxes and loans from London and its wealthiest citizens to finance the war. The destruction of the customs house, wharves and more than 13,000 buildings caused a significant drop in royal revenue and thus the King had a vested interest for London to be quickly rebuilt. Therefore, judging panels that were predominantly Royalists (more aligned with the King) were more likely to decree pragmatic rulings so as to facilitate the rebuilding of London. As a result, we can use the composition of the judging panels as an instrument for the share of initial cases in the parish that had pragmatic rulings. This gives us exogenous variations in legal rulings for each of the parishes. Of the 46 parishes in our regression with Fire Court cases, 17 of them (37.0%) had the majority of their initial cases presided by judging panels that consisted predominantly of Royalists.

Relevance of instrument

We estimate the following regression to examine the first-stage relationship between the composition of the judging panels in the initial cases and the share of initial cases in the parish that had pragmatic rulings:

$$PragmaticRulings_j \times PostFire_t = \alpha_j + \delta PostFire_t + \beta MajorityRoyalist_j \times PostFire_t + \gamma' X_{jt} + u_{ijt} \tag{2.12}$$

Table 2.15 presents the first-stage results which suggest that if the majority of the initial cases in the parish were heard by judging panels that were predominantly Royalists, then the share of initial cases in the parish that had pragmatic rulings increased by 9.5%-pts.

Table 2.15: First-stage – Effect of Royalist on legal rulings

VARIABLES	(1)	(2) Pragmatic X Post	(3)	(4)
Majority royalist in judging panels X Post	0.119***	0.087*	0.095***	0.095***
	(0.038)	(0.045)	(0.031)	(0.029)
Observations	31,582	31,582	31,582	31,582
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	46	46	46	46
KP F-stat	9.895	3.795	9.122	10.37

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

The first-stage relationship is robust to the inclusion of a series of parish controls interacted with $PostFire_t$, broader locations-by-post fixed effects and pre-Fire hearth terciles-by-post fixed effects. In addition, in most of the regressions, the first-stage has a KP F-statistic value of around 10. Figure F4 shows the binned scatter plot of the results in column 4.

Validity of instrument

Conditional Independence. The validity of our instrument depends crucially on whether there were other parish-level factors involved in determining the composition of the judging panels in the initial cases. We verify this by running a balancing test. This is similar to the type of statistical test that is used to verify random assignment in a randomized controlled trial. Table 2.16 shows the result of this balancing test. The coefficients on the share of peers, high-ranking military personnel and doctors seem sizable. However, this is because the mean values of these variables are extremely small. For example, the mean value for the share of peers is 0.005, that of high-ranking military personnel is 0.0004 and that for doctors is 0.003.

Table 2.16 shows that parish-level factors were not predictive of the composition of the judging panels in the initial cases. In column 4, for parishes where data were available, we included the average rent in the parish in 1638. The 1638 rental data comes from "The Inhabitants of London in 1638".[17] Column 4 shows that historical rents were not predictive of the composition of the judging panels. Importantly, across all of the columns, all of the estimates are statistically insignificant at the 1% level and are not jointly significant with p-values ranging from 0.84 to 0.93.

Table 2.16: Testing for random assignment of judging panels to parishes

	(1)	(2)	(3)	(4)
VARIABLES	Majority royalist in judging panels			
ln(No. properties before the Fire)	-0.033	0.004	0.034	0.029
	(0.064)	(0.086)	(0.095)	(0.133)
Share of peers	-2.071	-3.816	-6.501	-4.256
	(9.929)	(10.247)	(11.106)	(23.191)
Share of high-ranking military personnel	-30.800	-7.468	-10.069	-21.936
	(29.296)	(42.349)	(51.972)	(81.006)
Share of doctors	1.689	7.120	2.283	-7.803
	(17.258)	(19.069)	(19.263)	(27.936)
Broader location 1		-0.097	-0.075	0.098
		(0.324)	(0.322)	(0.494)
Broader location 2		0.161	0.169	0.220
		(0.288)	(0.300)	(0.518)
Pre-Fire hearth tercile 1			-0.177	-0.315
			(0.236)	(0.292)
Pre-Fire hearth tercile 2			-0.049	-0.063
			(0.210)	(0.250)
ln(Average rent in 1638)				-0.175
				(0.289)
Observations	46	46	46	37
Adjusted R-squared	-0.081	-0.109	-0.152	-0.259
F-stat for joint test	0.350	0.431	0.373	0.361
p-value for joint test	0.843	0.853	0.928	0.932

Notes: Robust standard errors. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Exclusion. This restriction requires that the composition of the judging panels in the initial cases affected the quality of building investment in the parish only through its effect on legal rulings. While it is not possible to formally test the exclusion condition, the fact that our instrument passes the balancing test is reassuring.

However, there could still be concerns that the exclusion restriction could be violated since it is possible that monarchist officials may had some influence in the assignment of judges to the cases. For example, these officials could be

[17] "The Inhabitants of London in 1638" was originally published by the Society of Genealogists in London in 1931 (see Dale (1931)) and can be accessed at the British History Online website. This publication was based on the manuscript "Settlement of Tithes, 1638", found in the Lambeth Palace Library. The manuscript contains a list of the householders in 93 out of the 107 parishes in the City of London, as well as the rentals paid for the houses and the tithes paid.

expecting some parishes to grow more, and so they wanted to make sure that the parliamentarian judges did not derail their plans. In addition, it could be the case that some parishes had more monarchist landowners and so the officials assigned monarchist judges to protect the interest of these landowners.

In order to address such concerns, we are in the process of collecting data to measure what is the share of peers in each parish that was loyal to the King. From the hearth tax data, we are able to identify the names of the peers (i.e., Duke, Duchess, Marquess, Marchioness, Earl, Countess, Viscount, Viscountess, Baron, Baroness, Lord, Lady, Sir, Dame and Ambassador) living in each parish. For the parishes in our regression, there are about 500 peers in total. We can then refer to various historical sources (e.g., the dictionary of national biography (2014)) to find out what were the views of these peers on the 1660 restoration of the monarchy. Once we have determined the share of monarchist peers in each parish, we can then include this variable as a control in the regression. This would hopefully help to control for the possibility that monarchist judges were being assigned to parishes where the monarchists wanted to have greater influence over or benefit more from.

In any case, even if the exclusion restriction is violated, our reduced form estimates can still be interpreted as the causal effect of the composition of the judging panels in the initial cases on the number of hearths per property in the parish.

Monotonicity. The monotonicity assumption requires a monotonic relationship between the instrument and the variable that is being instrumented. The monotonicity assumption ensures that our IV estimate can be interpreted as a local average treatment effect (LATE). In our context, this is the average causal effect among the subgroup of parishes that invested differently in their buildings because of the composition of the judging panels in the initial cases.

If the monotonicity assumption is violated, then in the classical IV framework, our results can only be interpreted as causal **constant** effects. On the other hand, in a heterogeneous treatment effects framework, if the monotonicity assumption is violated, Angrist et al. (1996) and Heckman and Vytlacil (2005) show that the IV estimates would still be a weighted average of marginal treatment effects. However, because the weights do not sum to one, this leads to an ill-defined local average treatment effect.

One testable implication of the monotonicity assumption is that the first-stage estimates should be non-negative for any subsample. To test this, we split the sample into various subsamples and estimated the first-stage relationship for each of these subsamples. The results are reported in table F15. In columns 1 and 2, we split the sample into whether the church in the parish was damaged by the Fire. In columns 3 to 5, we split the sample into three broader geographical locations (i.e., abutting the City of London walls, within the walls and outside the walls). In columns 6 to 8, we split the sample based on terciles of the number of hearths in each parish before the Fire. In columns 9 to 11, we split the sample based on terciles of the number of properties in each parish before the Fire. In columns 12 to 14, we split the sample based on terciles of the share of peers in the parishes. In columns 15 and 16, we split the sample into two based on the share of doctors in the parishes. Finally, in columns 17 and 18, we split the sample into two based on the share of high-ranking military personnel in the parishes. Out of these 18 subsamples, there are only three subsamples where the first-stage estimate is negative. In the other 15 subsamples, the first-stage estimates are positive, consistent with the monotonicity assumption.

IV results and discussion

Table 2.17 reports the results from the IV DiD regressions. In column 4, the results suggest that parishes where all the initial cases saw pragmatic Fire Court rulings experienced a highly statistically significant increase of around 200% more hearths compared to parishes where all the initial cases saw unpragmatic rulings. In other words, going from the 25th percentile parish in terms of the share of pragmatic rulings (0% pragmatic rulings) to the 75th percentile parish (20% pragmatic rulings) resulted in a 40% increase in the number of hearths. In absolute terms, this corresponds to an increase of 1.53 hearths. For completeness, we also report the reduced form estimates in table F16 and show the associated binned scatter plot of the residues (based on all the controls) in figure F5.

Table 2.17: IV – Effect of legal rulings on the number of hearths per property

	(1)	(2)	(3)	(4)
VARIABLES		ln(No. hearths)		
Pragmatic X Post Fire	2.276***	2.550***	2.247***	2.001**
	(0.698)	(0.917)	(0.753)	(0.789)
Observations	31,582	31,582	31,582	31,582
R-squared	0.010	0.016	0.020	0.026
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	46	46	46	46
KP F-stat	9.895	3.795	9.122	10.37

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

The IV results are around twice as large as the DiD results (1.53 hearths vs. 0.75 hearths). This could suggest two things. First, the DiD regression suffers from omitted variables and thus fails the parallel trend assumption. Second, even if the DiD estimate is unbiased, we should still expect the IV estimate to be different from it. This is because the IV estimate gives us the local average treatment effect for the compilers. Nevertheless, the fact that the IV estimates are positive and highly statistically significant re-affirms our DiD results. This gives us greater confidence in concluding that pragmatic legal rulings affected individuals' expectations about how much other individuals in their parish would be investing. This in turn resulted in a higher number of hearths per property in the parish.

Adding in other dimensions of the rulings as controls. To address concerns that our definition of pragmatic rulings fails to consider changes in rent and tenancy length, we include these as controls in our regression. Table F17 provides evidence that it is indeed pragmatic Fire Court rulings that are affecting individuals' expectations and not the other dimensions of the Fire Court's decisions.

Robustness checks

Dropping parishes which merged after the Fire. Similar to the other sections, in table F18 we re-run our IV analysis using only parishes that did not merge after the Fire. The coefficient estimates remain very similar to our baseline results.

Accounting for zeros in the outcome variable. To account for the zeros in the outcome variables, we apply the inverse hyperbolic sine transform to hearths. The results are reported in table F19. The estimated effects remain positive and continue to be highly statistically significant.

Trimming extreme values of the outcome variable. Another robustness check that we run is to drop the top and bottom 1 percentile of $ln\left(Hearths_{ijt}\right)$. The results of this robustness check is reported in table F20 and are similar to our baseline results.

2.6 Conclusion

The development of cities often involves the rejuvenation and replacement of outdated buildings. However, the sunk cost of existing durable structures often serves as an impediment. While disasters are destructive, an unintended silver lining is that they may help to remove development frictions. By lowering the opportunity cost of waiting to rebuild to zero, disasters could potentially spur the development of neighbourhoods and even cities. However, disasters do not necessarily guarantee higher quality buildings. What ultimately matters is what each individual expects other individuals to do. Our paper highlights this by providing causal evidence of how legal rulings can be a main driver in the formation of these expectations. While there is a relatively large theoretical literature on how legal institutions can affect expectations and hence the behaviour of individuals, there is relatively less empirical work on this. Our paper thus addresses this gap in the literature. Although the setting of our paper is 17th century England, even today, legal rulings continue to be a key aspect in society. This has policy implications as it suggests scope for laws to influence expectations and in so doing, facilitate the continual development of cities.

Chapter 3

Jury Composition and Selection

3.1 Introduction

Jury systems are typical for trials of serious crimes in Common Law countries, and jury selection is an important component of such systems. To ensure defendants are judged by unbiased panels of their peers, some randomness is typically introduced to select jurors. However, this introduces variance in verdicts because jurors can be different on many dimensions, and some may be intractably biased. Such biased jurors may have an outsized influence especially in systems that require unanimous verdicts.

Peremptory challenges (strikes) were introduced to tamper such variability by allowing attorneys, on either side, to veto potential jurors without the need for any justification. This allows the attorneys to remove extreme biased jurors and tamper variability caused by their inclusion.[1] Strikes also allow attorneys to eliminate non-extreme jurors who tend oppose their legal interests, and this may bias the composition of the final jury.

It is well documented that observable characteristics like gender (Anwar et al. (2016), Hoekstra and Street (2021)), race (Anwar et al. (2012); Flanagan (2017)) and age (Anwar et al. (2014)) can affect jury verdicts. Race, in particular seems to be an important determinant of striking behaviour and it is well documented

[1]There is a debate on whether attorneys are skilled at using information gathered during jury selection (e.g. verbal answers to questions, body language) to remove biased jurors. Morrison et al. (2016) demonstrated that legal professions were adept at identifying implicit racial bias simply from written answers to typical question posed during jury selection.

that prosecutors remove black jurors at higher rates than non-black jurors in civil trials (Diamond et al. (2009)), felony trials (Noye (2015); Wright et al. (2018)) and capital trials (Grosso et al. (2013); Baldu et al. (2001)).[2]

However, it is unclear whether strikes are directly responsible for biases in verdict. Diamond et al. (2009) found that opposing striking by the attorneys cancelled themselves out and produced no impact on the racial make up of the jury. Flanagan (2017) also found that the proportion of final black jurors was not different from that of jurors randomly picked from the pool; but black **male** jurors were under-represented in the final jury because prosecutors stuck them aggressively.[3] This points to a more general point that even if striking does not cause bias in observable characteristics, they might cause bias in unobservables which may favour one side.

While there is a substantial literature analysing potential biases induced by strikes (Barrett (2007)), there is little counterfactual analysis on what happens when we change the allocation of strikes. Flanagan (2015) suggests that, **theoretically**, having strikes may bias conviction rates towards prosecutors and increase the likelihood of extreme juries (increased variance). This is because any variance in the composition of the pool of potential jurors may be exacerbated by striking strategies. For instance, Anwar et al. (2012) show that having at least 1 black potential juror in the pool reduced the conviction rate of black defendants by 16 percentage points (pp).[4] One explanation for this is that prosecutors consider the expected type of replacement juror, when she chooses whether to strike a juror seated for the voir dire. The presence of a black in the pool changes this expectation and affects **all** her striking decision.

How the allocation of strikes affects the outcomes of jury trials is a salient issue because of the substantial variation in strike allocation across across and within countries. For instance, although peremptory challenges had been allowed in jury trials in England, it has been abolished by the Criminal Justice Act of 1988. On the

[2]The opposite is also true: that defence attorneys remove white jurors at higher rates than black jurors.

[3]The pool typically consists of 50 residents of a county, who were randomly summoned for jury duty. The final jury is selected from this pool using a jury selection system. Refer to section 3.2 for more details.

[4]Note that the jurors in the pool may not even be picked to be jurors seated for the voir dire (a process where the attorneys question potential jurors to identify any biases they may have). Jurors seated for the voir dire are picked from the pool, some may be struck by the attorneys. Refer to section 3.2 for more details.

other hand, nearly all states in the United States allow peremptory challenges but the number of strikes allocated differ across states (e.g., for misdemeanour trials the state of Texas allocates 5 strikes for each party, while the state of California allocates 6 strikes) and within states (e.g., 10 strikes are allocated for felony trials in Texas). Furthermore, the allocation of strikes remains a live policy issue: the state of Arizona abolished the use of peremptory strikes from jury trials from 1 January 2022 and the state of California reduced the number of strikes for misdemeanour trials from 10 to 6 in 2015. To better understand the implication of the variation in the allocation of strikes and policies that alter them, I build and estimate a model that allows me to conduct counterfactual analysis on how increasing the number of allocated strikes biases in favour of one side.

In estimating the model, I use data from North Carolina (NC) collected by the Jury Sunshine Project (Johnson (2018)). Rich sets of data on jury selection are archived and considered public information in NC, which is not the case in many other states. Besides data availability there are a few key advantages with using this data: (i) NC is the 9th most populous state in the US; (ii) NC has a jury selection process that is quite standard and widely applied across other states; (iii) NC has cross county heterogeneity in demographics, which makes the results more generalisable. I discuss the process of selection and the data in sections 3.2 and 3.3.

A key limitation of the data is the lack of information on jurors in the pool who were not seated for the voir dire (a process where the attorneys question potential jurors to identify any biases they may have). Conventionally, NC court clerk only recorded the names of jurors who were 'ever seated' for the voir dire (whether or not they struck after the voir dire) but not those in the pool who were never called to be seated for the voir dire. Studies estimating the impact of jury composition on verdicts typically exploit the random composition of the pool. Flanagan (2017) used the composition of jurors who were 'ever seated' for the voir dire as a measure of the composition of the pool (e.g. proportion of blacks in 'ever seated'$=\frac{\text{no of 'ever seated' black jurors}}{\text{no of 'ever seated' jurors}}$). This variable is difficult to interpret because the composition of a group in 'ever seated' for the voir dire could increase by randomly drawing more of them from the pool and/or by striking more of them when seated. The variable is also potentially endogenous to verdicts because the number of draws from the pool is related to the number of strikes used, and attorneys' use of strikes may be related to details of the case. For instance, prosecutors may strike blacks more aggressively if evidence in favour of conviction is weak. If so then, the 'ever seated' composition would have a higher proportion of blacks who were struck. [5]

While the goal of this paper is to study the impact of the allocation of strikes on the outcome of jury trials, to understand the context of jury trials in NC and to appropriately model the striking behaviour in these trials, I also replicate and try to improve upon Flanagan (2017). Hence, I propose an alternative variable: composition of the first 12 jurors seated for the voir dire (first draw). The first draw is random and made before any decision to strike and hence should not be endogenous. However, the first draw is random conditional on the composition in the pool; but I do not observe pool composition. As a second best solution I controlled for years, seasons, counties and judges and show that the composition of first draw was uncorrelated with other variables related to the trial. I discuss identification issues in section 3.9. Hoekstra and Street (2021) exploit a similar source of variation to identify the impact of own-gender jurors (jurors with the same gender as the defendant) on conviction rates. Instead of using the gender composition of the jury pool, the authors constructed the expected number of own-gender jurors using the **ordering** of potential jurors from the pool. They found a large impact of jurors' characteristics (in this case the expecter number of own-gender jurors) on conviction rates. This finding is consistent with studies that exploit the variation in the composition of the jury pool (e.g., Anwar et al. (2012)) and suggests that, while it would be ideal to have data on the composition of the jury pool, my approach is not invalid.

In section 3.5 I discuss my reduced form results. I find that attorneys struck racial groups at different probabilities. Prosecutors were twice more likely to strike blacks than whites, while defence attorneys four times more likely to strike white than blacks. This difference was more stark when the defendant was black. I also find that using the composition of the first draw attenuated the findings from using 'ever seated'. Specifically, an increase in the proportion of blacks in 'ever seated' from 0 to 1, decreased the probability of conviction by 36.9pp. The same change in composition of first draw reduced the probability by 9.8pp. These negative relationships hold only when the defendant was black.

[5] I use the term 'ever seated' to refer to jurors who were 'ever seated' for the voir dire and not the final seated jurors.

In section 3.6 I model the striking behaviour of the attorneys as a game between two players. Consistent with the reduced-form results, the key variables in the model are: (i) the race of jurors (black/non-black); and (ii) the race of defendants (black/non-black). While I was able to study the impact of race and gender in the reduced form regressions, I focus on the impact on race in the model to keep it parsimonious for computational tractability. The results from my model suggest: (i) that the defence benefits from more black jurors; (ii) that the game is one of strategic complementarity and (iii) under strong strategic complementarity, the defence can benefit from being allowed more strikes because she can use them to restore any disadvantage from the prosecutor's strikes.

Finally, I discuss further work to be done in section 3.7.

3.2 Jury selection in North Carolina

The office of the clerk of courts oversees the jury selection process; they ensure randomness in three key stages of jury selection. First, they issue summons randomly to eligible residents of the county and ensure a sufficient number turn up. Second, on the day of selection, they randomly allocate 50 jurors to each court case. Third, they ensure jurors seated for the voir dire are randomly picked from the pool.

3.2.1 Constructing a jury pool

Eligible residents who are called up may defer their duties or simply not turn up. This selection, out of jury duty, will be unrelated to the case because such information are not in the summon letter. However, the process may cause systematic under representation of certain subgroups.

To qualify as a juror, residents must not have been convicted or pleaded guilty to a felony, unless citizenship has been restored. So with a higher proportion of ex-convicts under parole, post-release supervision or probation will be under-represented.

Furthermore, the database for issuing summons is based on driving license records and state ID lists from voting records. These lists would under represent subgroups with lower voting and driving rates. For instance, in Burleigh county, North Dakota, Native Americans, who have lower driving rates, make up 4.2% of the population but only 1.9% of summoned jurors.

Finally, juror pay is notoriously low. The current rate is $12 for the first day of service, and $20 thereafter ($40 if one serves more than five days). $40 is equivalent to 5.5 hours of work at the state minimum wage, which suggests even minimum wage workers can earn more from avoiding jury duty. Hence, there are strong incentives for low-income households (i.e. near the poverty line) to avoid jury duty.

In my dataset, 15.9% of jurors seated for the voir dire were black, which is lower than 21.5% of blacks in NC from the 2010 Census. Hence, there may be an under representation of blacks in the jury pool.

Residents who turn up on the day indicated on the summon are then randomly allocated, in groups of 50, to individual courtrooms for selection (i.e. trial-specific jury pool). It is possible that jurors may opt out of certain pools - for instance, jurors may need to wait around for a longer time for a selection that happens later in the day. Again, this selection should be unrelated to the case because case information is not known at this point.

3.2.2 Selecting the final jury

The actual details of jury selection vary by state but NC courts use the version most common across states:

1. Judge or prosecutor briefs people in pool about broad case details.

2. 12 jurors are randomly selected, by the clerk, from the pool to be seated for the voir dire. I call this step the **first draw**.

3. Attorneys and judge direct questions at all jurors seated for the voir dire to ascertain their suitability for trial.

4. Judge can excuse jurors for cause. Attorneys can use peremptory strikes. I define jurors who were removed by strikes as **struck jurors**.

5. Replacement jurors are randomly selected from the pool to be seated for the voir dire.

6. Steps 3. and 4. are repeated until no party wants to strike or attorneys exhaust their strikes.

7. Jurors that were not struck form the final jury (final jury).

To excuse for cause, the judge will have to provide a justification. These are usually related to the stated biases, stated inability to serve, or proximity to the case (e.g. relative of victim) of the juror during questioning. In principle, the judge has unlimited powers to excuse for cause as long as she is justified in doing so.

The attorneys each have a limited number of strikes. I focus on felony cases, which allow six strikes for each side. When using their strikes, attorneys do not have to provide a justification.[6].

This process of jury selection can be thought of as going through a number of rounds. In the first round, 12 jurors are randomly picked; J_1 jurors are struck or excused. In the second round, J_1 replacements jurors are randomly picked and J_2 jurors are struck or excused etc.. Once a juror survives a round, she cannot be struck or excused in subsequent rounds. In the data, only 1% of selection ended in the first round, 13% by the second round, 42% by the third round, and 69% by the fourth round.

Within each round, the composition of jurors is random. However, the **number** of replacement jurors and the number of rounds is endogenous to the case. The number of replacement jurors is partly determined by the number of strikes used. Attorneys may be more or less aggressive in their usage of strikes based on the underlying characteristic of their case. Further, the selection process is more likely to extend into later rounds the more aggressive the attorneys are in striking jurors.

[6]A landmark Supreme Court ruling in Batson vs Kentucky, 1986, ruled that peremptory challenges could not be used to remove jurors based solely on their race. This ruling introduced the possibility of opposing attorneys challenging strikes (a Batson challenge) perceived as based solely on race. The attorney will then have to justify why she used the strike; and the judge can nullify the strike if the justification is not convincing. In practice, this rule has no bite because prosecutors can pick from the myriad of non-racial justification. The University of North Carolina School of Government, in its “Indigent Defence Manual Series” has found just two successful Batson challenge since 1986.

3.3 Data

The data was collected by the Jury Sunshine Project (JSP), led by Prof Ron Wright from Wake Forest University (Johnson (2018)). The team collected jury selection data on felony trials from the entire state of NC from 2010-2012.

3.3.1 Data collection and preparation

JSP had to first request for a master list of felony trials for the year 2011, which contained case numbers, from each of the county courts in NC. They then requested for all the case files based on the case numbers in the master list. The case files were folders with paper documents stored in the court's archives. Each folder typically contained: (i) details of the case, including the indictment, verdict and sentence; (ii) demographics of the defendant, including gender, race and year of birth; (iii) names of jurors that were seated for the voir dire and whether they were struck, recorded in a standardised form; and (iii) names of the attorneys, clerk and judge. Although the jury selection process was recorded in a standardised form, some court clerks did not indicate which party removed a juror.

The names of the jurors were then matched with names from the relevant county's voting register to obtain the race, gender and political affiliation of the juror.[7] About 80% of juror names were successfully matched, and of the 20% of unmatched jurors, JSP used the 'gender' package in R to predict their gender. JSP was also able to match the names of the prosecutors and judges to public datasets to obtain their race, gender and years of professional experience.

I classified the charge information using the US Department of Justice's National Incident-based Reporting System. This system is the standard used at the local, state and federal level in the US for collecting and reporting data on crime. I also classified capital trials as those with class A felony charges (e.g. murder), which may result in a death sentence in NC.

In my analysis, I excluded capital trials and trials with multiple defendants. Such cases allow attorneys significantly more than six strikes. Both attorneys have 14 strike in capital trials, and six additional strikes for each defendant beyond the

[7] Flanagan (2017) details the exact procedure JSP used to match the jurors' names

first. The database had 1,239 trials; I dropped 49 capital trials and 40 trials with multiple defendants (1,150 trials for analysis).

The NC statues for jury selection allow each party an additional strike for each alternate juror. Alternates replace any of the final jurors who drop out during trial. In the data, the number of alternate jurors varied from one to three. This means the total number of allocated strikes, including for alternate jurors, may vary from six to nine. To avoid this complication, I restricted my analysis to the 12 main jurors.

Jury trials in NC require unanimity in any verdict, therefore there are three possible outcomes: (i) unanimous guilty verdict; (ii) unanimous acquittal or; (iii) hung jury. In my analysis I define a guilty verdict as a unanimous guilty verdict.

3.3.2 Limitations

The major limitation of the data is the lack of juror information of jurors in the pool, who were never seated for the voir dire. Anwar et al (2012) demonstrate that the composition of the pool could have huge implications on verdicts. This means pool composition will cause an omitted variable bias in regressions with verdict as the dependent variable. I try to address this in my identification strategy and discuss the possible bias induced.

Since I am unable to identify the race of 20% of jurors, there will be measurement error in any measure of race composition. It is not clear whether black jurors had lower match rates. While JSP used voter registers for matching, and blacks may be under represented in these records, the black jurors summoned by the court had to be on voter registers to be summoned in the first place.

Many cases of failed matching were due to multiple matches, and matched records having different races.[8] It is not clear, whether black or non-black juror names would have a higher probability of multiple matches.

In this paper, I focus on the composition of black jurors. Since it is not clear

[8] If a juror name was matched to multiple records, but all the match records had the same reported race, the juror was recorded to have that race.

whether the composition of black jurors should be over or under-measured, I treat this mismeasurement as a classic measurement error and expect coefficient to be attenuated towards zero.

3.3.3 Summary statistics

Table 3.1: Summary statistics of key variables

	Mean	SD	Min	Max
Black defendant	0.569	0.495	0	1
Female defendant	0.114	0.318	0	1
Prop black jurors	0.159	0.129	0	0.71
Prop female jurors	0.510	0.109	0.15	0.89
Prosecutor strikes	2.037	1.759	0	6
Defence strikes	2.867	2.127	0	6
Excused for cause	1.874	2.624	0	25
Guilty verdict	0.672	0.470	0	1

Table 3.1 summaries the key variables in the data. For comparison, according to the 2010 US census, 21.5% of the population of NC was black or African American, and 51.3% was female. This suggests that among defendants, blacks were over represented at 56.9% and females were under represented at 11.4%. Among jurors, blacks were under represented at 15.9%. There was significant variation in the jury composition.

On average, defence attorneys used more strikes than prosecutors; and more jurors were struck by either attorney (4.9) than excused for cause (1.9). However, the standard deviation (SD) of the number of juror excused for cause was higher than the use of strikes by either party. This suggests significant differences in how judges choose to excuse jurors. In order to account for this variation, I use judge-specific dummies in my key regressions.

Among the 1,150 trials, I had 1,128 with charge details (table 3.2). The most common charges were drug related (18%), assault (15%), burglary (13%) and theft (9%).

Table 3.2: Charge details

	N	Percent
Assault	167	15
Burglary/B&E	141	13
Drug/Narcotic	208	18
DUI	45	4
Fraud	63	5.6
Homicide	45	4
Kidnapping	26	2.3
Larceny/Theft	106	9.4
Robbery	76	6.7
Sex Offence	132	12
Weapon Law Violation	49	4.3
Others - Class A	17	1.5
Others - Class B	53	4.7
Total	1128	100

3.4 Identification

Flanagan (2017) used the same data collected by JSP and used the gender-race composition of 'ever seated' jurors as a measure of the pool composition (e.g. $\frac{\text{no of 'ever seated' black jurors}}{\text{no of 'ever seated' jurors}}$). I argue that this variable is difficult to interpret and may potentially be endogenous; and propose using the composition of the first draw (composition of the first 12 jurors seated for the voir dire) as a more exogenous measure of jury composition.

3.4.1 The problem with using 'ever seated'

'Ever seated' composition is difficult to interpret because there are two kinds of 'ever seated' jurors: those who were not struck; and those who were struck. Therefore the composition of black jurors may increase due to: (i) more randomly drawn black jurors; or (ii) more struck black jurors.

Furthermore, this definition of composition is influenced by the number of strikes used, so it may also be endogenous. For instance, prosecutors may be more aggressive in striking black jurors if the evidence is not clear cut guilty (lower probability of conviction). This aggressiveness would result in a higher proportion of struck blacks in the composition.

3.4.2 First draw as a better alternative

There are several advantages of using the composition of the first draw. First, it is exogenous to any decision to strike because the first draw is realised prior to any questioning by the attorneys. Second, the denominator is constant at 12 which makes it comparable across trials. It is also possible to study the composition of juror draws at later rounds, but such cases would be subject to selection - more contentious cases may drag into later rounds. Also, the number of replacement jurors in later rounds may vary: having 50% blacks among 10 replacement jurors would have a far larger impact than 50% among 2 replacement jurors.

The composition of the first draw is conditional on the composition of the pool, which is missing from the dataset. Being unable to control for pool composition, my second best solution is to use county, year, season, and judge dummies to account for variation in the pool. County dummies are the most crucial variables because there is substantial cross-county variation in proportion of blacks in the population. Year and season dummies account for trends and seasonal differences in responding to jury summons due to job and family commitments. Judge dummies are included to account for any inadvertent differences in pool composition across judges.[9]

These controls may not fully capture differences in the pool composition. If so, then I expect there to be a negative bias between guilty verdicts and the composition of blacks in the first draw. This is because I expect a positive relationship between the proportion of blacks in the first draw and the pool (after all, the first 12 jurors are drawn from the pool); and a negative relationship between the proportion of blacks in the pool and guilty verdicts as shown in Anwar et al. (2012).

While the means of the variables are similar (15.9% using 'ever seated' jurors and 16% using first draw), the variances are different. Figure 3.1 plots the smoothed density of proportion of black jurors computed for 'ever seated' and first draw. The density at the tails are lower for 'ever seated', this could be because the attorneys' use of strikes are especially sensitive to outlier draws in the first draw. For instance, the defence would be more aggressive in striking non-blacks if there was zero blacks in the first draw. This squeezes the distribution of 'ever seated' blacks towards the

[9] For instance, judges may tend to schedule selections on different days of the week.

mean.

Figure 3.1: Proportion of blacks. Standard normal kernel with optimal bandwidth.

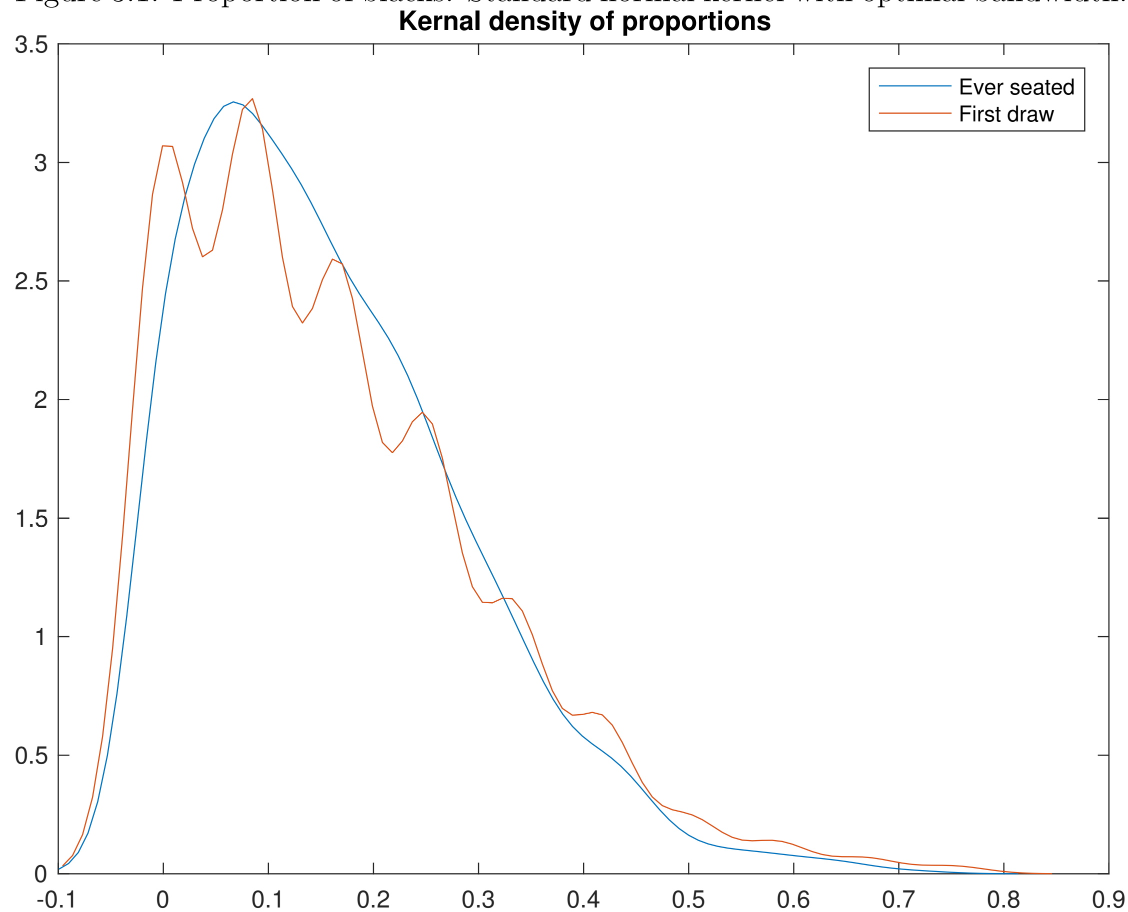


Table 3.3 documents the F-test to check whether charge and defendant characteristics predict the jury composition. For the black composition the characteristics together did not predict the proportion of blacks in 'ever seated' or first draw (P-values of 0.2 and 0.72). This suggests that using 'ever seated' might not be too problematic. However, I show that regression results using 'ever seated' were vastly different from using first draw, suggesting that unobservable variables may be driving the endogeneity.

Table 3.3: F-test of charge and defendant characteristics

	(1) Black 1st draw	(2) Black 'ever seated'	(3) Female 1st draw	(4) Female 'ever seated'
Burglary/B&E	-0.002	-0.007	0.014	0.018
	(0.011)	(0.008)	(0.023)	(0.018)
Drug/Narcotic	-0.011	-0.010	0.007	0.010
	(0.014)	(0.011)	(0.017)	(0.013)
DUI	-0.009	-0.001	0.004	-0.011
	(0.019)	(0.017)	(0.028)	(0.022)
Fraud	-0.018	-0.013	0.012	0.022
	(0.018)	(0.014)	(0.017)	(0.017)
Homicide	0.013	-0.006	0.021	0.025
	(0.013)	(0.007)	(0.025)	(0.015)
Kidnapping	-0.015	-0.001	-0.038	-0.004
	(0.020)	(0.018)	(0.032)	(0.022)
Larceny/Theft	0.004	0.005	0.018	0.015
	(0.015)	(0.012)	(0.022)	(0.015)
Robbery	-0.017	-0.017	-0.001	0.016
	(0.013)	(0.011)	(0.028)	(0.021)
Sex Offence	-0.028	-0.023	0.005	0.010
	(0.014)	(0.010)	(0.021)	(0.014)
Weapon Law Violation	0.007	0.003	0.021	0.029
	(0.020)	(0.014)	(0.028)	(0.021)
Others - Class A	0.013	0.005	-0.024	-0.007
	(0.026)	(0.021)	(0.039)	(0.034)
Others - Class B	-0.009	-0.012	0.028	0.019
	(0.023)	(0.015)	(0.033)	(0.022)
>1 charge	-0.004	-0.016	0.013	-0.010
	(0.009)	(0.009)	(0.017)	(0.009)
Female defendant	0.005	0.003	-0.000	-0.010
	(0.009)	(0.008)	(0.020)	(0.013)
Black defendant	0.005	0.005	0.005	-0.005
	(0.012)	(0.008)	(0.009)	(0.007)
Old defendant	0.001	-0.002	0.008	-0.001
	(0.008)	(0.007)	(0.009)	(0.008)
Observations	1128	1128	1128	1128
F-stat	0.76	1.32	0.60	1.93
P-value	0.72	0.20	0.88	0.03

3.4.3 Econometric models

At the individual juror level, I run the following regression to estimate differences in the probability of being struck, by gender and race:

$$\text{strike}_{jt} = \alpha_t + \beta_1 * \text{White-female}_{jt} + \beta_2 * \text{Black-male}_{jt} + \beta_3 * \text{Black-female}_{jt} + e_{jt} \quad (3.1)$$

Where $strike_{jt} = 1$ if juror j in trial t was struck, and 0 otherwise. I run separate regression for prosecutor and defence strikes. Since I observe several jurors within a trial, I am able to control for trial fixed-effects. I also run the regression using 'ever seated' jurors, and jurors in the first draw. The results should not differ much because, controlling for trial fixed-effects, the decision to strike jurors should not differ too much across different rounds.

At the trial level, I test whether the composition of blacks jurors affects verdicts by running the regression:

$$guilty_t = \beta * \text{prop. black}_t + \gamma' X_t + e_t \tag{3.2}$$

Where $guilty_t = 1$ if trial t resulted in a guilty verdict, and 0 otherwise. I compare the two definitions of proportion black detailed above ('ever seated' and first draw).

3.5 Results

3.5.1 Striking behaviour

Table 3.4 details the results from estimating equation 3.1. The excluded gender-race category is white males so we can interpret the coefficient of the constant as the probability of striking a white male.

White males had a 10.9% probability of being struck by a prosecutor, while black males and females were 12.6pp and 9.2pp more likely to be struck. This implies that prosecutors struck black males (23.3%) at twice the rate relative to white males. In contrast, white females were 2.8pp less likely to be struck.

Defence attorneys were more likely to strike white males compared to prosecutors. They were also 16.4pp less likely to strike black males and 15pp less likely to strike black females compared to white males. This implies that defence attorneys struck white males at more than four times the rate relative to white males (4.9%). There was not statistically significant difference between the defence striking white males and white females.

Since the defence substantially was more aggressive in striking whites relative to blacks, white jurors had a higher probability of being struck. For instance, the

Table 3.4: Juror-level regressions

	(1) prosecutor strike	(2) defence strike	(3) prosecutor strike	(4) defence strike
White Female	-0.028	0.008	-0.033	0.001
	(0.007)	(0.009)	(0.009)	(0.012)
Black Male	0.126	-0.164	0.116	-0.173
	(0.015)	(0.011)	(0.019)	(0.015)
Black Female	0.092	-0.150	0.096	-0.146
	(0.011)	(0.011)	(0.015)	(0.015)
Constant	0.109	0.213	0.119	0.224
	(0.004)	(0.005)	(0.005)	(0.007)
Observations	15608	15608	8930	8930
Sample	'Ever seated'	'Ever seated'	First draw	First draw

coefficients imply that white males had a 32.2% chance of being stuck while black makes had a 28.4% chance of being struck.[10] Note that this does not mean that that blacks will be over represented in the final jury, because the probability of drawing a black replacement juror is < 0.5.

These results suggests race is a more important observable determinant of striking behaviour than gender. They also show that prosecutors and defence attorneys have opposite preferences on black jurors. The coefficients on black males and females remain broadly similar when I restricted the sample to jurors in the first draw.

I further separated the sample by trials with black and non-black defendants (table 3.5). Broadly, in both samples, prosecutors (defence attorneys) struck black (white) jurors at a higher rate than white (black) jurors. However, both attorneys were more aggressive when the defendant was black. For instance, relative to white males, prosecutors were 11.3pp more likely to strike black females when the defendant was black, compared with 4pp more likely when the defendant was not black.

[10]The probabilities were 30.3% for white females and 26.4% for black females.

Table 3.5: Juror-level regressions by race of defendant

	(1) prosecutor strike	(2) defence strike	(3) prosecutor strike	(4) defence strike
White Female	-0.018	-0.002	-0.044	0.022
	(0.008)	(0.012)	(0.011)	(0.013)
Black Male	0.151	-0.201	0.060	-0.068
	(0.018)	(0.013)	(0.028)	(0.023)
Black Female	0.113	-0.193	0.040	-0.040
	(0.013)	(0.012)	(0.020)	(0.023)
Constant	0.099	0.241	0.125	0.169
	(0.005)	(0.007)	(0.007)	(0.007)
Observations	10020	10020	5588	5588
Sample	Black def	Black def	Non-black def	Non-black def

For 'ever seated' jurors.

3.5.2 How jury composition affects verdicts

Table 3.6 summarises the estimates of equation 3.2; adding controls did not seem to change the estimated coefficients much so I will focus on interpreting the estimates from models (5) and (6), with all the controls. There was a negative, and statistically significant (at 5% level) relationship between the proportion of 'ever seated' black jurors and guilty verdicts. The magnitude of the coefficient implies that an increase from the 25th percentile of proportion blacks, to the 75th percentile would lower the probability of conviction by 6.2pp.

The estimated coefficient for proportion of blacks in the first draw, was about a third the size of 'ever seated' blacks. Despite of the smaller standard errors, the coefficient was not statistically significant even at the 10% level. The magnitude of the coefficient implies that an increase from the 25th percentile of proportion blacks, to the 75th percentile would lower the probability of conviction by 1.9pp. Note that this smaller coefficient includes the possible negative bias induced by not directly controlling for pool composition. Hence we might expect the true coefficient to be even smaller.

What would account for the smaller estimates when we use the composition of the first draw? One possibility is the endogeneity of using 'ever seated' compositions. Suppose prosecutors were less likely to strike black jurors if the case was clear cut guilty (higher probability of guilty verdict); and this resulted in a lower proportion of 'ever seated' blacks, because there would be less struck blacks in the numerator. Then we would expect a negative bias in the coefficient.[11] The F-test shown earlier indicated that observed case characteristics were unable to predict the composition of 'ever seated' blacks, hence any endogenity should work through unobservable - e.g. strength of evidence, eyewitness testimonies.

Table 3.6: Trial-level regressions

Dependent: guilty verdict	(1)	(2)	(3)	(4)	(5)	(6)
Prop black - 'ever seated'	-0.369		-0.335		-0.337	
	(0.166)		(0.158)		(0.1563)	
Prop black - first draw		-0.0979		-0.116		-0.116
		(0.126)		(0.123)		(0.123)
Defendant characteristics					Yes	Yes
Charge details			Yes	Yes	Yes	Yes
Judge	Yes	Yes	Yes	Yes	Yes	Yes
Year/season	Yes	Yes	Yes	Yes	Yes	Yes
County	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1150	1150	1128	1128	1128	1128

Another possibility is the link between the variables and the racial composition of the final jury. Since prosecutors were more likely to strike black jurors, the probability that black jurors from the first draw making it into the final jury would be low. Indeed, when I regressed the proportion of blacks in the final jury on the two variables, the coefficient for the first draw is weaker at 0.64 compared to 1.02 for 'ever seated' (table 3.7).

Since the composition of first draw predicts that of the final jury, it is possible to use the former as an instrument for the latter. I have two concerns about doing so. First, the first draw may affect the jurors' impression of the attorneys. For instance, the final jurors' impression of the prosecutor might be tainted by watching her strike many black jurors from the first draw. Second, we know that the first draw is related to the pool composition, and the pool composition affects conviction rates (Anwar et al. (2012)). The inability to measure pool composition directly would cause the exclusion restriction to fail. As a reference, I include the

[11] The bias should be exacerbated by the behaviour of the defence attorney. In particular, I expect that defence attorneys were more likely to strike white jurors if the case was clear cut guilty; thus resulting in a lower proportion of 'ever seated' blacks, because there would be more struck whites in the numerator.

Table 3.7: Relationship with composition of final jury

Dependent: Prop black - final jury	(1)	(2)
Prop black - 'ever seated'	1.0275	
	(0.0311)	
Prop black - first draw		0.6370
		(0.0341)
Defendant characteristics	Yes	Yes
Charge details	Yes	Yes
Judge	Yes	Yes
Year/season	Yes	Yes
County	Yes	Yes
Observations	1128	1128

OLS and 2SLS comparisons in Appendix G.

When I split up the sample by black and non-black defendants, the estimated standard errors were too large to make precise inferences (table 3.8). But it suggests that the negative relationship between black jurors and guilty verdicts was driven by the sub-sample with black defendants.

Table 3.8: Trial-level regressions by race of defendant

Dependent: guilty verdict	(1)	(2)	(3)	(4)
Prop black - 'ever seated'	-0.385		0.021	
	(0.278)		(0.633)	
Prop black - first draw		-0.073		0.115
		(0.199)		(0.495)
Defendant characteristics	Yes	Yes	Yes	Yes
Charge details	Yes	Yes	Yes	Yes
Judge	Yes	Yes	Yes	Yes
Year/season	Yes	Yes	Yes	Yes
County	Yes	Yes	Yes	Yes
Observations	648	648	480	480
Sample	Black def	Black def	Non-black def	Non-black def

I also split up the sample into counties with a low and high proportion of black residents (table 3.9).[12] The results indicate a positive relationship between the proportion of blacks in the first draw and guilty verdicts in low proportion

[12]This was based on the 2010 census. I classified low/high based on whether the county's proportion of black residents was below/above 20%.

counties; but the opposite in high proportion counties. Both coefficients were significant at the 10% level.

These results could be driven by the differences in the proportion of black defendants - about 60% of defendants in low proportion counties were non-black defendants. They could also be driven by selection. The type of blacks who choose to live in communities with many non-blacks (e.g. rural communities) may have a different outlook on crime than other blacks (e.g., living in urban communities).

Table 3.9: Trial-level regression by proportion of blacks in county

Dependent: guilty verdict	(1)	(2)
Prop black - first draw	0.5910	-0.3121
	(0.3454)	(0.1722)
Defendant characteristics	Yes	Yes
Charge details	Yes	Yes
Judge	Yes	Yes
Year/season	Yes	Yes
County	Yes	Yes
Observations	465	663
Sample	Low prop of blacks	High prop of blacks

3.6 Model

I build a game theory model using key results from section 3.5 as a guide: the race of jurors and defendants affect verdicts. While the reduced form regressions suggest gender also plays a role in verdicts and the striking behaviour of the attorneys, I focus only on the impact on race in the model to keep the it parsimonious for computational tractability. The aim of this model is to compare the current allocation of six strikes (for each side), to the counterfactual of reducing the number of allocated strikes symmetrically.

3.6.1 Model set-up

There two are players in the game: p (prosecutor) and d (defence); both are maximising their utility from jury verdicts $\pi_j(Y, X_j, U_j)$, where $j \in \{p, d\}$. It is static one-round game where the players observe the characteristics of 12 jurors being drawn (i.e. first draw) and each has 6 strikes to veto jurors ($Y = (Y_p, Y_d)$ and $Y_j \in \{0, 1, 2, 3, 4, 5, 6\}$). Some characteristics are known to the econometrician (X, like race of jurors, race of the defendant, number of excuse for cause by the judge) while others are not (U_j, like the propensity of the jurors to respond well to evidence presented).

This static setup abstracts away from the dynamic nature of the actual jury selection process. The main reason is for tractability. The dynamic games literature focuses on games with fixed action sets in every round (e.g stay or exit the market); in a dynamic jury selection game, actions affect subsequent actions **sets**, so action sets become state variables.[13] Empirically, I still focus on the most relevant round: an average of 65.3% of prosecutors 63.5% of defence strikes, over the entire selection process, are used in the first round.

One way to interpret the utility in this context is of the **continuing utility** of moving to the next round. This would include (i) the utility from the final jury (not vetoed) and (ii) the expected utility from subsequent rounds. The average number of strikes in the first round is 3.9, so there is an average of 8.1 jurors who were not struck (i.e., part of the final jury) after the first round; so I expect (i) to be a greater contributor than (ii).

Also this is a game of complete information, which means that both players observe U_p and U_d. This is not unreasonable given both sides are present in the selection process (e.g. can listen to jurors' answers to questions from either side) and do not have prior information about the jurors. I do not fix a relationship between U_p and U_d, but estimate a parameter corresponds to their correlation. Given that the jury verdict is zero sum, I expect the two variables to be strongly negatively correlated.

[13] In earlier versions of this paper, I tried to set up such a dynamic game. One problem I encountered was that some subgames would have Pure Strategy Nash Equilibria (PSNE), while others did not.

3.6.2 Parametric model

I use the parametric specification used in Aradillas-Lopez and Rosen (2019) (henceforth AR). Their key contribution is showing how to identify and estimate a static game with ordered action sets. Jury selection is a suitable application of their approach because of the ordinal nature of the action sets (number of strikes).

$$\pi_j(Y, X_j, U_j) = Y_j * (\alpha + X_j\beta_j - \Delta Y_{-j} - \eta Y_j + U_j) \tag{3.3}$$

- Y_{-j} is the number of strikes of the other player; so Δ_j indicates whether strikes are strategic complements ($\Delta_j < 0$, where a striking is more valuable when others strike) or substitutes ($\Delta_j > 0$, where others striking reduces your incentive to strike). This is an interesting parameter because it tells us the nature of the game, if the players are striking "neutrally" akin to an excuse for cause (vetoing jurors unsuitable to serve) then I expect $\Delta_j > 0$ because a strike reduces the incentive for the other player to strike. However, if the players are striking to their own advantage, then I expect $\Delta_j > 0$, as a strike incentives the other player to use her strike to cancel out the advantage.

- η corresponds to the concavity of the utility function. I will constrain $\eta > 0$ to ensure the uniqueness of the best response (see below).

- X_j includes three variables: (i) the number of black jurors in the first draw; (ii) whether the defendant is black; and (iii) the number of strikes for cause by the judge. β_j is a reflection these characteristics may affect players differently (e.g. black jurors are disadvantageous for prosecutors but a boon for the defence).

In addition I restrict the distribution of the bivariate unobserved heterogeneity U to a Gaussian copula indexed by the parameter $\lambda \in (-1, 1)$. Specifically, U_p and U_d each have the logistic marginal Cumulative Distribution Function (CDF)

$$G(u_j) = \frac{\exp(u_j)}{1 + \exp(u_j)}$$

, and their joint CDF is

$$F(u_p, u_d; \lambda) = \Phi_G(\Phi^{-1}(u_p), \Phi^{-1}(u_d); \lambda) = \int_{-\infty}^{\Phi^{-1}(u_p)} \int_{-\infty}^{\Phi^{-1}(u_d)} \frac{1}{2\pi(1-\lambda^2)^{1/2}} * \frac{-(s^2 - 2\lambda st + t^2)}{2(1-\lambda^2)} dsdt$$

where Φ is the CDF of the standard normal distribution, and Φ_G is the standard bivariate normal distribution with correlation parameter λ restricted to the interval $(-1, 1)$. The benefit of using a Gaussian copula is that it permits a high positive or negative correlation between the U_p and U_d. [14] This is important because I expect the correlation to be strongly negative (i.e. close to -1) given the zero sum nature of a jury verdict.

There are two key parametric shape assumptions required for the existence and uniqueness of a Pure Strategy Nash Equilibrium (PSNE). The first is concavity:

$$\pi_j(y_j, y_{-j}, X_j, u_j) - \pi_j(y_j - 1, y_{-j}, X_j, u_j) > \pi_j(y_j + 1, y_{-j}, X_j, u_j) - \pi_j(y_j, y_{-j}, X_j, u_j) \tag{3.4}$$

The other is increasing difference, which requires that if $y_j^{'} > y_j$ and $u_j^{'} > u_j$

$$\pi_j(y_j^{'}, y_{-j}, X_j, u_j^{'}) - \pi_j(y_j, y_{-j}, X_j, u_j^{'}) > \pi_j(y_j^{'}, y_{-j}, X_j, u_j) - \pi_j(y_j, y_{-j}, X_j, u_j) \tag{3.5}$$

To see why we can work out a d step difference in utility, given y_{-j} and X_j, between playing action y_j versus $y_j - d$:

$$\Delta^d \pi_j(y, X_j, U_j) = \pi_j(y_j, y_{-j}, X_j, U_j) - \pi_j(y_j - d, y_{-j}, X_j, U_j) = \alpha + X_j \beta_j - \Delta y_{-j} + U_j - \eta(2dy_j - d^2) \tag{3.6}$$

$\Delta^1 \pi_j(y, X_j, U_j)$ is decreasing in y_j when $\eta > 0$, which satisfies concavity. Also, notice $\Delta^d \pi_j(y, X_j, U_j)$ is increasing in u_j, which satisfies increasing differences.

Translated to the context of jury selection. Concavity means that the marginal gain from additional strikes is decreasing. This is reasonable because I expect both sides to eliminate jurors most detrimental to their case first. Any subsequent strikes would therefore be targetted at less undesirable jurors. Increasing difference means that players are more likely to strike when the first draw is randomly more beneficial to their case (unobserved by the econometrician). This implies a complementarity between favourable draws and strikes. For instance, a player is more likely to

[14] For instance, XX use a Farlie-Gumbel-Morgenstern (FGM) copula which restricts the correlation to be $[-\frac{1}{3}, \frac{1}{3}]$. If I use the FGM copula, λ would be estimated at exactly $-\frac{1}{3}$ (with tight standard errors), which suggests copula is too limiting.

eliminate a "bad" juror when the others are "good" compared to when the others are "bad". Since the average number of jurors who were not struck after the first round (i.e., in the final jury) is 8.1 and 90% of cases have 7 jurors who were not struck, and hence form part of the final jury, after the first round, there is indeed a sense that the majority of jurors are "set in stone" after the first round and that if these final jurors are already not favourable then it is quite inefficient to trim out the "bad" jurors in the margin.

However, this assumption is not innocuous. It rules out the possibility that players try to aggressively trim the jury when the first draw is largely unfavourable.

For every y_j we can define a threshold u_j^- such that $\Delta^1\pi_j(y_j, y_{-j}, X_j, u_j^-) = 0$ and u_j^+ such that $\Delta^1\pi_j(y_j + 1, y_{-j}, X_j, u_j^+) = 0$. By increasing difference, we know that for any $u_j > u_j^- \iff \pi_j(y_j, y_{-j}, X_j, u_j) > \pi_j(y_j - d, y_{-j}, X_j, u_j)$ and $u_j < u_j^- \iff \pi_j(y_j, y_{-j}, X_j, u_j) < \pi_j(y_j - d, y_{-j}, X_j, u_j)$. Also, by concavity and increasing differences, we know that for $u_j < u_j^+ \iff 0 > \Delta^1\pi_j(y_j + 1, y_{-j}, X_j, u_j) > \Delta^1\pi_j(y_j + 2, y_{-j}, X_j, u_j) > ....$ Therefore we know that y_j is the best response (given y_{-j} and X_j) $\iff u_j \in [u_j^-, u_j^+]$. This leads to the condition for y_j.

$$\alpha + X_j\beta_j - \Delta Y_{-j} - \eta(2y_j - 1) \leq u_j \leq \alpha + X_j\beta_j - \Delta Y_{-j} - \eta(2y_j + 1) \tag{3.7}$$

AR show that these thresholds allow us to set identify the parameters in the model.

3.6.3 Identification and estimation

I follow AR in estimating the parameters via a two-step process. The first step uses Maximum Likelihood Estimation (MLE) to point identify a subset of the variables, specifically $(\beta_p^{'}, \beta_d^{'}, \eta - \delta, \lambda)$. The second step set identifies all the parameters using moment inequalities and allows us to make inference on the bounds of $(\Delta_p, \Delta_d, \eta, \delta)$.

Maximum Likelihood Estimation

First, notice that when $y_{-j} = 0$ the best response is only dependent on X_j. We know that

$$Pr(y_p = 0, y_d = 0) \iff u_p(-\infty, \eta - \delta - X_p\beta_p] \quad \& \quad u_p(-\infty, \eta - \delta - X_d\beta_d]$$

and therefore

$$Pr(y_p = 0, y_d = 0) = G(\eta - \delta - X_p\beta_p, \eta - \delta - X_d\beta_d)$$

. To estimate these parameters (θ_{MLE}) using MLE I set up the log-likelihood as

$$L(\theta_{MLE}) = \sum_{i=1}^{n} \mathbf{1}\{(y_p, y_d) = (0,0)\} \ln G(\eta - \delta - X_p\beta_p, \eta - \delta - X_d\beta_d) \\ + \mathbf{1}\{(y_p, y_d) \neq (0,0)\} \ln[1 - G(\eta - \delta - X_p\beta_p, \eta - \delta - X_d\beta_d)] \quad (3.8)$$

Partial identification and estimation

First, let us define a box $R(Y, X; \theta)$ as the box $[u_p^-, u_p^+] \cup [u_d^-, u_d+]$. This box contains the values of (u_p, u_d) that are consistent with the equilibrium actions $Y = (Y_p, Y_d)$, under the observed X and set of parameters θ. Consider an $\epsilon \in R(Y, X; \theta)$ a test set C which can be thought of an arbitrary box, it must be that

$$\mathbf{1}\{R(Y, X) \in C\} \leq \mathbf{1}\{\epsilon \in C\} \quad \forall C \subset \mathbf{R}^2$$

. This is simply saying it cannot be that the whole box supporting an equilibrium is in C but an element of it is not in C. However, the opposite is possible: that an element ϵ is in C but the entire box is not contained in C. Therefore identified parameters are those that fulfil the moment inequalities (one for each C):

$$Pr(R(Y, X; \theta) \subseteq C|X) \leq Pr(u \in C|X) \quad \forall C \subset \mathbf{R}^2 \quad (3.9)$$

AR detail an estimator of these parameters and I follow their steps. Briefly, the steps are

1. Choose a number of test sets C.
2. Choose a set of parameter values $\tilde{\theta}$.

3. Construct an estimator of $Pr(R(Y,X;\tilde{\theta}) \subseteq C) - Pr(u \in C)$.

4. Construct an estimator of the weighted number of violations of the moment inequalities 3.9 (i.e. $Pr(R(Y,X;\tilde{\theta}) \subseteq C) - Pr(u \in C) > 0$)

5. Construct a Wald test to check if the violations are statistically different from zero

6. If the violations are no different from zero then $\tilde{\theta}$ is accepted as possible, if the violations are different from zero then we reject $\tilde{\theta}$.

7. Repeat steps 2-6 with other sets of parameters

3.6.4 Results

Inference of MLE parameters

Table 3.10 summarises the parameters estimated using MLE. First, a higher number of black jurors decreases the utility of p and increases the utility of d; both estimates are significantly different from zero. This result is not surprising, and is consistent with the reduced form analysis. In particular, p suffers a bigger decline in utility than d gains from one more black juror. This could be due to the unanimity required for conviction: one deviant juror is sufficient to thwart a guilty verdict.

Second, p suffers from a black defendant while d benefits from it. In the data, there is no significant difference between the conviction rates of black vs non-black defendants (after controlling for the charge and the jury composition); so it is not clear that non-blacks are easier to convict. However, it may be that the race of the defendant is highly correlated with the type of charges, which are related to ease of conviction.

Third, both players seem to benefit from the judge excusing jurors for cause (though this is only significant for d). This could be due to the norm setting role of a judge strike: a judge that is more liberal with excusing is more likely to more liberal towards accepting players' strikes without question [15]. Indeed, the number of strikes used by both parties are positively and significantly correlated with the number of judge strikes.

[15] For instance, a strict judge may use the Batson rule to force players to rationalise their strikes.

Finally, λ is significantly negative and implies that u_p and u_d have a negative correlation coefficient of -0.59. This is consistent with my expectation that, in a zero sum game, what is good for one party should be bad for another.

Table 3.10: Estimated parameters from MLE

Parameter	Estimate	Standard Errors	95% Confidence Interval	
$\beta_{p,\text{number of black jurors}}$	-0.677	0.007	-0.691	-0.662
$\beta_{p,\text{black defendant}}$	-1.084	0.092	-1.264	-0.903
$\beta_{p,\text{number of judge strikes}}$	0.030	0.124	-0.214	0.273
$\beta_{d,\text{number of black jurors}}$	0.078	0.003	0.071	0.085
$\beta_{d,\text{black defendant}}$	0.447	0.028	0.392	0.501
$\beta_{d,\text{number of judge strikes}}$	1.598	0.027	1.545	1.652
$\eta - \delta$	0.205	0.012	0.182	0.227
λ	-0.606	0.029	-0.663	-0.548

Full inference

Table 3.11 summarises the results from the moments inequality estimation, with the MLE estimates and Confidence Intervals (CI) for comparison. For parameters estimated using MLE, their CI are largely similar to the CI estimated using the moment inequalities.

The CI of Δ_p and Δ_d allow us to reject the null of zero, so we can infer they are significantly negative. This means, for both sides, strikes are strategic complements: striking begets more striking. This could be interpreted as a tit-for-tat strategy, where one side strikes in respond to losing a favourable juror by striking off an unfavourable one in retaliation.

Also, the CI of η allow us to reject the null of zero, so we can infer it is significantly positive. This results is consistent with the shape requirement of concavity.

3.6.5 Counterfactual analysis

The model allows us to calculate the ex ante (before X and U are known) expected utility for p and d and how they change with the number of strikes allowed. The idea behind a symmetrical number of strikes is that it seems to be equally

Table 3.11: Estimated parameters from moment inequality

Parameter	MLE Estimate	MLE 95% CI		95% CI	
$\beta_{p,\text{number of black jurors}}$	-0.677	-0.691	-0.662	-1.263	-0.014
$\beta_{p,\text{black defendant}}$	-1.084	-1.264	-0.903	-1.261	-0.907
$\beta_{p,\text{number of judge strikes}}$	0.030	-0.214	0.273	-0.126	0.185
$\beta_{d,\text{number of black jurors}}$	0.078	0.071	0.085	0.000	0.984
$\beta_{d,\text{black defendant}}$	0.447	0.392	0.501	0.114	0.770
$\beta_{d,\text{number of judge strikes}}$	1.598	1.545	1.652	1.262	1.925
$\eta - \delta$	0.205	0.182	0.227	-0.289	0.701
λ	-0.606	-0.663	-0.548	-0.293	-0.926
Δ_p				-8.899	-4.094
Δ_d				-8.238	-3.113
η				3.123	5.635
δ				2.422	5.924

benefit both sides. This may not be true in practice as minority jurors (blacks) have an asymmetric impact. Figure 2 to figure 5 presents the range of expected utilities as I symmetrically increase the number of strikes allowed. I cannot provide a point estimate because some of the parameters are set identified. In generating the counterfactuals, I use point identified parameters when possible and use combinations of extreme values in the CI of set identified variables to generate a range of possibilities. In particular, I fix the concavity (η) and utility shifter (δ) at an arbitrary level and examine the counterfactuals at different levels of Δ_p and Δ_d. This is because p and d share the same η and δ so while they affect the level of utility for both players, they should not have a big asymmetric impact across the players. The results are summarised in the figures below, where the expect ex ante utility of both players are plotted against the number of strikes. Weak and strong refer to the magnitude of the strategic complementarity. For instance, a weak complementarity for p would be -4.094 - the lowest magnitude in the CI for Δ_p.

Unfortunately, the range across Δ_p and Δ_d can lead to asymmetries in favour of either player. If Δ_j is a large negative relative to Δ_{-j}, then we see that player j has an enormous advantage from having more strikes. This is because it allows her more strikes to exploit the complementarities. I interpret this as equivocating the assumption in the literature that allowing more strikes is advantageous to the prosecutor.

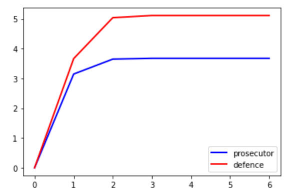


Figure 3.2: p weak d weak

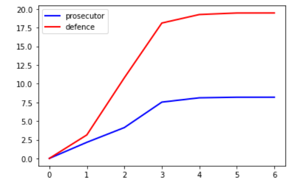


Figure 3.3: p weak d strong

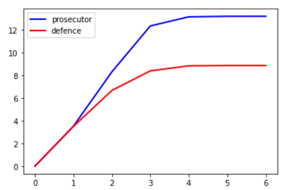


Figure 3.4: p strong d weak

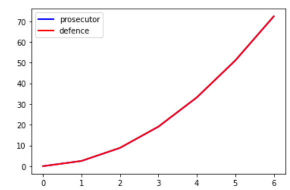


Figure 3.5: p strong d strong

3.7 Conclusion

My understanding is that there is currently no other work done on looking at how the jury composition of the first draw affects jury verdicts. It would be interesting to use the Florida data to compare whether the composition in the pool or in the first draw had a greater impact on verdicts. The pool composition could have a greater impact because it affects the expected type of replacement juror, and hence affects all strikes attorneys use. Whereas the first draw may only affect some of the strikes attorneys use. Also, given the relevance of gender in the jury selection process, the model can be extended to account for the differences in striking behaviour based on gender **and** race.

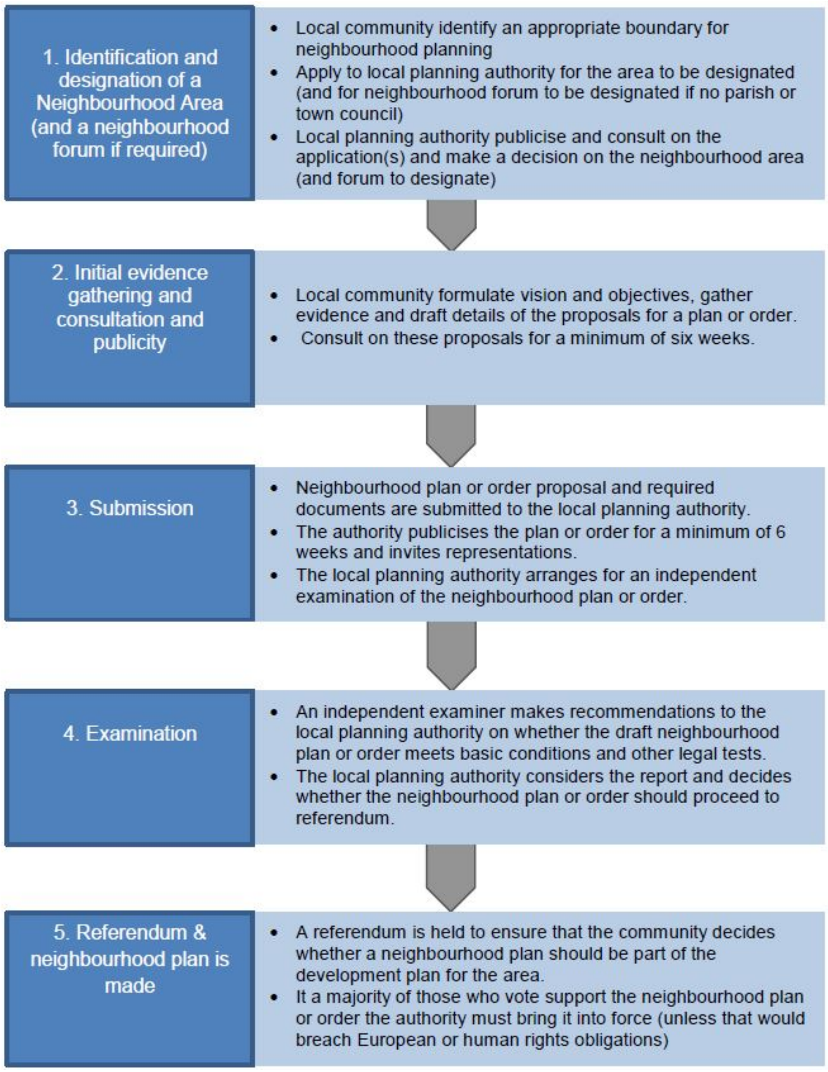
The results from my model suggest: (i) that the defence benefits from more black jurors; (ii) that the game is one of strategic complementarity and (iii) under strong strategic complementarity, the defence can benefit from being allowed more strikes because she can use them to restore any disadvantage from the prosecutor's strikes.

Appendix A

Localism Act (2011) and the Rise of NIMBY Behaviour

Neighborhood plans were introduced by the Localism Act (2011) to give locals more power in deciding how land is used in their neighborhood. There is considerable flexibility in what neighborhood plans can include. For example, they can range from just a few policies on design or retail uses to a whole suite of comprehensive plans which incorporate a diverse range of policies and site allocations for housing or other development. Figure A1 shows the stages involved in putting together a neighborhood plan.[1]

Figure A1: Neighborhood plan application process in England



Source: Plain English guide to the Planning System (2015)

[1]Based on https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/391694/Plain_English_guide_to_the_planning_system.pdf

As a first step, because there is no fixed definition of what constitutes a neighborhood, the local community and the local planning authority must identify and agree on an appropriate boundary for the neighborhood plans. In the second and third step, the local community then formulates a vision and objectives, gather evidence and draft the details of their neighborhood plan. The plan is then submitted to the local planning authority. Next, the local planning authority then arranges for an independent examiner to check if the neighborhood plan fits with local strategic and national policies and whether it complies with the law. If the neighborhood plan passes the independent examiner's assessment, it is put to a local referendum where the community votes to decide whether the neighborhood plan should be passed. If it passes the referendum, the neighborhood plan then becomes part of the statutory development plan which is used by the local planning authority when deciding planning applications.

Unfortunately, the advent of neighborhood plans may have made it easier for locals to engage in NIMBY behavior by putting in place regulations to protect their own interests at the expense of "outsiders". First, the definition of a neighborhood is highly nebulous and can be easily abused. Second, the wording in the neighborhood plans is extremely subjective and this could make it difficult for the independent examiner to evaluate it. Figure A2 shows an example of the subjective manner in which neighborhood plans are often phrased.[2]

Figure A2: Example of how neighborhood plans are phrased

A Neighbourhood Plan has many benefits. The Woodcote Neighbourhood Plan has been developed by volunteers from the village to:

- protect the village from uncontrolled, large scale, or poorly placed development;
- spread the development required by South Oxfordshire's Core Strategy around the village across several small sites;
- ensure that development is sympathetic to, and improves, the look and feel of the village;
- take steps to give residents preferred access to many of the new homes;
- minimise the loss of greenfield sites by, where possible, using previously developed sites;
- give the village the potential to access Community Infrastructure Levy funding to improve village facilities; and
- identify, in a **Woodcote Parish Action Plan 2013**, additional actions to improve Woodcote's facilities, services and local environment and to address issues beyond the scope of the Neighbourhood Plan.

Third, some of the objectives in neighborhood plans could be at direct odds

[2] This is from http://www.southoxon.gov.uk/sites/default/files/2014-05-23%20Woodcote%20Neighbourhood%20Plan_0.pdf. The use of this text is meant to be an example of how neighborhood plans are phrased. We are by no means making an assessment as to whether Woodcote's neighborhood plan is a good or bad one.

with national policies. For example, the National Planning Policy Framework has a five year housing land supply requirement where local planning authorities have to ensure that there is a regular supply of land that is suitable, available and deliverable for housing development. However, this is often in conflict with one of the common objectives of neighborhood plans which is to "protect the feel" of an area because protecting the feel of an area often involves limiting the height and size of housing developments.

It is no surprise that since neighborhood planning was introduced by the Localism Act (2011), there has been an increasing number of communities across England that have held a referendum to vote on their neighborhood plans. Table A1 shows that the number of referendum held has increased dramatically over time. The average turnout is low but those who turn up tend to vote overwhelmingly in support of the neighborhood plans.

Table A1: Summary statistics for neighborhood plans

Year	Number of referendum	Average turnout	Average voted Yes
2013	6	33.33	86.83
2014	36	33.78	87.36
2015	84	32.14	89.58
2016	166	31.23	88.86
2017	210	33.96	88.13
2018	208	31.38	88.65

Appendix B

Mathematical Appendix for Chapter 1

B1 Proof of proposition 1

First, we can work out the $P_\kappa(\kappa, s)$ by differentiating the market clearing with respect to κ, which gives

$$-f_r\left(W_r + \psi(\kappa, s) - P(\kappa, s)\right)*(\psi_\kappa(\kappa, s) - P_\kappa(\kappa, s)) = f_p\left(W_p + \phi(\kappa, s) - P(\kappa, s)\right)*(\phi_\kappa(\kappa, s) - P_\kappa(\kappa, s))$$

$$\implies P_\kappa(\kappa, s) = \theta_r(W_r, W_p, \kappa, s) * \psi_\kappa(\kappa, s) + \theta_p(W_r, W_p, \kappa, s) * \phi_\kappa(\kappa, s)$$

Where $f_k(.) = \frac{\partial}{\partial \kappa} F_k(.)$ and $\theta_k(.) = \frac{f_k(.)}{f_k(.) + f_{!k}(.)}$ for $k = \{r, p\}$. Therefore, as the exogenous amenity value of the location increases, the clearing price rises as a weighted average of the increase in utility of resident and potential homeowners.

Next we, define $\underline{\kappa} < 0$ as value of κ that clears the market at $P(\underline{\kappa}, 0) = 0$ satisfies the market clearing condition (i.e., the lowest value of κ such that the market clears). Values of $\kappa < \underline{\kappa}$ would imply that $P(\kappa, s) < 0$, which we do not consider because residents can simply abandon their properties, and take their outside option, without paying potential residents. Given $s = 0$, the net weighted social utility of rejecting (i.e., the weighted social utility of rejecting subtracted by that of accepting) is

$$\omega\tilde{s}\,(0,0) * (W_r + \psi(0,0)) + \dot{s}\,(0,0) * (W_p + \psi(0,0) - P(0,0))$$
$$- \left[\omega\tilde{s}(\underline{\kappa}, 0) * (W_r + \psi(\underline{\kappa}, 0)) + \dot{s}(\underline{\kappa}, 0) * (W_p + \phi(\underline{\kappa}, 0))\right]$$

$$\leq \omega \tilde{s}(\underline{\kappa},0) * (\psi(0,0) - \psi(\underline{\kappa},0)) + \omega \Delta \tilde{s} * (W_r + \psi(0,0))$$

$$-\Delta \tilde{s} * \max \{W_p + \phi(0,0) - P(0,0), W_p + \phi(\underline{\kappa},0)\}$$

Where $\Delta \tilde{s} = \tilde{s}(0,0) - \tilde{s}(\underline{\kappa},0) > 0$ because $\frac{\partial \tilde{s}}{\partial \kappa} \geq 0$. Since $\psi(0,0) > \psi(\underline{\kappa},0)$:

$$\leq \omega \Delta \tilde{s} * (W_r + \psi(0,0))$$

$$-\Delta \tilde{s} * \max \{W_p + \phi(0,0) - P(0,0), W_p + \phi(\underline{\kappa},0)\}$$

$$< 0$$

Because of the assumptions that $\omega \geq \frac{W_p}{W_r}$ and $\psi(\kappa,s) \geq \phi(\kappa,s)$. Therefore, L will reject the development.

Finally we can work out how the weighted social utility changes with an increase in κ:

$$\frac{\partial}{\partial \kappa} [\omega * \tilde{s}(\kappa,0) * (W_r + \psi(\kappa,0)) + (\dot{s}(\kappa,0)) * (W_p + \psi(\kappa,0) - P(\kappa,0))]$$

$$= f_r(.)\theta_p(.) * (\psi_\kappa(\kappa,0) - \phi_\kappa(\kappa,0)) * (\omega W_r - W_p + \omega\psi(\kappa,0) - \phi(\kappa,0) + P(\kappa,0))$$

$$+ (\omega \tilde{s} - \dot{s}\theta_r(.))\, \psi_\kappa(\kappa,0)$$

The top line is positive because of the assumptions $\psi_\kappa(\kappa,s) \geq \phi_\kappa(\kappa,s)$ and $\omega \geq \frac{W_p}{W_r}$. The bottom line can be negative if $(\omega\tilde{s} - \dot{s}\theta_r(.)) < 0$; but we know that

$$\frac{\partial}{\partial \kappa} (\omega \tilde{s} - \dot{s}\theta_r(.)) = (\omega + \theta_r(.)) \frac{\partial \tilde{s}}{\partial \kappa} \geq 0$$

Therefore, we know that there are two possibilities as we increase κ from $\underline{\kappa}$: (i) the weighted social utility of accepting increases with κ over the range $[\underline{\kappa}, \infty)$; or (ii) $\exists \bar{\kappa}$ such that the weighted social utility of accepting decreases over the range $[\underline{\kappa}, \bar{\kappa})$ and increases over the range $(\bar{\kappa}, \infty)$

If we denote the weighted social utility of accepting as $U(\kappa)$ and the weighted social utility of rejecting as the constant R then there are two possibilities.

Under (i), $U(\underline{\kappa}) < R$, $\lim_{\kappa\to\infty} U(\kappa) = \infty > R$ and $U_\kappa(.) > 0$. So $\exists \kappa^*(\underline{\kappa}, \infty)$ such that $U(\kappa^*) = R$, $U(\kappa) < R$ if $\kappa < \kappa^*$ and $U(\kappa) > R$ if $\kappa > \kappa^*$.

Under (ii), since $U_\kappa(.) < 0$ for $\kappa \in [\underline{\kappa}, \bar{\kappa}]$, $U(\bar{\kappa}) < U(\underline{\kappa}) < R$. So $U(\bar{\kappa}) < R$, $\lim_{\kappa\to\infty} U(\kappa) = \infty > R$ and $U_\kappa(.) > 0$. So $\exists \kappa^* \in (\bar{\kappa}, \infty)$ such that $U(\kappa^*) = R$, $U(\kappa) < R$ if $\kappa < \kappa^*$ and $U(\kappa) > R$ if $\kappa > \kappa^*$.

B2 Proof of proposition 2

Define:

$$H(\kappa, \omega) \equiv U(\kappa^*, \omega) - R = 0$$

Then, by the implicit function theorem:

$$\frac{\partial \kappa^*}{\partial \omega} = \frac{-H_\omega(\kappa^*, \omega)}{H_\kappa(\kappa^*, \omega)} = \frac{-\left[\tilde{s}(\kappa^*, 0) * (W_r + \psi(\kappa^*, 0)) - \tilde{s}(0,0) * (W_r + \psi(0,0))\right]}{U_\kappa(.)}$$

From the proof in B1 we know that at κ^*, $U_\kappa(.) > 0$. Further, if $\kappa^* < 0$ then $\tilde{s}(\kappa^*, 0) * (W_r + \psi(\kappa^*, 0)) < \tilde{s}0, 0) * (W_r + \psi(0,0))$ so $\frac{\partial \kappa^*}{\partial \omega} > 0$. Also, if $\kappa^* > 0$ then $\tilde{s}(\kappa^*, 0) * (W_r + \psi(\kappa^*, 0)) > \tilde{s}(0,0) * (W_r + \psi(0,0))$ so $\frac{\partial \kappa^*}{\partial \omega} < 0$.

Appendix C

Additional Tables for Chapter 1

Table C1: Dynamic Difference-in-differences

VARIABLES	(1) ln(price)	(2) ln(price)	(3) ln(price)	(4) ln(price)
Success X Relative Year -2	-0.004	-0.003	-0.004	-0.003
	(0.003)	(0.002)	(0.003)	(0.002)
Success X Relative Year -1	-0.007*	-0.006	-0.007*	-0.006*
	(0.004)	(0.004)	(0.004)	(0.004)
Success X Relative Year 0	-0.011**	-0.011**	-0.010*	-0.010**
	(0.005)	(0.005)	(0.005)	(0.005)
Success X Relative Year 1	-0.011**	-0.011**	-0.012**	-0.011**
	(0.006)	(0.005)	(0.006)	(0.005)
Success X Relative Year 2	-0.016**	-0.017***	-0.016**	-0.016***
	(0.006)	(0.006)	(0.006)	(0.006)
Success X Relative Year 3	-0.023***	-0.019***	-0.022***	-0.019***
	(0.008)	(0.007)	(0.008)	(0.007)
Observations	1,237,943	1,177,861	1,237,706	1,177,683
R-squared	0.576	0.843	0.609	0.856
Time FE	✓	✓	✓	✓
Appeal FE	✓	✓		
Appeal X Postcode sector FE			✓	✓
Controls		✓		✓

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C2: First-stage with controls

VARIABLES	(1) Success*Postappeal	(2) Success*Postappeal	(3) Success*Postappeal
Leniency*Postappeal	0.680*** (0.136)	0.664*** (0.135)	0.664*** (0.135)
Observations	1,236,401	1,176,342	1,176,164
Time FE	✓	✓	✓
Appeal FE	✓	✓	
Appeal X Postcode Sector FE			✓
Controls		✓	✓
KP F-stat	25.15	24.04	24.07

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C3: Effect of overturning the Local Authority's decision (OLS)

VARIABLES	(1) ln(price)	(2) ln(price)	(3) ln(price)	(4) ln(price)
Success*Postappeal	-0.010** (0.004)	-0.010** (0.004)	-0.009** (0.004)	-0.010** (0.004)
Observations	1,237,943	1,177,861	1,237,706	1,177,683
R-squared	0.576	0.843	0.609	0.856
Time FE	✓	✓	✓	✓
Appeal FE	✓	✓		
Appeal X Postcode sector FE			✓	✓
Controls		✓		✓

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C4: Effect of overturning the Local Authority's decision (IV)

VARIABLES	(1) ln(price)	(2) ln(price)	(3) ln(price)	(4) ln(price)
Success*Postappeal	0.073** (0.035)	0.060* (0.033)	0.072** (0.034)	0.062* (0.033)
Observations	1,236,401	1,176,342	1,236,164	1,176,164
R-squared	-0.002	0.623	-0.003	0.623
Time FE	✓	✓	✓	✓
Appeal FE	✓	✓		
Appeal X Postcode sector FE			✓	✓
Controls		✓		✓
KP F-stat	25.15	24.04	25.08	24.07

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C5: Effect by distance from appeal site (500m)

VARIABLES	(1) ln(price)	(2) ln(price)	(3) ln(price)	(4) ln(price)
Success*Postappeal	-0.011**	-0.009**	0.039	0.051
	(0.005)	(0.005)	(0.033)	(0.031)
Observations	381,188	360,933	380,811	360,563
R-squared	0.621	0.857	-0.001	0.615
Time FE	✓	✓	✓	✓
Appeal X Postcode sector FE	✓	✓	✓	✓
Controls		✓		✓
Sample	<500m	<500m	<500m	<500m
KP F-stat			31.68	28.96

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C6: Effect by distance from appeal site (500m to 1km)

VARIABLES	(1) ln(price)	(2) ln(price)	(3) ln(price)	(4) ln(price)
Success*Postappeal	-0.009**	-0.010**	0.095**	0.073*
	(0.004)	(0.004)	(0.042)	(0.038)
Observations	856,381	816,622	855,217	815,474
R-squared	0.626	0.861	-0.004	0.619
Time FE	✓	✓	✓	✓
Appeal FE				
Appeal X Postcode sector FE	✓	✓	✓	✓
Controls		✓		✓
Sample	500m to 1km	500m to 1km	500m to 1km	500m to 1km
KP F-stat			20.13	19.85

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C7: Effect by supply shock (small suply shock)

VARIABLES	(1) ln(price)	(2) ln(price)	(3) ln(price)	(4) ln(price)
Success*Postappeal	-0.008	-0.009*	0.081**	0.070**
	(0.005)	(0.005)	(0.036)	(0.035)
Observations	960,928	913,192	960,333	912,619
R-squared	0.626	0.860	-0.003	0.618
Time FE	✓	✓	✓	✓
Appeal FE				
Appeal X Postcode sector FE	✓	✓	✓	✓
Controls		✓		✓
Sample	Small shock	Small shock	Small shock	Small shock
KP F-stat			25.64	24.70

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C8: Effect by supply shock (large suply shock)

VARIABLES	(1) ln(price)	(2) ln(price)	(3) ln(price)	(4) ln(price)
Success*Postappeal	-0.004	-0.007	-0.009	-0.011
	(0.005)	(0.005)	(0.100)	(0.083)
Observations	276,778	264,491	275,831	263,545
R-squared	0.518	0.831	0.000	0.646
Time FE	✓	✓	✓	✓
Appeal FE				
Appeal X Postcode sector FE	✓	✓	✓	✓
Controls		✓		✓
Sample	Large shock	Large shock	Large shock	Large shock
KP F-stat			1.470	1.838

Notes: Standard errors clustered at outward code level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C9: Regression of indicator for successful appeal on proxy variables

VARIABLES	(1) Success
Number of live appeal cases	0.000
	(0.000)
Development with 1 residence	-0.021**
	(0.010)
Development with 2 residences	-0.006
	(0.010)
Number of residences	0.001***
	(0.000)
Number of residences squared	-0.000***
	(0.000)
Bespoke timeline	-0.071**
	(0.035)
Costs applied for	0.176***
	(0.010)
Number of days to resolve appeal	0.000
	(0.000)
Number of linked appeal calses	-0.012
	(0.012)
Intermediary involved	0.060***
	(0.009)
LPA historical number of appeals	0.007***
	(0.001)
LPA historical appeals success rate	-0.023***
	(0.005)
Observations	28,519
R-squared	0.095
Year FE	✓
Procedure FE	✓
Appeal X Year FE	✓
Key Words	✓

Notes: Standard errors clustered at LPA level.
Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Appendix D

The Effect of the Fire on Quantities

First, did the Fire result in fewer properties being rebuilt? To answer this, we run a difference-in-differences regression where we collapse the data to the parish-level:

$$ln\left(Properties_{jt}\right) = \alpha_j + \delta PostFire_t + \beta Burned_j \times PostFire_t + \gamma' X_{jt} + \epsilon_{jt}$$

$ln\left(Properties_{jt}\right)$ is the log number of properties in parish j in period t. The two periods are before the Fire and after the Fire. $Burned_j$ is an indicator variable that denotes whether the parish experienced damage from the Fire. $PostFire_t$ is an indicator variable for the period after the Fire. X_{jt} is a vector of controls. These include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. These variables are interacted with post-Fire. Broader locations-by-post fixed effects are also included to control for geographical characteristics. Finally, α_j are parish fixed effects. The standard errors are clustered at the parish-level.

table D1 presents the results of this regression. The coefficient estimates of β are negative. This is expected as the plague wiped out about a quarter of London's population so we should expect fewer properties to be rebuilt in the immediate aftermath since there are now fewer people to house. The results in column 4 suggest that burned parishes saw a highly statistically significant decrease of around 67.6% properties as compared to unburned parishes. In addition, the reduction in the number of properties is consistent with post-Fire regulations that stipulated that properties needed to be of a certain minimum size.

Table D1: Effect of Fire on the number of properties

	(1)	(2)	(3)	(4)
VARIABLES		ln(No. Properties)		
Parish Burned X Post Fire	-1.059***	-1.256***	-0.790***	-0.676**
	(0.240)	(0.267)	(0.283)	(0.258)
Observations	140	140	140	140
R-squared	0.205	0.354	0.429	0.460
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	70	70	70	70

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Second, did the total number of hearths in the parishes decline after the rebuilding? To answer this, we again run a difference-in-differences regression where we collapse the data to the parish-level:

$$ln\left(Hearths_{jt}\right) = \alpha_j + \delta PostFire_t + \beta Burned_j \times PostFire_t + \gamma' X_{jt} + \epsilon_{jt}$$

$ln\left(Hearths_{jt}\right)$ is the log number of hearths in parish j in period t. The other variables are the same as previously defined and the standard errors are clustered at the parish-level. table D2 presents the results from this regression. The results are similar to what happens to the total number of properties being rebuilt after the Fire (table D1). In particular, the coefficient estimates are negative.

Table D2: Effect of Fire on the number of hearths

	(1)	(2)	(3)	(4)
VARIABLES		ln(No. Hearths)		
Parish Burned X Post Fire	-0.866***	-1.043***	-0.643**	-0.518*
	(0.232)	(0.251)	(0.293)	(0.271)
Observations	140	140	140	140
R-squared	0.147	0.324	0.387	0.427
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	70	70	70	70

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Appendix E

Discussion about Jensen's Inequality

Running a regression with a log transformed dependent variable could result in an opposite treatment effect as compared to if we were to run the regression without taking logs. To see this, consider the following stylized example:[1]

Table E1: Stylized example about Jensen's inequality

Parish 1

Property	Hearths	ln(Hearths)
1	10	2.30
2	20	3.00
Parish Average	15	2.65

Parish 2

Property	Hearths	ln(Hearths)
1	5	1.61
2	30	3.40
Parish Average	18	2.51

In this example, the average number of hearths per property and the total number of hearths are higher in parish 2 than in parish 1. However, if we ran a regression using the log of each property's hearths on a parish dummy, we will find that parish 2 on average has fewer log hearths per property.

This stylized example shows the possibility that this could happen but it does not mean that this would definitely happen for other values. Therefore, what we do is

[1]We would like to thank David Weinstein for providing us with this stylized example.

to replicate this stylized example using the actual data that we have. In particular, we collapse the data into two groups – burned and unburned parishes. We then compare the differences of the averages (in both logs and without logs) across the burned and unburned groups in the pre- and post-Fire periods. Table E2 reports the averages from this exercise. It shows us that both a regression without logs and a regression with logs will give us a positive effect. In particular, from the regression without logs we will get a difference-in-differences effect of: $(6.07-4.70)-(4.74-4.41)=1.04$. In the regression with logs we get: $(1.80-1.55)-(1.56-1.48)=0.17$. Fortunately, the reversal of signs issue does not happen when we use the actual data.

Table E2: Stylized example using actual data

Unburned

Post-Fire	Parish Average: Hearths	Parish Average: ln(Hearths)
0	4.41	1.48
1	4.74	1.56

burned

Post-Fire	Parish Average: Hearths	Parish Average: ln(Hearths)
0	4.70	1.55
1	6.07	1.80

The second approach would be to directly run the quality regression without taking logs on the left hand-side variable. This guarantees that the regression will not suffer from Jensen's inequality issues but it comes at the expense of failing the parallel trends assumption and the results being potentially driven by the skewed data. Nevertheless, table E3 shows that the estimated coefficient from this regression is positive. Since both the regressions in logs and without logs give us positive coefficient estimates, this should allay the worry that the estimated effect could actually be positive without taking logs but negative with a log transformation due to Jensen's inequality.

Table E3: Effect of Fire on the number of hearths per property (no logs)

	(1)	(2)	(3)	(4)
VARIABLES	No. Hearths per Property			
Parish Burned X Post Fire	0.614	0.569	0.407	0.524
	(0.371)	(0.399)	(0.497)	(0.468)
Observations	79,730	79,730	79,730	79,730
R-squared	0.002	0.003	0.003	0.004
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	70	70	70	70

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix F

Additional Figures and Tables for Chapter 2

Figure F1: Case characteristics over time

Figure F1.1: Years left in tenancy

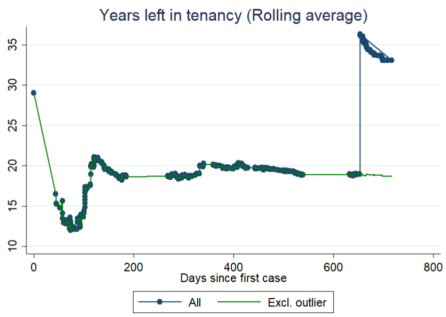


Figure F1.2: Pre-Fire rent

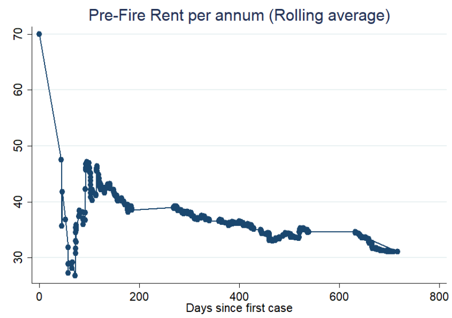


Figure F1.3: Pre-Fire fine

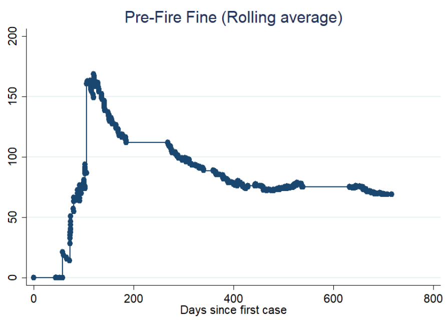


Figure F1.4: Pre-Fire improvements

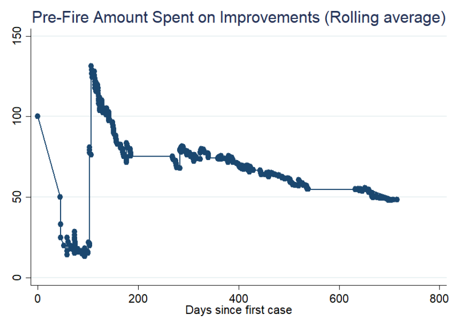


Figure F1.5: Degree from owner

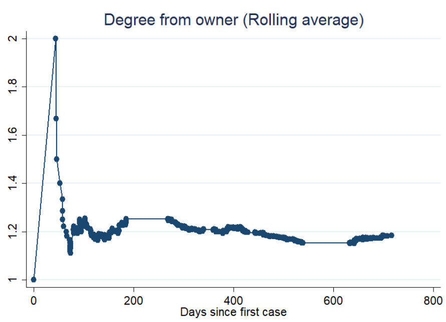


Figure F1.6: Number of parishes

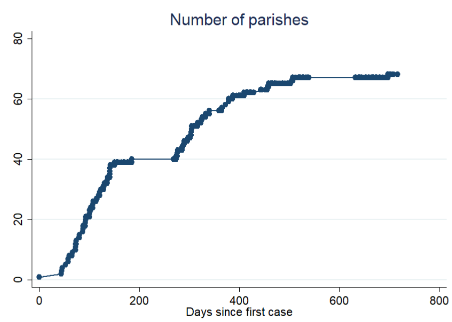


Figure F2: Binscatter of the effect of Fire on the number of hearths per property (All controls)

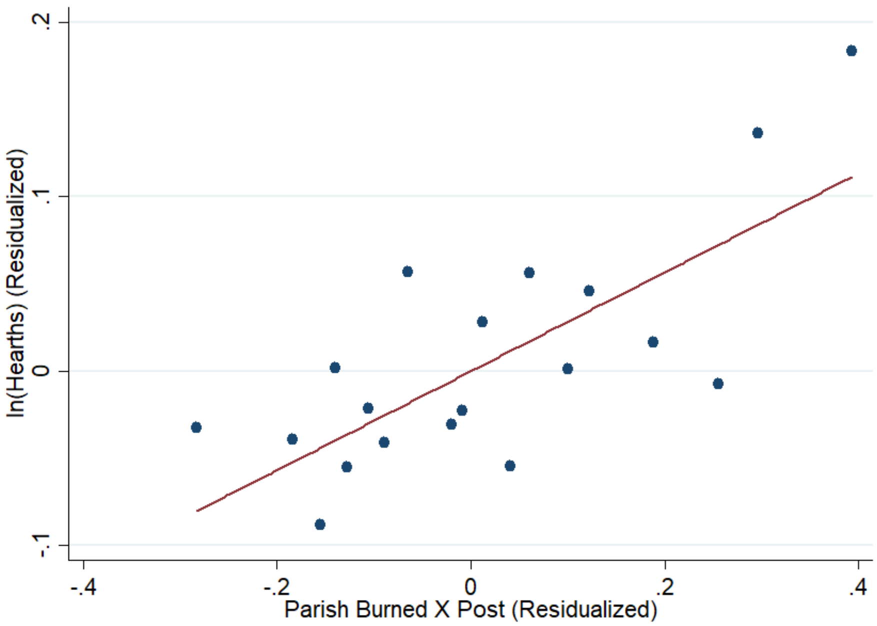


Figure F3: Binscatter of the effect of legal rulings on the number of hearths per property (All controls)

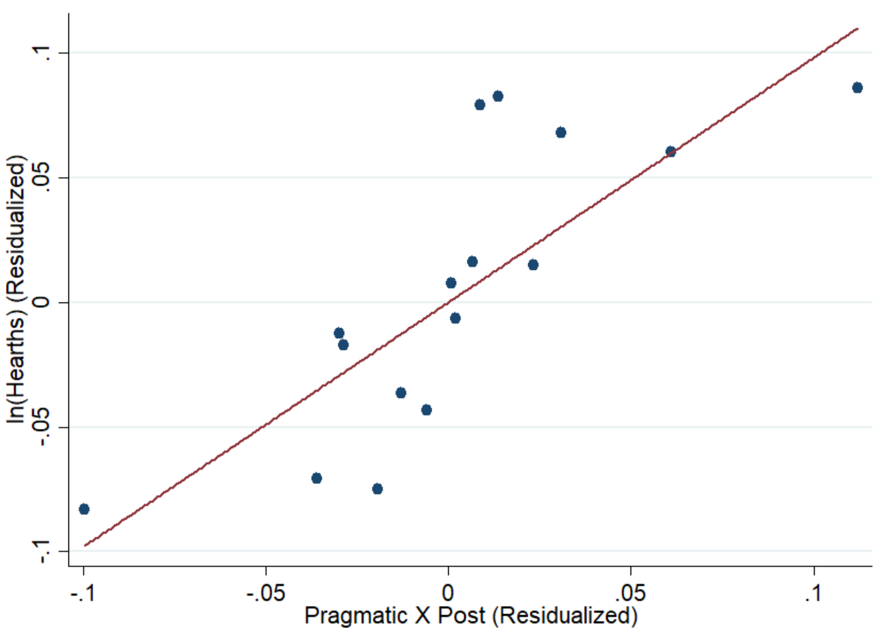


Figure F4: Binscatter of the first-stage (All controls)

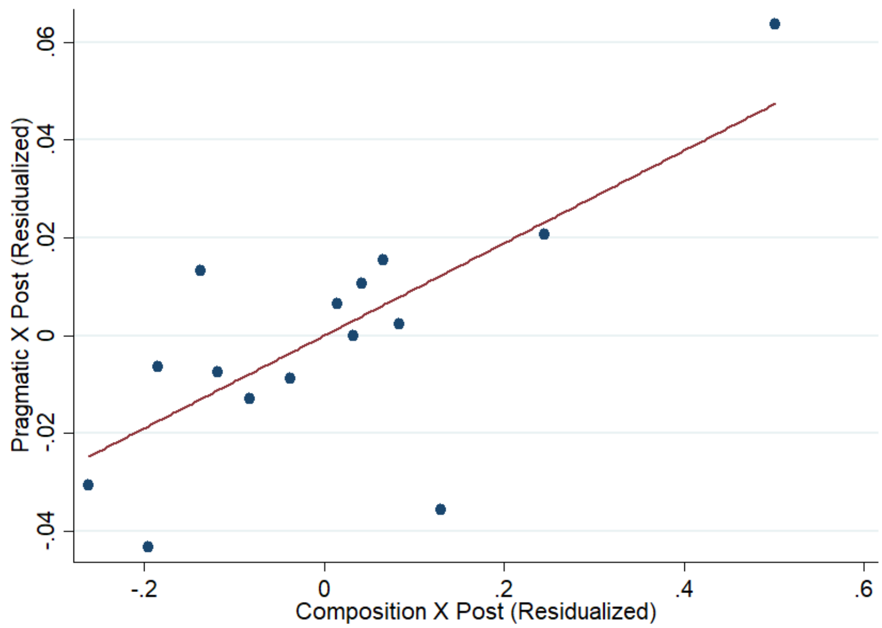


Figure F5: Binscatter of the reduced-form (All controls)

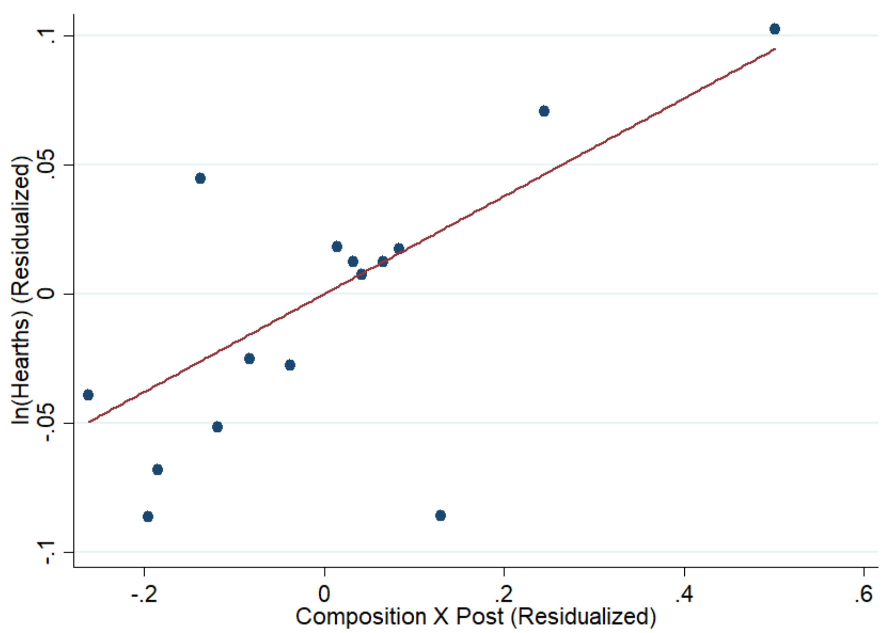


Table F1: Share of properties in each parish that went to the Fire Court

Parish	Cases	No. properties before Fire	Share
St Botolph Aldersgate	1	3969	0.000
St Giles Cripplegate	1	4967	0.000
St Andrew Holborn	4	1757	.002
All Hallows Staining	1	158	.006
St Antholin Budge Row & St John Walbrook	3	204	.015
St Bartholomew The Less	2	124	.016
St Mary Somerset & St Mary Mounthaw	4	223	.018
St Sepulchre Without Newgate	19	999	.019
All Hallows Barking	9	455	.02
Whitefriars Precinct	4	204	.02
St Bride Fleet Street	34	1614	.021
St Alphage London Wall	4	174	.023
St Martin Ludgate	6	241	.025
Holy Trinity The Less & St Michael Queenhithe	6	226	.027
St Andrew Hubbard & St Mary At Hill	7	255	.027
St Benet Pauls Wharf & St Peter Pauls Wharf	8	298	.027
St Mary Staining & St Michael Wood Street	3	112	.027
St Martin Vintry & St Michael Paternoster Royal	3	105	.029
St Alban Wood Street & St Olave Silver Street	8	257	.031
All Hallows The Great & All Hallows The Less	14	417	.034
St Mary Aldermary & St Thomas Apostle	4	109	.037
St Dunstan In The West	40	1001	.04
St Botolph Billingsgate & St George Botolph Lane	6	148	.041
St Gabriel Fenchurch Street & St Margaret Pattens	6	148	.041
St Swithin London Stone & St Mary Bothaw	7	171	.041
St Dunstan In The East	16	378	.042
Christchurch Newgate Street & St Leonard Foster Lane	24	468	.051
St Magnus The Martyr & St Margaret New Fish Street	12	235	.051
St Peter Le Poer	6	117	.051
St Nicholas Olave & St Nicholas Cole Abbey	6	107	.056
St Matthew Friday Street & St Peter Westcheap	7	117	.06
St Martin Pomeroy & St Olave Old Jewry	7	109	.064
St Michael Le Querne & St Vedast Foster Lane	17	238	.071
St Andrew By The Wardrobe & St Anne Blackfriars	12	167	.072
St Lawrence Jewry & St Mary Magdalen Milk Street	17	231	.074
St Mary Colechurch & St Mildred Poultry	8	108	.074
St Mary Magdalen Old Fish Street	5	68	.074
St Clement Eastcheap & St Martin Orgar	5	65	.077
St Mary Aldermanbury	13	153	.085
St Mary Le Bow & All Hallows Honey Lane & St Pancras Soper Lane	20	194	.103
St Margaret Moses & St Mildred Bread Street	12	107	.112
St Stephen Walbrook & St Benet Sherehog	15	109	.138
St Augustine Watling Street & St Faith Under St Paul	29	203	.143
St Gregory By St Paul	53	364	.146
St Lawrence Pountney & St Mary Abchurch	5	17	.294
All Hallows Bread Street & St John The Evangelist Friday Street	8	27	.296

Table F2: Comparing parishes before and after the Fire (burned vs unburned)

VARIABLES	(1) ln(No. hearths)	(2)
Parish Burned	0.101	0.311***
	(0.084)	(0.070)
Observations	42,174	34,919
R-squared	0.141	0.197
Parish controls	✓	✓
Broader location controls	✓	✓
Number of clusters	70	70
Sample	Pre-Fire	Post-Fire

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table F3: Effect of Fire on the number of hearths per property (Dropping parishes which merged after the Fire)

VARIABLES	(1)	(2) ln(No. hearths)	(3)	(4)
Parish Burned X Post Fire	0.241**	0.224**	0.222*	0.277**
	(0.105)	(0.109)	(0.129)	(0.120)
Observations	69,466	69,466	69,466	69,466
R-squared	0.007	0.008	0.008	0.010
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	40	40	40	40

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table F4: Effect of Fire on the number of hearths per property
(Using different control groups - Nearby sample)

VARIABLES	(1)	(2)	(3)	(4)
	ln(No. hearths)			
Parish Burned X Post Fire	0.336***	0.296***	0.327***	0.392***
	(0.094)	(0.080)	(0.116)	(0.121)
Observations	48,103	48,103	48,103	48,103
R-squared	0.013	0.021	0.022	0.025
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	61	61	61	61

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table F5: Effect of Fire on the number of hearths per property
(Using different control groups - Further away sample)

VARIABLES	(1)	(2)	(3)	(4)
	ln(No. hearths)			
Parish Burned X Post Fire	0.225**	0.136	0.137	0.236*
	(0.105)	(0.134)	(0.151)	(0.137)
Observations	62,466	62,466	62,466	62,466
R-squared	0.007	0.009	0.010	0.015
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	60	60	60	60

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table F6: Effect of Fire on the number of hearths per property (Applying the inverse hyperbolic sine transform to hearths)

	(1)	(2)	(3)	(4)
VARIABLES	ln(No. hearths)			
Parish Burned X Post Fire	0.124	0.101	0.056	0.086
	(0.112)	(0.134)	(0.159)	(0.159)
Observations	79,730	79,730	79,730	79,730
R-squared	0.002	0.007	0.008	0.009
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	70	70	70	70

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table F7: Effect of Fire on the number of hearths per property (Using a Poisson pseudo-likelihood regression)

	(1)	(2)	(3)	(4)
VARIABLES	No. hearths			
Parish Burned X Post Fire	0.129	0.103	0.070	0.101
	(0.088)	(0.108)	(0.129)	(0.121)
Observations	79,730	79,730	79,730	79,730
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	70	70	70	70

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table F8: Effect of Fire on the number of hearths per property (Trimming extreme values of the outcome variable)

VARIABLES	(1)	(2)	(3)	(4)
	ln(No. hearths)			
Parish Burned X Post Fire	0.124**	0.101*	0.086	0.103
	(0.056)	(0.055)	(0.065)	(0.065)
Observations	64,402	64,402	64,402	64,402
R-squared	0.004	0.005	0.005	0.006
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	70	70	70	70

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table F9: Comparing parishes before and after the Fire (by legal rulings)

VARIABLES	(1)	(2)
	ln(No. hearths)	
Pragmatic	-0.036	0.769**
	(0.289)	(0.304)
Observations	21,017	10,565
R-squared	0.140	0.196
Parish controls	✓	✓
Broader location controls	✓	✓
Number of clusters	46	46
Sample	Pre-Fire	Post-Fire

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table F10: Effect of legal rulings on the number of hearths per property (Adding in other dimensions of the rulings as controls)

VARIABLES	(1)	(2) ln(No. hearths)	(3)	(4)
Pragmatic X Post Fire	1.050***	1.031***	0.926***	0.835***
	(0.294)	(0.306)	(0.249)	(0.224)
Avg. change in tenancy length X Post Fire	0.003	0.000	-0.000	-0.001
	(0.003)	(0.002)	(0.002)	(0.002)
Avg. change in rent X Post Fire	-0.016**	-0.013*	-0.012**	-0.009*
	(0.007)	(0.007)	(0.006)	(0.005)
Observations	31,582	31,582	31,582	31,582
R-squared	0.018	0.026	0.027	0.031
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	46	46	46	46

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table F11: Effect of legal rulings on the number of hearths per property (Dropping parishes which merged after the Fire)

VARIABLES	(1)	(2) ln(No. hearths)	(3)	(4)
Pragmatic X Post Fire	2.128***	2.038***	1.917***	2.025***
	(0.657)	(0.249)	(0.304)	(0.409)
Observations	24,384	24,384	24,384	24,384
R-squared	0.024	0.037	0.037	0.037
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	17	17	17	17

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table F12: Effect of legal rulings on the number of hearths per property (Applying the inverse hyperbolic sine transform to hearths)

VARIABLES	(1)	(2) ln(No. hearths)	(3)	(4)
Pragmatic X Post Fire	1.122**	1.091***	1.015***	0.915***
	(0.422)	(0.338)	(0.246)	(0.254)
Observations	32,383	32,383	32,383	32,383
R-squared	0.009	0.017	0.018	0.021
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	46	46	46	46

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table F13: Effect of legal rulings on the number of hearths per property (Using a Poisson pseudo-likelihood regression)

VARIABLES	(1)	(2) No. hearths	(3)	(4)
Pragmatic X Post Fire	0.921**	0.934***	0.885***	0.784***
	(0.424)	(0.285)	(0.210)	(0.206)
Observations	32,383	32,383	32,383	32,383
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	46	46	46	46

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table F14: Effect of legal rulings on the number of hearths per property (Trimming extreme values of the outcome variable)

VARIABLES	(1)	(2) ln(No. hearths)	(3)	(4)
Pragmatic X Post Fire	0.852***	0.709***	0.663***	0.620***
	(0.220)	(0.193)	(0.151)	(0.148)
Observations	25,965	25,965	25,965	25,965
R-squared	0.010	0.015	0.015	0.016
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	46	46	46	46

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table F15: First-stage by different subsamples

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
			Pragmatic X Post Fire			
Majority royalist in judging panels X Post	0.135***	0.046	0.038	-0.010	0.179***	0.125**
	(0.048)	(0.049)	(0.062)	(0.060)	(0.032)	(0.051)
Observations	20,737	10,845	1,726	8,419	21,437	19,984
Parish FE	✓	✓	✓	✓	✓	✓
Post FE	✓	✓	✓	✓	✓	✓
Sample	Church not destroyed	Church destroyed	Abutting walls	Within walls	Outside walls	Hearth tercile 1
Number of clusters	11	35	5	33	8	14

VARIABLES	(7)	(8)	(9)	(10)	(11)	(12)
			Pragmatic X Post Fire			
Majority royalist in judging panels X Post	0.036	0.133**	0.022	-0.096	0.159***	-0.011
	(0.097)	(0.054)	(0.099)	(0.064)	(0.036)	(0.072)
Observations	4,094	7,504	3,094	5,470	23,018	6,690
Parish FE	✓	✓	✓	✓	✓	✓
Post FE	✓	✓	✓	✓	✓	✓
Sample	Hearth tercile 2	Hearth tercile 3	Size tercile 1	Size tercile 2	Size tercile 3	Peers tercile 1
Number of clusters	16	16	19	17	10	21

VARIABLES	(13)	(14)	(15)	(16)	(17)	(18)
			Pragmatic X Post Fire			
Majority royalist in judging panels X Post	0.153***	0.159***	0.073	0.166***	0.062	0.188***
	(0.035)	(0.043)	(0.069)	(0.024)	(0.065)	(0.021)
Observations	17,727	7,165	16,383	15,199	19,135	12,447
Parish FE	✓	✓	✓	✓	✓	✓
Post FE	✓	✓	✓	✓	✓	✓
Sample	Peers tercile 2	Peers tercile 3	Doctors quantile 1	Doctors quantile 2	Military quantile 1	Military quantile 2
Number of clusters	11	14	27	19	39	7

Notes: All regressions include parish FEs and post FE. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table F16: Reduced-form – Effect of Royalist majority on the number of hearths per property

VARIABLES	(1)	(2) ln(No. hearths)	(3)	(4)
Majority royalist in judging panels X Post	0.270***	0.221**	0.213***	0.189**
	(0.098)	(0.087)	(0.069)	(0.074)
Observations	31,582	31,582	31,582	31,582
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	46	46	46	46

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table F17: IV – Effect of legal rulings on the number of hearths per property (Adding in other dimensions of the rulings as controls)

VARIABLES	(1)	(2) ln(No. hearths)	(3)	(4)
Pragmatic X Post Fire	2.198*	2.483**	2.141**	1.951**
	(1.152)	(1.058)	(0.835)	(0.866)
Avg. change in tenancy length X Post Fire	0.004	0.001	0.001	-0.000
	(0.004)	(0.003)	(0.002)	(0.002)
Avg. change in rent X Post Fire	-0.005	-0.004	-0.006	-0.003
	(0.016)	(0.011)	(0.009)	(0.008)
Observations	31,582	31,582	31,582	31,582
R-squared	0.012	0.017	0.022	0.027
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	46	46	46	46
KP F-stat	6.427	3.548	8.863	10.63

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table F18: IV – Effect of legal rulings on the number of hearths per property (Dropping parishes which merged after the Fire)

	(1)	(2)	(3)	(4)
VARIABLES		ln(No. hearths)		
Pragmatic X Post Fire	2.065***	1.886***	1.904***	1.945***
	(0.549)	(0.404)	(0.361)	(0.441)
Observations	24,384	24,384	24,384	24,384
R-squared	0.024	0.036	0.037	0.037
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	17	17	17	17
KP F-stat	16.12	7.611	40.30	25.22

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table F19: IV – Effect of legal rulings on the number of hearths per property (Applying the inverse hyperbolic sine transform to hearths)

	(1)	(2)	(3)	(4)
VARIABLES		ln(No. hearths)		
Pragmatic X Post Fire	1.866***	2.271**	1.909**	1.712*
	(0.636)	(0.942)	(0.719)	(0.872)
Observations	32,383	32,383	32,383	32,383
R-squared	0.005	0.010	0.015	0.019
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	46	46	46	46
KP F-stat	10.03	3.831	9.195	10.43

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table F20: IV – Effect of legal rulings on the number of hearths per property (Trimming extreme values of the outcome variable)

VARIABLES	(1)	(2)	(3)	(4)
	ln(No. hearths)			
Pragmatic X Post Fire	1.663***	1.688***	1.560***	1.490***
	(0.386)	(0.556)	(0.438)	(0.439)
Observations	25,965	25,965	25,965	25,965
R-squared	0.001	0.005	0.008	0.010
Parish FE	✓	✓	✓	✓
Post FE	✓	✓	✓	✓
Parish controls X Post FE		✓	✓	✓
Broader location X Post FE			✓	✓
Pre-fire hearth tercile X Post FE				✓
Number of clusters	46	46	46	46
KP F-stat	12.07	4.716	9.581	10.83

Notes: Parish controls include the number of properties in the parish before the Fire, the share of peers, high-ranking military personnel and doctors living in the parish. Standard errors are clustered at the parish level. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Appendix G

Comparing OLS to 2SLS Estimates

Table F1: Comparing OLS to 2SLS

	(1) OLS	(2) 2SLS - 'ever seated'	(3) 2SLS - first draw
Prop black - final jury	-0.2616	-0.3278	-0.1823
	(0.1221)	(0.1328)	(0.1672)
Defendant characteristics	Yes	Yes	Yes
Charge details	Yes	Yes	Yes
Judge	Yes	Yes	Yes
Year/season	Yes	Yes	Yes
County	Yes	Yes	Yes
Observations	1128	1128	1128

It might be instructive to compare the IV results to the OLS to understand the possible biases (table F1). The 2SLS regression using 'ever seated' blacks as an instrument produced estimates more negative than the OLS. The 2SLS estimate using blacks in the first draw was less negative than that of the OLS.

My prior is that the variation in the firs draw is less endogenous and hence I interpret these results as suggesting there is a negative bias in the OLS. As a corollary, 'ever seated' seems to be endogenous because it did not reduce the negative bias but exacerbated it instead.

Bibliography

Acemoglu, D., Johnson, S., and Robinson, J. (2005). The rise of europe: Atlantic trade, institutional change, and economic growth. *American economic review*, 95(3):546–579.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American economic review*, 91(5):1369–1401.

Aizer, A. and Doyle Jr, J. J. (2015). Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. *The Quarterly Journal of Economics*, 130(2):759–803.

Akai, N. and Sakata, M. (2002). Fiscal decentralization contributes to economic growth: evidence from state-level cross-section data for the united states. *Journal of urban economics*, 52(1):93–108.

Albouy, D. and Ehrlich, G. (2018). Housing productivity and the social cost of land-use restrictions. *Journal of Urban Economics*, 107(July 2017):101–120.

Angelucci, C., Meraglia, S., and Voigtländer, N. (2017). How merchant towns shaped parliaments: From the norman conquest of england to the great reform act. Technical report, National Bureau of Economic Research.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434):444–455.

Anwar, S., Bayer, P., and Hjalmarsson, R. (2012). The impact of jury race in criminal trials. *Quarterly Journal of Economics*, 127(2):1017–1055.

Anwar, S., Bayer, P., and Hjalmarsson, R. (2014). The Role of Age in Jury Selection and Trial Outcomes. *The Journal of Law and Economics*, 57(4):1001–1030.

Anwar, S., Bayer, P., and Hjalmarsson, R. (2016). A Jury of Her Peers: The Impact of the First Female Jurors on Criminal Convictions.

Aradillas-Lopez, A. and Rosen, A. M. (2019). Inference in ordered response games with complete information.

Au, C.-C. and Henderson, J. V. (2006). Are chinese cities too small? *The Review of Economic Studies*, 73(3):549–576.

Autor, D. H., Palmer, C. J., and Pathak, P. A. (2014). Housing market spillovers: Evidence from the end of rent control in cambridge, massachusetts. *Journal of Political Economy*, 122(3):661–717.

Baldu, D. C., Woodworthz, G., Zuckenan, D., Weiner, N. A., and Broffit, B. (2001). The Use of Peremptory Challenges in Capital Murder Trials: A Legal and Empirical Analysis. *Journal of Constitutional Law*, 3(1).

Barrett, B. E. (2007). Detecting Bias in Jury Selection. *The American Statistician*, 61(4):296–301.

Basu, K. (2000). *Prelude to political economy: A study of the social and political foundations of economics*. Oxford University Press.

Baum-Snow, N. and Marion, J. (2009). The effects of low income housing tax credit developments on neighborhoods. *Journal of Public Economics*, 93(5-6):654–666.

Bhuller, M., Dahl, G. B., Løken, K. V., and Mogstad, M. (2020). Incarceration, recidivism, and employment. *Journal of Political Economy*, 128(4):1269–1324.

Bleakley, H. and Lin, J. (2012). Portage and path dependence. *The quarterly journal of economics*, 127(2):587–644.

Calabrese, S., Epple, D., and Romano, R. (2007). On the political economy of zoning. *Journal of Public Economics*, 91(1-2):25–49.

Card, D., Devicienti, F., and Maida, A. (2014). Rent-sharing, holdup, and wages: Evidence from matched panel data. *Review of Economic Studies*, 81(1):84–111.

Che, Y.-K. and Hausch, D. B. (1999). Cooperative investments and the value of contracting. *American Economic Review*, 89(1):125–147.

Cheshire, P., Nathan, M., and Overman, H. (2014). *Urban Economics and Urban Policy*. Number 15105 in Books. Edward Elgar Publishing.

Cheshire, P. and Sheppard, S. (2002). The welfare economics of land use planning. *Journal of Urban Economics*, 52(2):242–269.

Collins, W. J. and Shester, K. L. (2013). Slum clearance and urban renewal in the united states. *American Economic Journal: Applied Economics*, 5(1):239–73.

Collinson, R. and Ganong, P. (2018). How do changes in housing voucher design affect rent and neighborhood quality? *American Economic Journal: Economic Policy*, 10(2):62–89.

Cooter, R. (1998). Expressive law and economics. *The Journal of Legal Studies*, 27(S2):585–607.

Corkindale, J. (1999). Land development in the united kingdom: Private property rights and public policy objectives. *Environment and Planning A*, 31:2053–2070.

Dahl, G. B., Kostøl, A. R., and Mogstad, M. (2014). Family welfare cultures. *The Quarterly Journal of Economics*, 129(4):1711–1752.

Dale, T. C. (1931). *Inhabitants of London in 1638*. Society of Genealogists:London.

Davis, D. R. and Weinstein, D. E. (2002). Bones, bombs, and break points: the geography of economic activity. *American economic review*, 92(5):1269–1289.

Davoodi, H. and Zou, H.-f. (1998). Fiscal decentralization and economic growth: A cross-country study. *Journal of Urban economics*, 43(2):244–257.

Dell, M. (2010). The persistent effects of peru's mining mita. *Econometrica*, 78(6):1863–1903.

Dericks, G. H. and Koster, H. R. (2021). The billion pound drop: the blitz and agglomeration economies in london. *Journal of Economic Geography*, 21(6):869–897.

Diamond, R. and McQuade, T. (2018). Who Wants Affordable Housing in their Backyard? An Equilibrium Analysis of Low Income Property Development. *Journal of Political Economy*.

Diamond, R., Mcquade, T., and Qian, F. (2019). The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco. *American Economic Review*, 109(9).

Diamond, S. S., Peery, D., Dolan, F. J., and Dolan, E. (2009). Achieving Diversity on the Jury: Jury Size and the Peremptory Challenge. *Journal of Empirical Legal Studies*, 6(3):425–449.

dictionary of national biography, O. (2014). Oxford dictionary of national biography (2014). https://www.history.ox.ac.uk/oxford-dictionary-national-biography.

Djankov, S., La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2003). Courts. *The Quarterly Journal of Economics*, 118(2):453–517.

Dobbie, W., Goldin, J., and Yang, C. (2018). The Effects of Pre-Trial Detention on Conviction , Future Crime , and Employment: Evidence from Randomly Assigned Judges. *American Economic Review*, 108(2):201–240.

Doyle Jr, J. J. (2008). Child protection and adult crime: Using investigator assignment to estimate causal effects of foster care. *Journal of political Economy*, 116(4):746–770.

Epple, D., Romer, T., and Filimon, R. (1988). Community development with endogenous land use controls. *Journal of Public Economics*, 35(2):133–162.

Field, J. F. (2008). *Reactions and responses to the Great Fire: London and England in the later 17th century*. PhD thesis, Newcastle University.

Field, J. F. (2017). *London, Londoners and the Great Fire of 1666: Disaster and Recovery*. Routledge.

Fischel, W. A. (2001). Why Are There NIMBYs ?

Flanagan, F. X. (2015). Peremptory Challenges and Jury Selection. *The Journal of Law and Economics*, 58(385):1–26.

Flanagan, F. X. (2017). Race , Gender , and Juries : Evidence from North Carolina.

Galasso, A. and Schankerman, M. (2015). Patents and cumulative innovation: Causal evidence from the courts. *The Quarterly Journal of Economics*, 130(1):317–369.

Garrioch, D. (2016). 1666 and london's fire history: a re-evaluation. *The Historical Journal*, 59(2):319–338.

Gazette, T. L. (1666). The london gazette (1666). https://www.bl.uk/learning/timeline/item103652.html.

Glaeser, E. L., Gyourko, J., and Saks, R. E. (2005). Why have housing prices gone up? *American Economic Review*, 95(2):329–333.

Greenstone, M. and Gallagher, J. (2008). Does hazardous waste matter? Evidence from the housing market and the superfund program.

Grosso, C. M., Brien, B. O., and Capital, P.-b. N. C. (2013). A Stubborn Legacy : The Overwhelming Importance of Race in Jury Selection in 173 Post-Batson North Carolina Capital Trials A Stubborn Legacy : The Overwhelming Importance of Race in Jury Selection in. *Iowa Law Review*, 97(10).

Grout, P. A. (1984). Investment and wages in the absence of binding contracts: A nash bargaining approach. *Econometrica: Journal of the Econometric Society*, pages 449–460.

Gyourko, J. and Molloy, R. (2015). *Regulation and Housing Supply*, volume 5. Elsevier B.V., 1 edition.

Hadfield, G. K. and Weingast, B. R. (2012). What is law? a coordination model of the characteristics of legal order. *Journal of Legal Analysis*, 4(2):471–514.

Hamilton, B. W. (1975). Zoning and property taxation in a system of local governments. *Urban studies*, 12(2):205–211.

Hart, O. and Moore, J. (1988). Incomplete contracts and renegotiation. *Econometrica: Journal of the Econometric Society*, pages 755–785.

Heblich, S., Trew, A., and Zylberberg, Y. (2021). East-side story: Historical pollution and persistent neighborhood sorting. *Journal of Political Economy*, 129(5):1508–1552.

Heckman, J. J. and Vytlacil, E. (2005). Structural equations, treatment effects, and econometric policy evaluation 1. *Econometrica*, 73(3):669–738.

Henderson, J. V., Regan, T., and Venables, A. J. (2021). Building the city: from slums to a modern metropolis. *The Review of Economic Studies*, 88(3):1157–1192.

Hilber, C. A. and Robert-Nicoud, F. (2013). On the origins of land use regulations: Theory and evidence from US metro areas. *Journal of Urban Economics*, 75(1):29–43.

Hilber, C. A. and Vermeulen, W. (2016). The Impact of Supply Constraints on House Prices in England. *Economic Journal*, 126(591):358–405.

Hoekstra, M. and Street, B. (2021). The effect of own-gender jurors on conviction rates. *The Journal of Law and Economics*, 64(3):513–537.

Hornbeck, R. and Keniston, D. (2017). Creative destruction: Barriers to urban growth and the great boston fire of 1872. *American Economic Review*, 107(6):1365–98.

Hsieh, C. T. and Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2):1–39.

Humphries, J. E., Mader, N., Tannenbaum, D., and Dijk, W. V. (2018). Does eviction cause poverty ? Quasi-experimental evidence from Cook County , IL.

Jacoby, H. G. and Mansuri, G. (2008). Land tenancy and non-contractible investment in rural pakistan. *The Review of Economic Studies*, 75(3):763–788.

Jedwab, R., Johnson, N. D., and Koyama, M. (2019). Pandemics, places, and populations: Evidence from the black death.

Jha, S. (2015). Financial asset holdings and political attitudes: evidence from revolutionary england. *The Quarterly Journal of Economics*, 130(3):1485–1545.

Johnson, L. M. (2018). Accessing Jury Selection Data in a Pre-Digital Environment. *American Journal of Trial Advocacy*, 40(1).

Jones, P. E. (1966). *The fire court: calendar to the judgments and decrees of the court of judicature appointed to determine differences between landlords and tenants as to rebuilding after the great fire, volume I.* The Corporation of London: London.

Jones, P. E. (1970). *The fire court: calendar to the judgments and decrees of the court of judicature appointed to determine differences between landlords and tenants as to rebuilding after the great fire, volume II.* The Corporation of London: London.

Kline, P. and Moretti, E. (2014). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *The Quarterly journal of economics*, 129(1):275–331.

Koster, H. (2020). The welfare effects of greenbelt policy: Evidence from england.

Kostøl, A., Mogstad, M., Setzler, B., et al. (2019). Disability benefits, consumption insurance, and household labor supply. *American Economic Review*, 109(7):2613–54.

Krugman, P. (1991). History versus expectations. *The Quarterly Journal of Economics*, 106(2):651–667.

La Porta, R., Lopez-de Silanes, F., Pop-Eleches, C., and Shleifer, A. (2004). Judicial checks and balances. *Journal of Political Economy*, 112(2):445–470.

La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2008). The economic consequences of legal origins. *Journal of economic literature*, 46(2):285–332.

Libecap, G. D. and Lueck, D. (2011). The Demarcation of Land and the Role of Coordinating Institutions. *Journal of Political Economy*, 119(3):426–467.

McAdams, R. H. (2000). A focal point theory of expressive law. *Virginia Law Review*, pages 1649–1729.

McAdams, R. H. (2005). The expressive power of adjudication. *U. Ill. L. Rev.*, page 1043.

Michaels, G. and Rauch, F. (2018). Resetting the urban network: 117–2012. *The Economic Journal*, 128(608):378–412.

Miguel, E. and Roland, G. (2011). The long-run impact of bombing vietnam. *Journal of development Economics*, 96(1):1–15.

Morrison, M., DeVaul-Fetters, A., and Gawronski, B. (2016). Stacking the Jury. *Personality and Social Psychology Bulletin*, 42(8):1129–1141.

Myerson, R. B. (2004). Justice, institutions, and multiple equilibria. *Chi. J. Int'l L.*, 5:91.

Nash Jr, J. F. (1950). The bargaining problem. *Econometrica: Journal of the econometric society*, pages 155–162.

North, D. C. and Weingast, B. R. (1989). Constitutions and commitment: the evolution of institutions governing public choice in seventeenth-century england. *The journal of economic history*, 49(4):803–832.

Noye, U. (2015). Blackstrikes A Study of the Racially Disparate Use of Peremptory Challenges by the Caddo Parish District Attorney's office. Technical report.

of London, M. (2011). Creating fire! https://www.museumoflondon.org.uk/discover/creating-fire-fire.

Ortalo-Magné, F. and Prat, A. (2014). On the political economy of urban growth: Homeownership versus affordability. *American Economic Journal: Microeconomics*, 6(1 D):154–181.

Parkhomenko, A. (2020). Local Causes and Aggregate Implications of Land Use Regulation.

Pike, A., Rodríguez-Pose, A., Tomaney, J., Torrisi, G., and Tselios, V. (2012). In search of the 'economic dividend'of devolution: spatial disparities, spatial economic policy, and decentralisation in the uk. *Environment and Planning C: Government and Policy*, 30(1):10–28.

Porter, S. (1996). *The great fire of London*. Sutton Publishing: Gloucestershire.

Quigley, J. M. and Raphael, S. (2005). Regulation and the high cost of housing in california. *American Economic Review*, 95(2):323–328.

Rauch, J. E. (1993). Does history matter only when it matters little? the case of city-industry location. *The quarterly journal of economics*, 108(3):843–867.

Reddaway, T. F. (1940). *The rebuilding of London after the great fire*. Jonathan Cape: London.

Redding, S. J., Sturm, D. M., and Wolf, N. (2011). History and industry location: evidence from german airports. *Review of Economics and Statistics*, 93(3):814–831.

Richardson, J. (2001). *The annals of London: a year-by-year record of a thousand years of history*. Cassell: London.

Rodríguez-Pose, A. and Bwire, A. (2004). The economic (in) efficiency of devolution. *Environment and Planning A*, 36(11):1907–1928.

Rossi-hansberg, E., Sarte, P.-D., and Owens III, R. (2010). Housing Externalities Pierre-Daniel Sarte and Raymond Owens III. *The Journal of Poitical Economy*, 118(3):1–8.

Sainty, J. (1993). *The judges of England 1272-1990: a list of judges of the superior courts*. The Selden Society: London.

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *Quarterly Journal of Economics*, 125(3):1253–1296.

Satchell, M., Kitson, P., Newton, G., Shaw-Taylor, L., and Wrigley, T. (2018). *The judges of England 1272-1990: a list of judges of the superior courts*. UK Data Archive: Colchester, Essex.

Shertzer, A., Twinam, T., and Walsh, R. P. (2018). Zoning and the economic geography of cities. *Journal of Urban Economics*, 105(November 2017):20–39.

Siodla, J. (2015). Razing san francisco: The 1906 disaster as a natural experiment in urban redevelopment. *Journal of Urban Economics*, 89:48–61.

Stansel, D. (2005). Local decentralization and local economic growth: A cross-sectional examination of us metropolitan areas. *Journal of Urban Economics*, 57(1):55–72.

Tidmarsh, J. (2016). The english fire courts and the american right to civil jury trial. *U. Chi. L. Rev.*, 83:1893.

Tirole, J. (1986). Procurement and renegotiation. *Journal of Political Economy*, 94(2):235–259.

Turner, M., Haughwout, A., and van der Klaauw, W. (2014). Land Use Regulation and Welfare. *Econometrica*, 82(4):1341–1403.

Wright, R. F., Chavis, K., and Parks, G. S. (2018). The Jury Sunshine Project : Jury Selection Data as a Political Issue. *University of Illinois Law Review*, 2018(4).

Xie, D., Zou, H.-f., and Davoodi, H. (1999). Fiscal decentralization and economic growth in the united states. *Journal of Urban economics*, 45(2):228–239.

Zhang, T. and Zou, H.-f. (1998). Fiscal decentralization, public spending, and economic growth in china. *Journal of public economics*, 67(2):221–240.