Online Appendix for "Neighborhood Effects: Evidence from Wartime Destruction in London" (Not for Publication)

Stephen J. Redding*
Princeton University, NBER and CEPR

Daniel M. Sturm†
London School of Economics and CEPR

July 2024

Table of Contents

*Dept. Economics and SPIA, Julis Romo Rabinowitz Building, Princeton, NJ 08544. United States. Tel: 1 609 258 4016. Email: reddings@princeton.edu.

†Dept. Economics, Houghton Street, London, WC2A 2AE. United Kingdom. Tel: 44 20 7955 7522. Email: d.m.sturm@lse.ac.uk.

A Introduction

This online appendix contains additional supplementary material for the paper. In Section B, we provide further details on our theoretical model, including the derivations of all expressions in the paper. In Section C, we report extensions of this theoretical model. In Section D, we present further information on our estimation of the model's parameters. In Section E, we give additional details on our counterfactuals for wartime destruction and neighborhood effects. In Section F, we report additional empirical results and robustness checks. In Section G, we provide further details on the data sources and definitions.

B Theoretical Appendix

In this section of the Online Appendix, we provide further details on our theoretical model from Section 5 of the paper. The subsections follow the same structure as the subsections in the paper.

B1 Preferences

No additional derivations required.

B2 Production

No additional derivations required.

B3 Residence and Workplace Decisions

In this subsection of the Online Appendix, we derive the commuting probabilities and expected utility, as reported in the paper.

B3.1 Distribution of Utility

From the indirect utility function in equation (5) in the paper, we have the following monotonic relationship between idiosyncratic amenities ($z_{nit}^{o}(\psi)$) and utility ($u_{nit}^{o}(\omega)$):

$$z_{nit}^{o}(\psi)=\frac{u_{nit}^{o}(\psi)\,\kappa_{nit}^{o}\left(P_{nt}^{Y}\right)^{\alpha^{o}}\left(Q_{nt}\right)^{1-\alpha^{o}}}{B_{nt}^{o}w_{it}^{o}}. \tag{B.1}$$

Combining equation (B.1) the our assumption that idiosyncratic amenities ($z_{nit}^{o}(\psi)$) are drawn from an independent extreme value (Fréchet) distribution in equation (7) in the paper, we obtain the following distribution of utility for residence n, workplace i and occupation o:

$$G_{nit}^{o}(u)=e^{-\Psi_{nit}^{o}u^{-\epsilon^{o}}}, \qquad \Psi_{nit}^{o}\equiv\left(B_{nt}^{o}w_{it}^{o}\right)^{\epsilon^{o}}\left(\kappa_{nit}^{o}\left(P_{nt}^{Y}\right)^{\alpha^{o}}\left(Q_{nt}\right)^{1-\alpha^{o}}\right)^{-\epsilon^{o}}. \tag{B.2}$$

From all possible pairs of residence and workplace, each worker chooses the bilateral commute that offers the maximum utility. Since the maximum of a sequence of Fréchet distributed random variables is itself Fréchet distributed, the distribution of utility across all possible pairs of residence and workplace for occupation o is:

$$1 - G_t^o(u) = 1 - \prod_{k \in \mathbb{N}} \prod_{\ell \in \mathbb{N}} e^{-\Psi_{k\ell t}^o u^{-\epsilon^o}},$$

where the left-hand side is the probability that a worker has a utility greater than u, and the right-hand side is one minus the probability that the worker has a utility less than u for all possible pairs of residence and employment locations. Therefore we have:

$$G_t^o(u) = e^{-\Psi_t^o u^{-\epsilon^o}}, \qquad \Psi_t^o = \sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \Psi_{k\ell t}^o. \tag{B.3}$$

Given this Fréchet distribution for utility, expected utility is:

$$U_t^o = \mathbb{E}_t \left[u_t^o \right] = \int_0^\infty \epsilon^o \Psi_t^o u^{-\epsilon^o} e^{-\Psi_t^o u^{-\epsilon^o}} du, \tag{B.4}$$

where $\mathbb{E}_t[\cdot]$ denotes the expectations operator with respect to the distribution of idiosyncratic utility. Now define the following change of variables:

$$y = \Psi_t^o u^{-\epsilon^o}, \qquad dy = -\epsilon^o \Psi_t^o u^{-(\epsilon^o+1)} du. \tag{B.5}$$

Using this change of variables, expected utility can be written as:

$$U_t^o = \mathbb{E}_t \left[u_t^o \right] = \int_0^\infty \left(\Psi_t^o \right)^{1/\epsilon^o} y^{-1/\epsilon^o} e^{-y} dy, \tag{B.6}$$

which can be in turn written as:

$$U_t^o = \mathbb{E}_t \left[u_t^o \right] = \vartheta^o \left(\Psi_t^o \right)^{1/\epsilon^o}, \qquad \vartheta^o = \Gamma \left(\frac{\epsilon^o - 1}{\epsilon^o} \right), \tag{B.7}$$

where $\Gamma(\cdot)$ is the Gamma function. We thus obtain equation (13) in the paper:

$$U_t^o = \mathbb{E}_t \left[u_t^o \right] = \vartheta^o \left(\Psi_t^o \right)^{1/\epsilon^o} = \vartheta^o \left[\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \left(B_{kt}^o w_{\ell t}^o \right)^{\epsilon^o} \left(\kappa_{k\ell t}^o Q_{kt}^{1-\alpha^o} \right)^{-\epsilon^o} \right]^{1/\epsilon^o}, \tag{B.8}$$

where we have used our choice of numeraire ($P_{nt}^Y = 1$).

B3.2 Residence and Workplace Choices

Using the distribution of utility for pairs of residence and employment locations for a given occupation o, the probability that a worker chooses the bilateral commute from n to i out of all possible bilateral commutes is:

$$\begin{aligned}\lambda^o_{nit} &= \Pr\left[u^o_{nit} \geq \max\{u^o_{k\ell t}\}; \forall k, \ell\right], \\ &= \int_0^\infty \prod_{\ell \neq i} G^o_{n\ell t}(u) \left[\prod_{k \neq n} \prod_{\ell \in \mathbb{N}} G^o_{k\ell t}(u)\right] g^o_{nit}(u) du, \\ &= \int_0^\infty \prod_{k \in \mathbb{N}} \prod_{\ell \in \mathbb{N}} \epsilon^o \Psi^o_{nit} u^{-(\epsilon^o+1)} e^{-\Psi^o_{k\ell t} u^{-\epsilon^o}} du, \\ &= \int_0^\infty \epsilon^o \Psi^o_{nit} u^{-(\epsilon^o+1)} e^{-\Psi^o_t u^{-\epsilon^o}} du.\end{aligned} \tag{B.9}$$

Note that:

$$\frac{d}{du}\left[\frac{1}{\Psi^o_t} e^{-\Psi^o_t u^{-\epsilon^o}}\right] = \epsilon^o u^{-(\epsilon^o+1)} e^{-\Psi^o_t u^{-\epsilon^o}}. \tag{B.10}$$

Using this result to evaluate the integral above, the probability that the worker in occupation o chooses to live in location n and work in location i is given by the following expression, which corresponds to equation (12) in the paper:

$$\lambda^o_{nit} = \frac{E^o_{nit}}{\overline{E}^o_t} = \frac{\Psi^o_{nit}}{\Psi^o_t} = \frac{\left(B^o_{nt} w^o_{it}\right)^{\epsilon^o} \left(\kappa^o_{nit} Q^{1-\alpha^o}_{nt}\right)^{-\epsilon^o}}{\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \left(B^o_{kt} w^o_{\ell t}\right)^{\epsilon^o} \left(\kappa^o_{k\ell t} Q^{1-\alpha^o}_{kt}\right)^{-\epsilon^o}}, \tag{B.11}$$

where E^o_{nit} is the measure of commuters from residence n to workplace i in occupation o; $\overline{E}^o_t$ is the overall measure of workers from occupation o in the city as a whole; we have used our choice of numeraire ($P^Y_{nt} = 1$); and the term $\left(B^o_{nt} w^o_{it}\right) / \left(\kappa^o_{nit} Q^{1-\alpha^o}_{nt}\right)$ in the numerator and denominator captures amenity-adjusted real income.

Summing across workplaces i in equation (B.11), we obtain the probability that a worker from occupation o chooses to live in residence n:

$$\lambda^{Ro}_{nt} = \frac{R^o_{nt}}{\overline{E}^o_t} = \frac{\sum_{i \in \mathbb{N}} \left(B^o_{nt} w^o_{it}\right)^{\epsilon^o} \left(\kappa^o_{nit} Q^{1-\alpha^o}_{nt}\right)^{-\epsilon^o}}{\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \left(B^o_{kt} w^o_{\ell t}\right)^{\epsilon^o} \left(\kappa^o_{k\ell t} Q^{1-\alpha^o}_{kt}\right)^{-\epsilon^o}}, \tag{B.12}$$

where R^o_{nt} is the measure of residents from occupation o in location n. Intuitively, a location attracts a larger share of residents within an occupation (λ^{Ro}_{nt}) if it has higher amenities (B^o_{nt}), cheaper residential floor space (Q_{nt}), and lower commuting costs (κ_{nit}) to workplaces with higher wages (w^o_{it}).

Similarly, summing across residences n in equation (B.11), we obtain the probability that a worker from occupation o chooses workplace i:

$$\lambda^{Eo}_{it} = \frac{E^o_{it}}{\overline{E}^o_t} = \frac{\sum_{n \in \mathbb{N}} \left(B^o_{nt} w^o_{it}\right)^{\epsilon^o} \left(\kappa^o_{nit} Q^{1-\alpha^o}_{nt}\right)^{-\epsilon^o}}{\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \left(B^o_{kt} w^o_{\ell t}\right)^{\epsilon^o} \left(\kappa^o_{k\ell t} Q^{1-\alpha^o}_{kt}\right)^{-\epsilon^o}}, \tag{B.13}$$

where E^o_{it} is the measure of workers from occupation o employed in location i. Therefore, a location's attracts a larger share of employment within an occupation (λ^{Eo}_{it}) if has higher wages (w^o_{it}) and lower commuting costs (κ^o_{nit}) to residences with higher amenities (B^o_{nt}) and lower prices of residential floor space (Q_{nt}).

For the measure of workers from occupation o employed in location i (E^o_{it}), we can evaluate the conditional probability that they commute from location n (conditional on working in i):

$$\begin{aligned}
\lambda^{Eo}_{nit|i} = \frac{\lambda^o_{nit}}{\lambda^{Eo}_{it}} &= \Pr\left[u^o_{nit} \geq \max\{u^o_{rit}\}; \forall r\right], \\
&= \int_0^\infty \prod_{r \neq n} G^o_{rit}(u) g^o_{nit}(u) du, \\
&= \int_0^\infty e^{-\Psi^{Eo}_{it} u^{-\epsilon^o}} \epsilon^o \Psi^o_{nit} u^{-(\epsilon^o+1)} du.
\end{aligned} \tag{B.14}$$

where

$$\Psi^{Eo}_{it} \equiv \sum_{k \in \mathbb{N}} \left(B^o_{kt} w^o_{it}\right)^{\epsilon^o} \left(\kappa^o_{kit} \left(P^Y_{kt}\right)^{\alpha^o} \left(Q_{kt}\right)^{1-\alpha^o}\right)^{-\epsilon^o}. \tag{B.15}$$

Using the result (B.10) to evaluate the integral in equation (B.14), the probability that a worker in occupation o commutes from residence n to workplace i conditional on working in location i is:

$$\lambda^{Eo}_{nit|i} = \frac{\lambda^o_{nit}}{\lambda^{Eo}_{it}} = \frac{\left(B^o_{nt} w^o_{it}\right)^{\epsilon^o} \left(\kappa^o_{nit} \left(P^Y_{nt}\right)^{\alpha^o} \left(Q_{nt}\right)^{1-\alpha^o}\right)^{-\epsilon^o}}{\sum_{k \in \mathbb{N}} \left(B^o_{kt} w^o_{it}\right)^{\epsilon^o} \left(\kappa^o_{kit} \left(P^Y_{kt}\right)^{\alpha^o} \left(Q_{kt}\right)^{1-\alpha^o}\right)^{-\epsilon^o}}, \tag{B.16}$$

which simplifies to:

$$\lambda^{Eo}_{nit|i} = \frac{\left(B^o_{nt}\right)^{\epsilon^o} \left(\kappa^o_{nit} Q^{1-\alpha^o}_{nt}\right)^{-\epsilon^o}}{\sum_{k \in \mathbb{N}} \left(B^o_{kt}\right)^{\epsilon^o} \left(\kappa^o_{kit} Q^{1-\alpha^o}_{kt}\right)^{-\epsilon^o}}, \tag{B.17}$$

where again we have used our choice of numeraire ($P^Y_{nt} = 1$).

For the measure of residents in occupation o in location n (R^o_{nt}), we can evaluate the conditional probability that they commute to location i (conditional on living in location n):

$$\begin{aligned}
\lambda^{Ro}_{nit|n} = \frac{\lambda^o_{nit}}{\lambda^{Ro}_{nt}} &= \Pr\left[u^o_{nit} \geq \max\{u^o_{n\ell t}\}; \forall \ell\right], \\
&= \int_0^\infty \prod_{\ell \neq i} G^o_{n\ell t}(u) g^o_{nit}(u) du, \\
&= \int_0^\infty e^{-\Psi^{Ro}_{nt} u^{-\epsilon^o}} \epsilon^o \Psi^o_{nit} u^{-(\epsilon^o+1)} du,
\end{aligned} \tag{B.18}$$

where

$$\Psi^{Ro}_{nt} \equiv \sum_{\ell \in \mathbb{N}} \left(B^o_{nt} w^o_{\ell t}\right)^{\epsilon^o} \left(\kappa^o_{n\ell t} \left(P^Y_{nt}\right)^{\alpha^o} \left(Q_{nt}\right)^{1-\alpha^o}\right)^{-\epsilon^o}. \tag{B.19}$$

Using the result (B.10) to evaluate the integral in equation (B.18), the probability that a worker in occupation o commutes to location i conditional on living in location n is:

$$\lambda_{nit|n}^{Ro} = \frac{\lambda_{nit}^{o}}{\lambda_{nt}^{Ro}} = \frac{\left(B_{nt}^{o} w_{it}^{o}\right)^{\epsilon^{o}} \left(\kappa_{nit}^{o} \left(P_{nt}^{Y}\right)^{\alpha^{o}} \left(Q_{nt}\right)^{1-\alpha^{o}}\right)^{-\epsilon^{o}}}{\sum_{\ell \in \mathbb{N}} \left(B_{nt}^{o} w_{\ell t}^{o}\right)^{\epsilon^{o}} \left(\kappa_{n\ell t}^{o} \left(P_{nt}^{Y}\right)^{\alpha^{o}} \left(Q_{nt}\right)^{1-\alpha^{o}}\right)^{-\epsilon^{o}}}, \tag{B.20}$$

which simplifies to:

$$\lambda_{nit|n}^{Ro} = \frac{\left(w_{it}^{o}/\kappa_{nit}^{o}\right)^{\epsilon^{o}}}{\sum_{\ell \in \mathbb{N}} \left(w_{\ell t}^{o}/\kappa_{n\ell t}^{o}\right)^{\epsilon^{o}}}. \tag{B.21}$$

Commuter market clearing requires that the measure of workers from occupation o employed in each location i (E_{it}^{o}) equals the sum across all locations n of the measures of residents from occupation o (R_{nt}^{o}) times their conditional probabilities of commuting to i ($\lambda_{nit|n}^{Ro}$):

$$\begin{aligned} E_{it}^{o} &= \sum_{n \in \mathbb{N}} \lambda_{nit|n}^{Ro} R_{nt}^{o} \\ &= \sum_{n \in \mathbb{N}} \frac{\left(w_{it}^{o}/\kappa_{nit}^{o}\right)^{\epsilon^{o}}}{\sum_{\ell \in \mathbb{N}} \left(w_{\ell t}^{o}/\kappa_{n\ell t}^{o}\right)^{\epsilon^{o}}} R_{nt}^{o}, \end{aligned} \tag{B.22}$$

where, since there is a continuous measure of workers residing in each location, there is no uncertainty in the supply of workers to each employment location.

Residential income per capita in occupation o conditional on living in location n (v_{nt}^{o}) equals the wages in all possible workplace locations weighted by the probabilities of commuting to those locations conditional on living in n:

$$\begin{aligned} v_{nt}^{o} &= \mathbb{E}_{t}\left[w_{nt}^{o} | n\right] \\ &= \sum_{i \in \mathbb{N}} \lambda_{nit|n}^{Ro} w_{it}^{o}, \\ &= \sum_{i \in \mathbb{N}} \frac{\left(w_{it}^{o}/\kappa_{nit}^{o}\right)^{\epsilon^{o}}}{\sum_{\ell \in \mathbb{N}} \left(w_{\ell t}^{o}/\kappa_{n\ell t}^{o}\right)^{\epsilon^{o}}} w_{it}^{o}, \end{aligned} \tag{B.23}$$

which corresponds to equation (15) in the paper. Intuitively, expected worker income is high in locations that have low commuting costs (low κ_{nit}^{o}) to high-wage (w_{it}^{o}) employment locations.

Another implication of the Fréchet distribution of utility is that the distribution of utility in occupation o conditional on residing in location n and commuting to location i is the same across all bilateral pairs of locations with positive residents and employment in that occupation, and is equal to the distribution of utility for that occupation for the economy as a whole. To establish this result, note that the distribution of utility in occupation o conditional on residing in location n and commuting to location i is:

$$= \frac{1}{\lambda_{nit}^{o}} \int_{0}^{u} \prod_{s \neq i} G_{nst}^{o}(u) \left[\prod_{k \neq n} \prod_{\ell \in \mathbb{N}} G_{k\ell t}^{o}(u)\right] g_{nit}^{o}(u) du, \tag{B.24}$$

$$= \frac{1}{\lambda^o_{nit}} \int_0^u \left[\prod_{k \in \mathbb{N}} \prod_{\ell \in \mathbb{N}} e^{-\Psi^o_{k\ell t} u^{-\epsilon^o}} \right] \epsilon^o \Psi^o_{nit} u^{-(\epsilon^o+1)} du,$$

$$= \frac{\Psi_t}{\Psi^o_{nit}} \int_0^u e^{-\Psi^o_t u^{-\epsilon^o}} \epsilon^o \Psi^o_{nit} u^{-(\epsilon^o+1)} du,$$

$$= e^{-\Psi^o_t u^{-\epsilon^o}}.$$

On the one hand, lower land prices in location n or a higher wage in location i raise the utility of a worker in occupation o with a given realization of idiosyncratic amenities b, and hence increase the expected utility of residing in n and working in i. On the other hand, lower land prices or a higher wage induce workers in occupation o with lower realizations of idiosyncratic amenities b to reside in n and work in i, which reduces the expected utility of residing in n and working in i. With a Fréchet distribution of utility, these two effects exactly offset one another. Pairs of residence and employment locations with more attractive characteristics attract more commuters on the extensive margin until expected utility in occupation o is the same across all pairs of residence and employment locations within the economy.

We assume that commuting costs are a power function of travel times (τ_{nit}) using the transport network ($(\kappa^o_{nit})^{-\epsilon^o} = \tau_{nit}^{-\epsilon^o \kappa} = \tau_{nit}^{-\phi^o}$); $\phi^o \equiv \epsilon^o \kappa$ is the product of the elasticity of commuting flows to commuting costs (ϵ^o) and the elasticity of commuting costs to travel time (κ).

B4 Floor Space Market Clearing

No additional derivations required.

B5 General Equilibrium

In this subsection of the Online Appendix, we provide a further characterization of the general equilibrium of the model.

B5.1 Sufficient Condition for Uniqueness of Equilibrium

We now provide a sufficient condition for the existence of a unique equilibrium for the special case of the model with neither neighborhood effects nor agglomeration forces, in which case residential amenities and productivity in each location depend only on the exogenous characteristics of that location: $\overline{B}^o_{nt} = B^o(b_{nt}, D_{nt})$ and $\overline{A}_{nt} = A(a_{nt}, D_{nt})$. We combine the general equilibrium conditions of the model to obtain a system of equations that takes the required form to apply Theorem 1 from Allen, Arkolakis and Li (2024):

$$x_{ih} = f_{ijh}(x_j) = \sum_{j \in \mathbb{J}} \mathcal{K}_{ijh} \prod_{h' \in \mathbb{H}} x_{jh'}^{\alpha_{ihh'}}. \tag{B.25}$$

In our specification, types $j \in \mathbb{J}$ can be a combination of locations $i \in \mathbb{N}$ and occupations $o \in \mathbb{O}$, such that $|\mathbb{J}| = |\mathbb{N}| \times |\mathbb{O}|$. Interactions $h \in \mathbb{H}$ include residents by occupation, employment by occupation, output, the price of residential floor space, and the price of commercial floor space. We begin by rewriting each of the general equilibrium conditions in the form required to apply this theorem.

Population Mobility As a preliminary step, note that the population mobility condition for each occupation (B.8) can be re-written as:

$$\left(\frac{U_t^o}{\vartheta^o}\right)^{\epsilon^o} = \left[\sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{N}} \left(\overline{B}_{nt}^o w_{it}^o\right)^{\epsilon^o} \left(\tau_{nit}^{\kappa} Q_{nt}^{1-\alpha^o}\right)^{-\epsilon^o}\right]^{\frac{1}{\epsilon^o}}. \tag{B.26}$$

We now use this expression for expected utility for each occupation (U_t^o) to rewrite the other general equilibrium conditions of the model.

Residential Choice Probabilities Using the population mobility condition (B.26), the residential choice probabilities (B.12) for each occupation and location can be re-written as:

$$R_{nt}^o = \xi_t^o \left(\overline{B}_{nt}^o \left(\Phi_{nt}^{Ro}\right)^{\frac{1}{\epsilon^o}} Q_{nt}^{\alpha^o - 1}\right)^{\epsilon^o}, \tag{B.27}$$

where ξ_t^o is an endogenous scalar:

$$\xi_t^o \equiv \overline{E}_{it}^o \left(\frac{U_t^o}{\vartheta^o}\right)^{-\epsilon^o}, \tag{B.28}$$

and Φ_{nt}^{Ro} is a measure of residential commuting market access for each occupation that is a commuting cost weighted average of wages in each workplace:

$$\Phi_{nt}^{Ro} \equiv \sum_{i \in \mathbb{N}} \tau_{nit}^{-\epsilon^o \kappa} \left(w_{it}^o\right)^{\epsilon^o}. \tag{B.29}$$

Workplace Choice Probabilities Using the population mobility condition (B.26), the workplace choice probabilities (B.13) for each occupation and location can be re-written as:

$$E_{it}^{Eo} = \xi_t^o \left(w_{it}^o \left(\Phi_{it}^{Eo}\right)^{\frac{1}{\epsilon^o}}\right)^{\epsilon^o}, \tag{B.30}$$

where we defined the endogenous scalar ξ_t^o in equation (B.28) above and Φ_{it}^{Eo} is a measure of workplace commuting market access for each occupation that is a commuting cost weighted average of amenities and the price of residential floor space in each residence:

$$\Phi_{it}^{Eo} \equiv \sum_{n \in \mathbb{N}} \tau_{nit}^{-\epsilon^o \kappa} \left(\overline{B}_{nt}^o Q_{nt}^{\alpha^o - 1}\right)^{\epsilon^o}. \tag{B.31}$$

Residential Commuting Market Access From the workplace choice probabilities (B.30) for each occupation, we have the following relationship:

$$\left(w_{it}^{o}\right)^{\epsilon^{o}} = \frac{1}{\xi^{o}} \frac{E_{it}^{o}}{\Phi_{it}^{o}}.$$

Using this relationship, we can re-write residential commuting market access for each occupation (Φ_{nt}^{Ro}) in equation (B.29) as follows:

$$\Phi_{nt}^{Ro} = \frac{1}{\xi^{o}} \sum_{i \in \mathbb{N}} \tau_{nit}^{-\epsilon^{o}\kappa} \frac{E_{it}^{o}}{\Phi_{it}^{o}}. \tag{B.32}$$

Workplace Commuting Market Access From the residential choice probabilities (B.27) for each occupation, we have the following relationship:

$$\left(\overline{B}_{nt}^{o} Q_{nt}^{\alpha^{o}-1}\right)^{\epsilon^{o}} = \frac{1}{\xi_{t}^{o}} \frac{R_{nt}^{o}}{\Phi_{nt}^{Ro}}$$

Using this relationship, we can re-write workplace commuting market access for each occupation (Φ_{it}^{Eo}) in equation (B.31) as follows:

$$\Phi_{it}^{Eo} = \frac{1}{\xi^{o}} \sum_{n \in \mathbb{N}} \tau_{nit}^{-\epsilon^{o}\kappa} \frac{R_{nt}^{o}}{\Phi_{nt}^{Ro}}. \tag{B.33}$$

Output From the unit cost function in equation (8), the production technology for the final good implies that output (Y_{it}) is given by:

$$Y_{it} = \overline{A}_{it} \left(\frac{\mathbb{L}_{it}}{\beta}\right)^{\beta} \left(\frac{H_{it}^{E}}{1-\beta}\right)^{1-\beta}, \tag{B.34}$$

where $\mathbb{L}_{it}$ is a composite labor input.

From the labor cost index in equation (9), this composite labor input takes the following constant elasticity of substitution (CES) form:

$$\mathbb{L}_{it} = \left[\left(\gamma^{L} E_{it}^{L}\right)^{\frac{\sigma-1}{\sigma}} + \left(\gamma^{M} E_{it}^{M}\right)^{\frac{\sigma-1}{\sigma}} + \left(\gamma^{H} E_{it}^{H}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \tag{B.35}$$

which can be re-written as follows:

$$\mathbb{L}_{it}^{\frac{\sigma-1}{\sigma}} = \sum_{o \in \mathbb{O}} \left(\gamma^{o} E_{it}^{o}\right)^{\frac{\sigma-1}{\sigma}}. \tag{B.36}$$

Residential Income Using the definition of residential commuting market access for each occupation from equation (B.29) in equation (B.23), average residential income for each occupation (v_{nt}^{o}) can be written as:

$$\Phi_{nt}^{Ro} v_{nt}^{o} = \sum_{i \in \mathbb{N}} \tau_{nit}^{-\epsilon^{o}\kappa} \left(w_{it}^{o}\right)^{\epsilon^{o}+1}. \tag{B.37}$$

Wages From the Cobb-Douglas production technology (B.34) and composite labor input (B.35), the requirement that the wage for each occupation in each location (w_{it}^{o}) equals labor's value marginal product can be written as:

$$w_{it}^{o} = \beta \gamma^{o} \left(\gamma^{o} E_{it}^{o}\right)^{-\frac{1}{\sigma}} \mathbb{L}_{it}^{-\left(\frac{\sigma-1}{\sigma}\right)}, \tag{B.38}$$

where we have again used our choice of numeraire ($P_{it}^{Y} = 1$).

Residential and Commercial Floor Space Prices The prices of residential (Q_{nt}) and commercial (q_{nt}) floor space are determined by the market clearing conditions for residential and commercial floor space in equations (16) and (17) in the paper, respectively, given the supplies of residential and commercial floor space (H_{nt}^{R}, H_{nt}^{E}).

System of General Equilibrium Conditions Using equations (B.29), (B.31), (B.32), (B.33), (B.34), (B.36), (B.37), (B.38), (16) and (17), the system of general equilibrium conditions of the model can be written in the form of equation (B.25) as follows:

$$\begin{aligned}
R_{nt}^{o} &= \xi_{t}^{o} \left(B_{nt}^{o} \left(\Phi_{nt}^{Ro}\right)^{\frac{1}{\epsilon^{o}}} Q_{nt}^{\alpha^{o}-1} \right)^{\epsilon^{o}}, \\
E_{nt}^{Eo} &= \xi_{t}^{o} \left(w_{nt}^{o} \left(\Phi_{nt}^{Eo}\right)^{\frac{1}{\epsilon^{o}}} \right)^{\epsilon^{o}}, \\
\Phi_{nt}^{Ro} &= \frac{1}{\xi^{o}} \sum_{i \in \mathbb{N}} \tau_{nit}^{-\epsilon^{o}\kappa} E_{it}^{o} \left(\Phi_{it}^{o}\right)^{-1}, \\
\Phi_{nt}^{Eo} &= \frac{1}{\xi^{o}} \sum_{i \in \mathbb{N}} \tau_{nit}^{-\epsilon^{o}\kappa} R_{it}^{o} \left(\Phi_{it}^{Ro}\right)^{-1}, \\
Y_{nt} &= \overline{A}_{nt} \left(\frac{\mathbb{L}_{nt}}{\beta}\right)^{\beta} \left(\frac{H_{nt}^{E}}{1-\beta}\right)^{1-\beta}, \\
\mathbb{L}_{it}^{\frac{\sigma-1}{\sigma}} &= \sum_{o \in \mathbb{O}} \left(\gamma^{o} E_{it}^{o}\right)^{\frac{\sigma-1}{\sigma}}, \\
Q_{nt} &= \sum_{o \in \mathbb{O}} (1-\alpha^{o}) v_{nt}^{o} R_{nt}^{o} \left(H_{nt}^{R}\right)^{-1}, \\
w_{nt}^{o} &= \beta \gamma^{o} \left(\gamma^{o} E_{nt}^{o}\right)^{-\frac{1}{\sigma}} \mathbb{L}_{nt}^{-\left(\frac{\sigma-1}{\sigma}\right)}, \\
\Phi_{nt}^{Ro} v_{nt}^{o} &= \sum_{i \in \mathbb{N}} \tau_{nit}^{-\epsilon^{o}\kappa} \left(w_{it}^{o}\right)^{\epsilon+1}, \\
q_{nt} &= \frac{1-\beta}{\beta} \sum_{o \in \mathbb{O}} w_{nt}^{o} E_{nt}^{o} (H_{nt}^{E})^{-1},
\end{aligned}$$

where the supply of residential floor space (H_{nt}^{R}), the supply of commercial floor space (H_{nt}^{E}), amenities ($\overline{B}_{nt}^{o}$) and productivity ($\overline{A}_{nt}$) are exogenous in this baseline specification.

The exponents on the variables on the left-hand side of this system of equations can be represented as the following matrix:

$$\mathbf{\Lambda} = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \ldots & \Lambda_{1H} \\ \Lambda_{21} & \Lambda_{22} & \ldots & \Lambda_{2H} \\ \vdots & \vdots & \vdots & \vdots \\ \Lambda_{H1} & \Lambda_{H2} & \ldots & \Lambda_{HH} \end{bmatrix}.$$

The exponents on the variables on the right-hand side of this system of equations can be represented as the following matrix:

$$\mathbf{\Gamma} = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} & \ldots & \Gamma_{1H} \\ \Gamma_{21} & \Gamma_{22} & \ldots & \Gamma_{2H} \\ \vdots & \vdots & \vdots & \vdots \\ \Gamma_{H1} & \Gamma_{H2} & \ldots & \Gamma_{HH} \end{bmatrix}$$

Let $\mathrm{Y} \equiv |\mathbf{\Gamma}\mathbf{\Lambda}^{-1}|$ and denote the spectral radius (eigenvalue with the largest absolute value) of this matrix by $\rho(\mathrm{Y})$. From Theorem 1 in Allen, Arkolakis and Li (2024), a sufficient condition for the existence of a unique equilibrium (up to scale) is $\rho(\mathrm{Y}) \leq 1$.

C Theoretical Extensions

In this section of the Online Appendix, we report theoretical extensions of our baseline model. In Subsection C1, we consider non-homothetic preferences. In Section C2, we derive the relationship between residential amenities and wartime destruction in our baseline model from a construction sector technology.

C1 Non-homothetic Preferences

Given the limited household expenditure data before the Second World War, our baseline specification assumes separate homothetic Cobb-Douglas preferences for each worker group, with expenditure shares that differ exogenously across the three groups. We now show that similar results hold using the non-homothetic Cobb-Douglas preferences (Heterothetic Cobb-Douglas preferences) of Bohr, Mestieri and Robert-Nicoud (2023). We assume the following non-homothetic direct utility function for a worker residing in location n and working in location i:

$$\ln u_{ni} = \ln B_n - \ln \kappa_{ni} + \ln z_{ni} + \alpha_Y(u_{ni}) \ln c_{ni}^Y + \alpha_H(u_{ni}) \ln c_{ni}^H, \tag{C.1}$$

where B_n denotes residential amenities; κ_{ni} denotes commuting costs; z_{ni} is an idiosyncratic preference shock to the utility from residing in location n and working in location i; c_{ni}^Y is consumption of the tradable final good; c_{ni}^H is consumption of residential floor space; $(\alpha_Y(u_{ni}), \alpha_H(u_{ni}))$ are continuous and differentiable functions; we assume that this direct utility satisfies the conditions for the solution to (C.1) to be unique and for u_{ni} to be monotone and quasi-concave in (c_{ni}^{Y}, c_{ni}^{H}); and we also impose the following convenient cardinalization of utility that $\alpha_Y(u_{ni}) + \alpha_H(u_{ni}) = 1$, as in Bohr, Mestieri and Robert-Nicoud (2023).

The consumer's expenditure minimization problem is:

$$\min_{\{c_{ni}^{Y}, c_{ni}^{H}\}} \left\{ P^{Y} c_{ni}^{Y} + Q_n c_{ni}^{H} \right\},$$

$$\text{s.t.} \qquad \ln u_{ni} = \ln B_n + \alpha_Y(u_{ni}) \ln c_{ni}^{Y} + \alpha_H(u_{ni}) \ln c_{ni}^{H} - \ln \kappa_{ni} + \ln z_{ni},$$

where P^{Y} is the common price of the freely-traded final good across all locations and Q_n is the price of residential floor space. The Lagrangian is:

$$\mathcal{L} = P^{Y} c_{ni}^{Y} + Q_n C_{ni}^{H} - \xi \left[\alpha_Y(u_{ni}) \ln c_{ni}^{Y} + \alpha_H(u_{ni}) \ln c_{ni}^{U} - u_{ni} \right].$$

The first-order conditions are:

$$P^{Y} - \xi \frac{\alpha_Y(u_{ni})}{c_{ni}^{Y}} = 0,$$

$$Q_n - \xi \frac{\alpha_H(u_{ni})}{c_{ni}^{H}} = 0,$$

$$\alpha_Y(u_{ni}) \ln c_{ni}^{Y} + \alpha_H(u_{ni}) \ln c_{ni}^{H} - u_{ni} = 0.$$

These first-order conditions imply:

$$P^{Y} c_{ni}^{Y} = \xi \alpha_Y(u_{ni}),$$

$$Q_n c_{ni}^{H} = \xi \alpha_H(u_{ni}).$$

Using these relationships, we obtain the following expressions for equilibrium expenditure shares:

$$s_Y(u_{ni}) = \frac{P^{Y} c_{ni}^{Y}}{w_i} = \frac{P^{Y} c_{ni}^{Y}}{P^{Y} c_{ni}^{Y} + Q_n c_{ni}^{H}} = \frac{\alpha_Y(u_{ni})}{\alpha_Y(u_{ni}) + \alpha_H(u_{ni})}, \tag{C.2}$$

$$s_H(u_{ni}) = \frac{Q_n c_{ni}^{H}}{w_i} = \frac{Q_n c_{ni}^{H}}{P^{Y} c_{ni}^{Y} + Q_n c_{ni}^{H}} = \frac{\alpha_H(u_{ni})}{\alpha_Y(u_{ni}) + \alpha_H(u_{ni})}, \tag{C.3}$$

where expenditure equals the wage (w_i).

From equations (C.2) and (C.3), these expenditure shares $(s_Y(u_{ni}), s_H(u_{ni}))$ depend on utility alone and imply the following equilibrium consumption choices:

$$c_{ni}^{Y} = \frac{s_Y(u_{ni}) w_i}{P^{Y}},$$

$$c_{ni}^{H} = \frac{s_H(u_{ni}) w_i}{Q_n}.$$

Substituting these equilibrium consumption choices into the direct utility function (C.1), we obtain the following indirect utility function:

$$\ln u_{ni} = \ln B_n - \ln \kappa_{ni} + \ln z_{ni} + \ln w_i + \alpha_Y (u_{ni}) \ln \left(\frac{s_Y (u_{ni})}{P^Y} \right) + \alpha_H (u_{ni}) \ln \left(\frac{s_H (u_{ni})}{Q_n} \right). \quad \text{(C.4)}$$

Recalling that $\alpha_Y (u_{ni}) + \alpha_H (u_{ni}) = 1$, the expenditure shares (C.2) and (C.3) imply:

$$s_Y (u_{ni}) = \alpha_Y (u_{ni}),$$

$$s_H (u_{ni}) = \alpha_H (u_{ni}).$$

Using these results, we can rewrite indirect utility (C.4) as follows:

$$\ln u_{ni} = \ln B_n - \ln \kappa_{ni} + \ln z_{ni} + \ln w_i + s_Y (u_{ni}) \ln \left(\frac{s_Y (u_{ni})}{P^Y} \right) + s_H (u_{ni}) \ln \left(\frac{s_H (u_{ni})}{Q_n} \right). \quad \text{(C.5)}$$

For workers in different occupations who receive different wages (w_i), this non-homothetic specification (C.5) will induce differences in expenditure shares on consumption goods ($s_Y (u_{ni})$) and residential floor space ($s_H (u_{ni})$). For the empirically-relevant case in which higher-income workers have lower shares of residential floor space in expenditure ($s_H (u_{ni})$), this implies that higher-income workers' location decisions will be less sensitive to differences in the price of residential floor space (Q_n) than those of lower-income workers.

Note that this non-homothetic Cobb-Douglas specification (C.5) corresponds closely to our baseline specification with separate homothetic Cobb-Douglas utility for each worker group, and exogenous differences in expenditure shares, which can be written as:

$$\ln u^o_{ni} = \ln B^o_n - \ln \kappa^o_{ni} + \ln z^o_{ni} + \ln w^o_i + s^o_Y \ln \left(\frac{1}{P^Y} \right) + s^o_H \ln \left(\frac{1}{Q_n} \right), \quad \text{(C.6)}$$

where $s^o_Y = \alpha^o$ and $s^o_H = (1 - \alpha^o)$, and our baseline specification allows amenities, commuting costs and the distribution of idiosyncratic preferences to differ across occupations, as well as wages and expenditure shares.

C2 Quality-Adjusted Floor Space

In the paper, we assume a direct relationship between residential amenities and wartime destruction in equation (6). In this section of the Online Appendix, we derive this relationship from a technology for the construction sector.

C2.1 Price and Supply of Residential Floor Space

We observe the price of residential floor space (Q_{nt}) in the data. We assume that workers' location choices depend on the quality-adjusted price of residential floor space ($\widetilde{Q}_{nt}$), which is not directly observed in the data:

$$\widetilde{Q}_{nt} = Q_{nt}/\varsigma_{nt}, \tag{C.7}$$

where ς_{nt} is the quality of residential floor space.

We assume that indirect utility for worker ψ from occupation o residing in location n and working in location i ($u^o_{nit}(\psi)$) depends on her wage (w^o_{it}), the price of the homogenous final consumption good (P^Y_{nt}), the quality-adjusted price of residential floor space ($\widetilde{Q}_{nt}$), commuting costs (κ^o_{nit}), residential fundamentals (b^o_{nt}), neighborhood effects ($\mathbb{B}_{nt}$), and an idiosyncratic amenity draw ($z^o_{nit}(\psi)$) for each worker, according to the following Cobb-Douglas functional form:

$$u^o_{nit}(\psi) = \frac{b^o_{nt}\mathbb{B}^{\eta^o_R}_{nt} z^o_{nit}(\psi) w^o_{it}}{\kappa^o_{nit}\left(P^Y_{nt}\right)^{\alpha^o}\left(\widetilde{Q}_{nt}\right)^{1-\alpha^o}}. \qquad 0 < \alpha^o < 1. \tag{C.8}$$

We assume that residential floor space is supplied by a competitive construction sector. We assume that the perceived quality of residential floor space (ς_{nt}) depends on whether it was constructed before or after the Second World War:

$$\varsigma_{nt} = e^{\widetilde{\eta}^o_D D_{nt}}, \tag{C.9}$$

where D_{nt} is the fraction of the pre-war built-up area destroyed during the Second World War ($D_{nt} \in [0,1]$); $\widetilde{\eta}^o_D$ parameterizes the impact of post-war reconstruction on perceived building quality; and the exponential specification ensures that the quality of residential floor space is positive for all $D_{nt} \in [0,1]$.

We assume that the quantity of residential floor space (H^R_{nt}) depends on inputs of land (K_n) and capital (M_{nt}) according to the following Cobb-Douglas construction technology:

$$H^R_{nt} = M^{\mu}_{nt}\left((1-\theta_{nt})K_n\right)^{1-\mu}, \qquad 0 < \mu < 1, \tag{C.10}$$

where $(1-\theta_{nt})$ is the fraction of land allocated to commercial use, and capital is assumed to be in perfectly elastic supply from the wider economy at a constant price of R_t.

C2.2 Residence-Workplace Choices

Using equation (C.7) to substitute for the quality-adjusted price of residential floor space ($\widetilde{Q}_{nt}$), and equation (C.9) to substitute for the quality of floor space (ς_{nt}), we can re-write indirect utility (C.8) in the same form as equation (5) in the paper:

$$u^o_{nit}(\psi) = \frac{B^o_{nt} z^o_{nit}(\psi) w^o_{it}}{\kappa^o_{nit}\left(P^Y_{nt}\right)^{\alpha^o}\left(Q_{nt}\right)^{1-\alpha^o}}. \qquad 0 < \alpha^o < 1, \tag{C.11}$$

where residential amenities (B_{nt}^{o}) are given by:

$$B_{nt}^{o}=e^{\eta_D^o D_{nt}}\mathbb{B}_{nt}^{\eta_R^o}b_{nt}^{o},$$

where we have defined $\eta_D^o \equiv (1-\alpha^o)\widetilde{\eta}_D^o$; $\mathbb{B}_{nt}$ represents neighborhood effects; η_R^o parameterizes the strength of these neighborhood effects for occupation o; and b_{nt}^{o} corresponds to residential fundamentals.

C2.3 Supply of Residential Floor Space

We now derive the implications of this specification for the supply of residential floor space. From cost minimization and zero profits in construction, equilibrium payments for capital are a constant share of payments for residential floor space ($Q_{nt}H_{nt}^{R}$):

$$R_t M_{nt}=\mu Q_{nt}H_{nt}^{R}. \tag{C.12}$$

Using this equilibrium condition (C.12) to substitute for capital (M_{nt}) in the construction technology (C.10), we obtain a constant elasticity supply function for residential floor space:

$$H_{nt}^{R}=\left(\frac{\mu}{R_t}\right)^{\frac{\mu}{1-\mu}}Q_{nt}^{\frac{\mu}{1-\mu}}\left(1-\theta_{nt}\right)K_n. \tag{C.13}$$

An advantage our estimation procedure is that we are not required to make assumptions about the determinants of the supplies of residential and commercial floor space (H_{nt}^{R}, H_{nt}^{R}). Instead, we use the model to back out the implied values of these variables given the observed data on the other endogenous variables of the model.

When we undertake counterfactuals, our baseline specification holds the supplies of residential and commercial floor space (H_{nt}^{R}, H_{nt}^{R}) fixed, which is motivated by our empirical setting, in which the reallocation of land between residential and commercial use was heavily restricted following the Town and Country Planning Act of 1942. This baseline specification corresponds to the case in which $(1-\theta_{nt})$ is exogenously determined and $\mu=0$ in equation (C.13).

In robustness checks, we undertake counterfactuals allowing for endogenous responses in the supply of residential floor space to changes in its price, which corresponds to case in which $(1-\theta_{nt})$ is exogenously determined and $\mu>0$ in equation (C.13).

D Quantitative Analysis

In this section of the Online Appendix, we provide further details on our quantitative analysis from Section 6 of the paper. Our quantitative analysis has a sequential structure, such that we undertake our analysis in a number of steps. Each step uses results from the previous one, and imposes the minimal set of additional assumptions relative to the previous step.

This estimation procedure has a number of advantages. First, we are not required to make assumptions about the impact of wartime destruction on productivity or agglomeration forces in production to estimate the neighborhood effects parameters, because we condition on observed variables that directly control for these production characteristics. Second, we are not required to make assumptions about whether the model has a unique equilibrium or multiple equilibrium in this estimation, because we condition on the observed equilibrium in the data. Given this observed equilibrium and the structure of the model, we are able to estimate the neighborhood effects parameters, regardless of whether or not there could have been another (unobserved) equilibrium for the same parameter values.

D1 Preference and Production Parameters (Step 1)

We begin by calibrating the model's standard preference and production technology parameters using historical data from our empirical setting.

D1.1 Preference Parameters (α^o)

We calibrate the housing expenditure shares ($1 - \alpha^o$) using a British Ministry of Labor household expenditure survey from 1937-8. Of the roughly 10,000 households surveyed in the United Kingdom, the data for 623 households has survived, and has been digitized by a team lead by Ian Gazeley, a Professor in Economic History at the London School of Economics (Gazeley et al. 2016). Each household was surveyed four times where possible. Weekly income for these households ranges from 143 pence (approximately £1 in British currency before the metric conversion in 1971) to 5,845 pence (approximately £24.3). Housing expenditure in the survey is measured net of any income from lodgers. We compute expenditure on floor space as including net housing expenditure and expenditure on gas, coal, electricity and other fuel sources. We separate households into our three groups using the thresholds of £3 and £5 per week that were used by NSOL for the boundaries between the low, middle and high-income groups. Based on this classification, we calibrate the expenditure shares on residential floor space for each group using the mean expenditure shares across households within each group: $(1 - \alpha^L) = 0.26$ ($N = 155$); $(1 - \alpha^M) = 0.22$ ($N = 222$); and $(1 - \alpha^H) = 0.16$ ($N = 111$). Therefore, we find an intuitive pattern, in which residential floor space accounts for a lower share of expenditure for higher-income households.

As a check on these calibrated values, we use data from the general household expenditure survey undertaken by the Ministry of Labor in 1953-4. This large-scale survey reports disaggregated results by geographic region, including the Greater London region, which includes the LCC area and its surrounding outer suburbs. Again we compute expenditure on floor space as including expenditure on fuel, light and power (including gas, electricity, coal, coke and oil and hire of gas and electric appliances). We update the thresholds of £3 and £5 per week that were used by NSOL for the boundaries between the low, middle and high-income groups to 1953 prices using the Bank of England's inflation calculator, which yields thresholds of £7.7 and £12.8. Based on this classification, we find similar expenditure shares on residential floor space for our three groups as using the 1937-8 data: $(1-\alpha^L) = 0.24$; $(1-\alpha^M) = 0.17$; and $(1-\alpha^H) = 0.14$. Hence, we again find that the share of residential floor space in expenditure declines with household income. We use the 1937-8 expenditure shares discussed above for our baseline calibration.[1]

D1.2 Production Parameter (β)

We assume a value for the share of labor in production costs of $\beta = 0.55$, which lies in the middle of the range of 0.43-0.63 in Antràs and Voth (2003), and is close to the labor share of 56 percent for Britain in 1913 in Matthews, Feinstein and Odling-Smee (1982). The remaining share of production costs of $(1-\beta) = 0.45$ is attributed to commercial floor space, which we interpret as including capital (machinery, equipment, buildings and structures) and land.

D2 Commuting Parameters (Step 2)

We next estimate the model's commuting parameters using data on bilateral commuting flows. Pre-war commuting data are not disaggregated by worker group and are only available for the relatively aggregated spatial units of the 29 LCC boroughs. Therefore, we use post-war data on bilateral commuting flows by worker group, which are available for 356 Middle Super Output Areas in the LCC area from the 2011 Population Census. We compare our model's predictions for pre-war commuting patterns to the available data in specification checks below.

Re-writing the commuting probabilities (B.11), we estimate the following gravity equation for each group of workers separately:

$$\lambda_{nit}^o = \eta_{nt}^{Ro} \eta_{it}^{Lo} \tau_{nit}^{-\phi^o} \zeta_{nit}^o, \tag{D.1}$$

where recall that we assume that commuting costs are a power function of travel times using the transport network ($(\kappa_{nit}^o)^{-\epsilon^o} = \tau_{nit}^{-\epsilon^o \kappa} = \tau_{nit}^{-\phi^o}$); $\phi^o = \epsilon^o \kappa$ is a composite elasticity that equals the product of the elasticity of commuting flows to commuting costs (ϵ^o) and the elasticity of commuting costs to travel time (κ); η_{nt}^{Ro} are residence fixed effects that capture amenities (B_{nt}^o) and the cost of living ($Q_{nt}^{-(1-\alpha^o)}$) and vary by occupation; η_{it}^{Lo} are workplace fixed effects that capture wages (w_{nt}^o) and vary by occupation; we use the property that the denominator in equation (B.11) equals expected utility (U_t^o) from equation (B.8) to absorb this denominator into the fixed effects;

[1] We find a similar pattern in which the share of housing in expenditure declines with income in the data reported in Parliamentary Papers (1877), although those earlier data only include workers with relatively low income.

and ζ^o_{nit} is a stochastic error. We cluster the standard errors by residence and workplace to allow for correlated error components by residence and workplace.

Two related challenges in estimating the commuting gravity equation (D.1) using data for disaggregated spatial units and socioeconomic groups are zero bilateral commuting flows and granularity, such that the sample and population means of random variables can differ in small samples. To address these concerns, our baseline specification follows Santos Silva and Tenreyro (2006) and Dingel and Tintelnot (2023) in estimating the gravity equation (D.1) using the Poisson pseudo maximum likelihood (PPML) estimator and using model-predicted commuting shares in our counterfactuals below. As a robustness check, we also report results from a specification estimating the gravity equation (D.1) in logs using the linear fixed effects estimator.

Another challenge in estimating the commuting gravity equation (D.1) is that travel time depends on the transport network, which is likely to be endogenous, because railway lines in London were constructed by profit-seeking private-sector companies. In particular, bilateral pairs that have more commuters for unobserved reasons in the error term (ζ^o_{nit}) could have more bilateral transport connections, and hence lower bilateral travel times (τ_{nit}). To address this concern, we follow Heblich et al. (2020) in instrumenting bilateral travel times with bilateral straight-line distance. In our baseline PPML specification, we use the control function approach proposed by Wooldridge (2014), in which we include in the second-stage equation (D.1) the residuals from a first-stage regression of bilateral travel times on bilateral distance. In our robustness check using a log linear specification, we instrument for bilateral travel times with bilateral distance using two-stage least squares (2SLS).

In Columns (1)-(3) of Panel A of Table D.1, we report the results of estimating the gravity equation (D.1) for low, middle and high-income workers, respectively, using the PPML estimator and without instrumenting for bilateral travel time. For all three groups of workers, we find negative and statistically significant elasticities of bilateral commuting flows with respect to bilateral travel times. Lower-income workers have commuting elasticities that are larger in absolute magnitude than higher-income workers, which reflects the net effect of several forces. On the one hand, lower-income workers could have lower opportunity costs of time, which implies commuting elasticities that are smaller in absolute magnitude. On the other hand, lower-income workers' commuting decisions are plausibly more sensitive to differences in real income relative to idiosyncratic preferences, which implies commuting elasticities that are larger in absolute magnitude. We find that the second of these two sets of forces dominates, which is in line with the empirical findings in Kreindler and Miyauchi (2023) and Tsivanidis (2023).

In Columns (4)-(6) of Panel A, we report the results of estimating the commuting gravity equation (D.1) in logs using the linear fixed effects estimator. Again we estimate negative and statistically significant elasticities of bilateral commuting flows with respect to bilateral travel times. We

Table D.1: Commuting Gravity Equation by Occupation

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A	λ_{ni}^{L}	λ_{ni}^{M}	λ_{ni}^{H}	$\log \lambda_{ni}^{L}$	$\log \lambda_{ni}^{M}$	$\log \lambda_{ni}^{H}$
Travel time	-2.788^{***}	-2.303^{***}	-1.791^{***}	-1.377^{***}	-1.746^{***}	-0.824^{***}
	(0.037)	(0.078)	(0.087)	(0.033)	(0.030)	(0.035)
Occupation	Low	Mid	High	Low	Mid	High
Estimator	PPML	PPML	PPML	OLS	OLS	OLS
Workplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Residence FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	126,380	126,380	126,380	37,453	62,713	34,567
Pseudo R-squared	0.629	0.801	0.842	–	–	–
R-squared	–	–	–	0.595	0.758	0.751
Panel B	λ_{ni}^{L}	λ_{ni}^{M}	λ_{ni}^{H}	$\log \lambda_{ni}^{L}$	$\log \lambda_{ni}^{M}$	$\log \lambda_{ni}^{H}$
Travel time	-2.923^{***}	-2.411^{***}	-1.873^{***}	-1.423^{***}	-1.810^{***}	-0.863^{***}
	(0.041)	(0.097)	(0.104)	(0.035)	(0.032)	(0.038)
Control function	1.122^{***}	0.901^{***}	0.717^{***}	–	–	–
	(0.097)	(0.192)	(0.231)			
Occupation	Low	Mid	High	Low	Mid	High
Estimator	PPML	PPML	PPML	2SLS	2SLS	2SLS
Workplace FEs	Yes	Yes	Yes	Yes	Yes	Yes
Residence FEs	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic	–	–	–	21,723	22,241	16,484
Observations	126,380	126,380	126,380	37,453	62,713	34,567

Notes: Table reports the results of estimating the gravity equation (D.1) using data on bilateral commuting flows between Middle Super Output Areas (MSOAs) in the LCC area for low, mid and high-income occupations from the 2001 population census; Panel A reports results using PPML and OLS; Panel B reports results using a control function for PPML (Wooldridge 2014) and 2SLS to instrument for bilateral travel time with bilateral straight-line distance; the second-stage R-squared is not reported for the IV specifications, because it does not have a meaningful interpretation; all regressions include workplace and residence fixed effects (FEs) that vary by occupation; standard errors in parentheses are clustered by residence and workplace; *, ** and *** denote significance at the 10, 5 and 1 percent levels, respectively.

find that these estimated elasticities in the log linear specification are somewhat smaller in absolute value than in our baseline PPML specification. Additionally, these estimated elasticities in the log linear specification no longer decline monotonically with income across the three groups, although the estimated elasticity is smaller in absolute value for high-income workers than for low-income workers. Both sets of results highlight the relevance of allowing for granularity and zero bilateral commuting flows using the PPML estimator.

In Columns (1)-(3) of Panel B of Table D.1, we report the results of estimating our preferred specification using the PPML estimator and the control function approach of Wooldridge (2014) to instrument bilateral travel times with bilateral distance. In Columns (4)-(6) of Panel B, we report our robustness test using log linear specification and 2SLS. In both cases, we continue to find negative and statistically significant elasticities of bilateral commuting flows with respect to bilateral travel times. In the first-stage regression, we find that bilateral distance is a powerful predictor of bilateral travel times, with a first-stage F-statistic well above the conventional threshold of ten.

In the second-stage regression, we find estimated commuting elasticities that are marginally larger in absolute magnitude once we instrument (comparing Panels A and B). This marginal increase in the absolute magnitude of the coefficients when we instrument suggests that a greater incentive to invest in routes with more commuters for unobserved reasons in the error term may have been offset by other factors. In particular, the historical literature emphasizes the noncooperative behavior of the private-sector railways, and their attempts to carve out geographical territories of dominance through a proliferation of branch lines. This struggle for areas of geographic dominance could have led to overinvestment in routes that were less attractive in terms of their unobserved characteristics in the error term, thereby resulting in IV coefficients that are marginally larger in absolute magnitude. We use the PPML control function estimates from Columns (1)-(3) of Panel B as our baseline specification: $\phi^L = 2.92$, $\phi^M = 2.41$, and $\phi^H = 1.87$.

Finally, we separate the composite elasticity of commuting flows to travel times ($\phi^o = \epsilon^o \kappa$) into its two components. We allow the commuting decisions of high, middle and low-income workers to respond differentially to commuting costs (through variation in ϵ^o). But we assume that travel time affects commuting costs in the same way for all three groups of workers (common κ). Given these assumptions, we calibrate the preference dispersion parameter for middle-income workers as $\epsilon^M = 5.25$ based on the estimate using the construction of London's 19th-century railway network in Heblich et al. (2020). We then recover the implied preference dispersion parameters for low and high-income workers from our gravity equation estimates above, using our assumption of a common κ: $\epsilon^L = (\phi^L/\phi^M)\epsilon^M = 6.36$ and $\epsilon^H = (\phi^H/\phi^M)\epsilon^M = 4.07$.[2]

D3 Wages, Commuting and Employment (Step 3)

Given our estimated commuting parameters, we next solve for the model's predictions for wages (w^o_{it}), commuting flows (E^o_{nit}) and employment (E^o_{it}) for each occupation in the initial pre-war equilibrium, which we use as inputs in our counterfactuals below.

Pre-war Wages, Commuting and Employment We set the elasticity of substitution across occupations equal to the conventional value of $\sigma = 1.41$ from Katz and Murphy (1992). We calibrate the labor cost weights (γ^o) such that the aggregate shares of the three occupations in the total wage bill for the LCC area are consistent with their aggregate shares of residential rateable

[2] These values for the preference dispersion parameters lie within the range of existing empirical estimates from 2.18 to 8.3 in Ahlfeldt et al. (2015), Dingel and Tintelnot (2023), Severen (2023), and Kreindler and Miyauchi (2023).

values. Under our assumption of Cobb-Douglas preferences, residential rateable values are a constant multiple of residential income, which allows us to recover the shares of the occupations in total residential income from their shares in total residential rateable values.

Given the observed commercial rateable values ($\mathbb{V}_{it}^{E}$), residents (R_{it}^{o}) and travel times (τ_{nit}), we solve for pre-war wages by occupation (w_{it}^{o}) for each Output Area from the commuter market clearing condition in equation (19) in the paper. Given these solutions for pre-war wages by occupation (w_{it}^{o}), we compute pre-war conditional commuting probabilities by occupation ($\lambda_{nit|n}^{Ro}$) using equation (B.21) and our estimates of commuting costs ($(\kappa_{nit}^{o})^{-\epsilon^{o}} = \tau_{nit}^{-\phi^{o}}$):

$$\lambda_{nit|n}^{Ro} = \frac{\left(w_{it}^{o}\right)^{\epsilon^{o}} \tau_{nit}^{-\phi^{o}}}{\sum_{\ell \in N} \left(w_{\ell t}^{o}\right)^{\epsilon^{o}} \tau_{n\ell t}^{-\phi^{o}}}. \tag{D.2}$$

Finally, using these solutions for pre-war conditional commuting probabilities ($\lambda_{nit|n}^{Ro}$) from equation (D.2), together with observed residents (R_{nt}^{o}) and total city population ($\overline{E}_{t}^{o}$) for each occupation, we calculate pre-war unconditional commuting probabilities (λ_{nit}^{o}) by occupation:

$$\lambda_{nit}^{o} = \frac{\lambda_{nit|n}^{o} R_{nt}^{o}}{\bar{E}_{t}^{o}}, \tag{D.3}$$

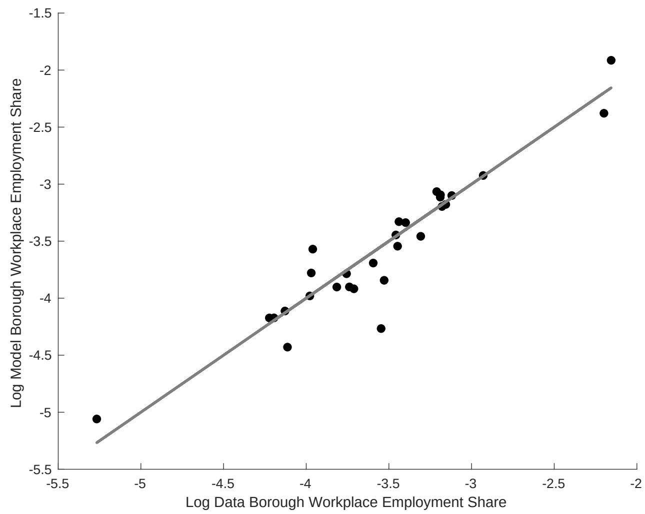
and employment by occupation:

$$E_{it}^{o} = \sum_{n \in \mathbb{N}} \lambda_{nit|n}^{Ro} R_{nt}^{o}. \tag{D.4}$$

We now report two specification checks on our model's predictions for pre-war employment and commuting patterns. In Figure D.1, we compare our model's predictions for the log pre-war employment share by workplace ($\lambda_{it}^{E} = E_{it}/\overline{E}_{t}$) for each LCC borough against the corresponding value of this variable in our pre-war bilateral commuting data for 1921. We aggregate across the three occupations and report results for boroughs, because our pre-war bilateral commuting data are not disaggregated by occupation and are only available for the relatively aggregated spatial units of the 29 LCC boroughs.[3] Our model's predictions are based on the commuter market clearing condition in equation (19) in the paper, using information on pre-war residents and commercial rateable values. Therefore, there is no necessary reason why these model predictions need exactly equal the observed values of employment by workplace in the data. Indeed, there are several reasons why these two variables could differ, including the fact that residents and commercial rateable values are measured during the 1930s, whereas our pre-war bilateral commuting data are for the earlier year of 1921. Nonetheless, we observe a strong and approximately log linear relationship between our model's predictions and the observed data, with a correlation coefficient of 0.94.

[3]The NSOL includes a sample survey that reports residence and workplace borough, as used in Seltzer and Wadsworth (2023), but these data are only available for low-income workers in a handful of Eastern boroughs.

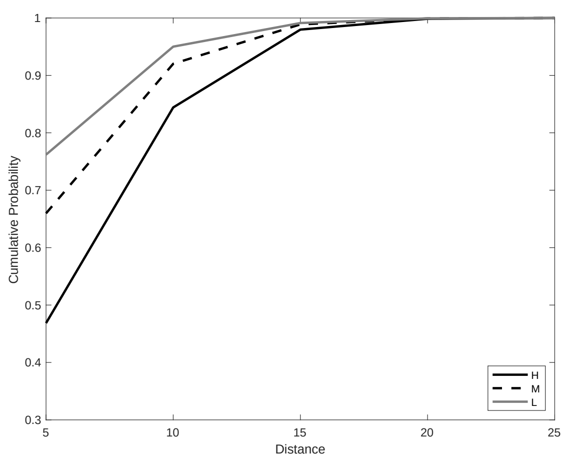
Figure D.1: Pre-war Shares of Boroughs in Employment by Workplace in the Model and Data



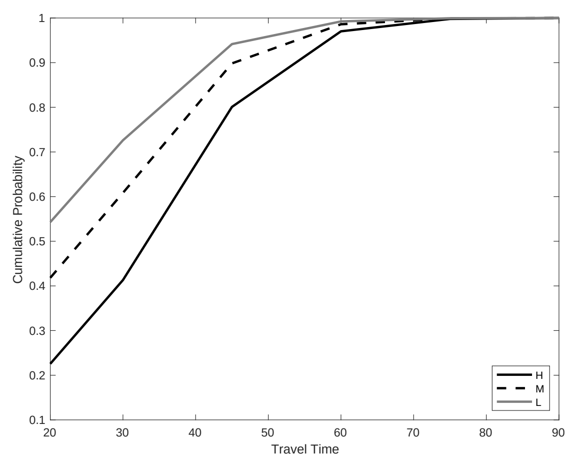
Note: Vertical axis shows our model's predictions for the log pre-war employment share by workplace aggregating across occupations ($\lambda_{it}^{E} = \sum_{o \in \mathbb{O}} E_{it}^{o} / \sum_{o \in \mathbb{O}} \overline{E}_{t}^{o}$) for each borough in the LCC area; horizontal axis shows the corresponding log employment share in our 1921 bilateral commuting data.

Figure D.2: Model Predictions for Pre-war Cumulative Commuting Probabilities by Distance and Travel Time

(a) Distance



(b) Travel Time

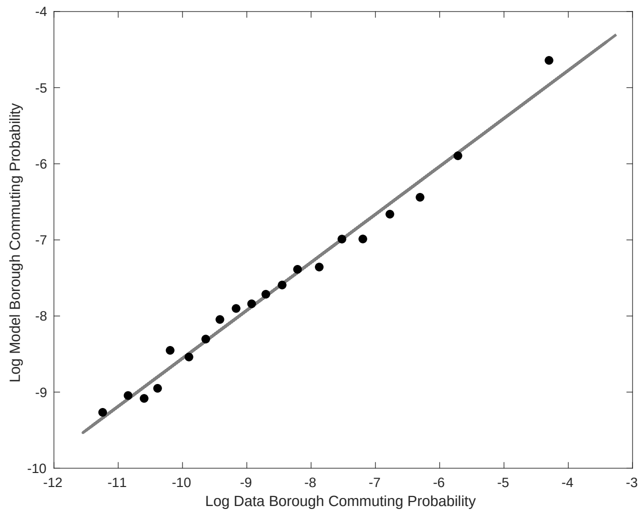


Note: Cumulative commuting probabilities in the model by distance (left panel) and travel time (right panel) for low, middle and high-income workers separately for the LCC area as a whole. Cumulative commuting probabilities computed from our model inversion using our pre-war data (Steps 1-3).

We next provide further evidence on our model's predictions for pre-war commuting patterns. In Figure D.2a, we display cumulative commuting probabilities in the model by distance, for low, middle and high-income workers in the LCC area. In Figure D.2b, we show analogous cumulative commuting probabilities in the model by travel time for these three groups of workers.

From these two figures, most workers commute for less than 10 kilometers and for less than 45 minutes. Higher-income workers typically commute over longer distances and for longer travel times than lower-income workers, which is consistent with them typically living in the outer more prosperous suburbs of the LCC area.

Figure D.3: Binscatter of Log Commuting Probabilities in the Model and Data Between Boroughs



Note: Vertical axis shows our model's prediction for the log pre-war bilateral commuting probability aggregating across occupations ($\lambda_{nit} = \sum_{o \in \mathbb{O}} E^o_{nit} / \sum_{o \in \mathbb{O}} \overline{E}^o_t$) for each pair of boroughs in the LCC area; horizontal axis shows the corresponding log bilateral commuting probability in our 1921 bilateral commuting data; solid circles correspond to ventiles of the distribution; solid gray line is the linear regression relationship.

Finally, we compare our model's commuting predictions to the available pre-war data on total bilateral commuting flows between the 29 boroughs of the LCC area. In a first step, we aggregate our model's predictions across the three worker groups to obtain total bilateral commuting flows between Output Areas. In a second step, we aggregate our model's predictions across pairs of Output Areas within each pair of boroughs. In Figure D.3, we show a binscatter of our model's predictions against log commuting probabilities between boroughs in our 1921 bilateral commuting data. Again, there are several reasons why our model's predictions and the observed data could differ from one another, including the fact that our model predictions use information on residents and commercial rateable values from the 1930s, whereas the pre-war commuting data are from 1921. Nonetheless, we find a strong and approximately log linear relationship between our model's predictions and the data, with a correlation coefficient of 0.87.

Taken together, both specification checks provide empirical support for our model's predictions for pre-war employment and commuting patterns.

Post-war Wages and Commuting For the post-war period, we observe employment (E^o_{it}) and residents (R^o_{it}) by occupation for each Output Area, but we do not observe wages (w^o_{it}) by occupation for each Output Area, and we only observe bilateral commuting flows (E^o_{nit}) by occupation at the more aggregated level of Middle Super Output Areas. Therefore, we solve for implied post-war wages by occupation and Output Area and bilateral commuting flows by occupation between pairs of Output Areas using a similar procedure as for the pre-war period above.

Again we use the model's commuter market clearing condition (B.22), which implies that employment by occupation for each workplace (E^o_{it}) equals the sum across residences of the number of workers in that occupation commuting to that workplace ($\lambda^o_{nit|n} R^o_{nt}$). Using our estimates of commuting costs ($(\kappa^o_{nit})^{-\epsilon^o} = \tau_{nit}^{-\phi^o}$), we can write this commuter market clearing condition as:

$$E^o_{it} = \sum_{n \in \mathbb{N}} \lambda^o_{nit|n} R^o_{nt} = \sum_{n \in \mathbb{N}} \frac{\left(w^o_{it}\right)^{\epsilon^o} \tau_{nit}^{-\phi^o}}{\sum_{\ell \in \mathbb{N}} \left(w^o_{\ell t}\right)^{\epsilon^o} \tau_{n\ell t}^{-\phi^o}} R^o_{nt}. \tag{D.5}$$

To ensure that employment by workplace (E^o_{it}) on the left-hand side and employment by residence (R^o_{nt}) on the right-hand side are both defined for workers living within the LCC area, we scale down employment by workplace (E^o_{it}) by the share of workers that commute from outside the LCC area for each Middle Super Output Area. After this adjustment, the sum of employment by workplace equals the sum of employment by residence for the LCC area as a whole. Given these adjusted values for employment (E^o_{it}), and observed residents (R^o_{it}) and travel times (τ_{nit}), this commuter market clearing condition (D.5) determines unique wages (w^o_{it}) by occupation for each Output Area (up to a choice of units in which to measure wages).

Using these solutions for post-war wages (w^o_{it}) by occupation from equation (D.5), we compute post-war conditional commuting probabilities ($\lambda^{Ro}_{nit|n}$) by occupation using equation (B.21) and our estimates of commuting costs ($(\kappa^o_{nit})^{-\epsilon^o} = \tau_{nit}^{-\phi^o}$):

$$\lambda^{Ro}_{nit|n} = \frac{\left(w^o_{it}\right)^{\epsilon^o} \tau_{nit}^{-\phi^o}}{\sum_{\ell \in N} \left(w^o_{\ell t}\right)^{\epsilon^o} \tau_{n\ell t}^{-\phi^o}}. \tag{D.6}$$

Finally, using these solutions for post-war conditional commuting probabilities ($\lambda^{Ro}_{nit|n}$) from equation (D.6), together with observed residents (R^o_{nt}) and total city population ($\overline{E}^o_t$) for each occupation, we calculate post-war unconditional commuting probabilities (λ^o_{nit}) by occupation:

$$\lambda^o_{nit} = \frac{\lambda^o_{nit|n} R^o_{nt}}{\overline{E}^o_t}, \tag{D.7}$$

where our model's predictions are necessarily equal to our observed data on post-war employment by occupation (E^o_{it}) from our solution of the commuter market clearing condition (D.5).

D4 Amenities (Step 4)

Given these solutions for wages (w^o_{nt}) by location and occupation, we next use the structure of the model to solve for residential amenities (B^o_{nt}) by location and occupation.

Re-writing the residential choice probabilities (B.12) using expected utility (B.8), we obtain the following closed-form expression for residential amenities for each worker group (B_{nt}^{o}) in terms of the observed shares of residents (λ_{n}^{Ro}), observed residential floor space prices (Q_n) and a measure of residents' commuting market access (RMA_{nt}^{o}):

$$\ln B_{nt}^{o} = \ln\left(\frac{U_t^o}{\delta^o}\right) + \frac{1}{\epsilon^o}\ln\left(\lambda_{nt}^{Ro}\right) + (1-\alpha^o)\ln Q_{nt} - \ln RMA_{nt}^{o}. \tag{D.8}$$

Residents commuting market access (RMA_{nt}^{o}) for each occupation is a travel time weighted average of wages in each workplace for that occupation:

$$RMA_{nt}^{o} = \left[\sum_{\ell\in\mathbb{N}} \left(w_{\ell t}^{o}\right)^{\epsilon^o} \tau_{n\ell t}^{-\phi^o}\right]^{\frac{1}{\epsilon^o}}, \tag{D.9}$$

where we have again used our estimates of commuting costs ($(\kappa_{nit}^{o})^{-\epsilon^o} = \tau_{nit}^{-\phi^o}$).

D5 Wartime Destruction and Neighborhood Effects (Step 5)

In our fifth and final step, we estimate the impact of wartime destruction on residential amenities and the strength of neighborhood effects.

D5.1 General Specification of Neighborhood Effects

We begin by considering our general specification of neighborhood effects, in which we estimate the direct and spillover effects of wartime destruction on residential amenities using equation (4) in the paper, without taking a stand on whether these spillover effects occur through the surrounding composition of either people or buildings.

We start with a Placebo specification, in which we regress pre-war amenities on subsequent wartime destruction, including fixed effects for 1 km hexagons. As reported in Table D.2 below, we find no evidence that pre-war amenities are correlated with future wartime destruction, which again provides further validation for our use of wartime destruction as an exogenous source of variation.

We next turn to our main causal regression, in which we regress post-war amenities on wartime destruction in both the own location and the 100-500 meter buffers, including fixed effects for 1 km hexagons. As reported in Table 4 in the paper, we find that the direct effects of wartime destruction are negative and statistically significant for high-income workers. In contrast, these direct effects are positive and weakly significant for low-income workers. This pattern of estimated coefficients is consistent with the idea that the construction of council housing in bombed locations reduces relative amenities in these locations for high-income workers.

Table D.2: Randomness of Wartime Destruction (Pre-war Amenities)

	(1) $\ln B^{H}_{n,\text{Pre}}$	(2) $\ln B^{H}_{n,\text{Pre}}$	(3) $\ln B^{M}_{n,\text{Pre}}$	(4) $\ln B^{M}_{n,\text{Pre}}$	(5) $\ln B^{L}_{n,\text{Pre}}$	(6) $\ln B^{L}_{n,\text{Pre}}$
Destruction in own area	-0.006	0.006	0.008	0.015	0.019	0.021
	(0.030)	(0.027)	(0.019)	(0.020)	(0.017)	(0.017)
Destruction in 100m buffer		-0.048		-0.030		-0.013
		(0.057)		(0.039)		(0.038)
Destruction in 200m buffer		0.038		-0.015		0.002
		(0.071)		(0.046)		(0.044)
Destruction in 300m buffer		-0.031		0.006		-0.028
		(0.079)		(0.058)		(0.053)
Destruction in 400m buffer		0.112		0.015		0.048
		(0.086)		(0.062)		(0.061)
Destruction in 500m buffer		-0.087		-0.067		-0.042
		(0.095)		(0.062)		(0.059)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
R Squared	0.46	0.46	0.32	0.32	0.36	0.36
Observations	8,717	8,714	8,718	8,715	8,719	8,716

Note: The unit of observation for all regressions is an Output Area as defined in the 2001 UK Census. The dependent variable in columns (1) and (2) is pre-war log amenities for high-income workers ($\ln B^{H}_{n,\text{Pre}}$), in columns (3) and (4) it is pre-war log amenities for middle-income workers ($\ln B^{M}_{n,\text{Pre}}$), and in columns (5) and (6) it is pre-war log amenities for low-income workers ($\ln B^{L}_{n,\text{Pre}}$). The explanatory variable is the fraction of the pre-war built-up area in each Output Area and five buffers of 100 meter width around each Output Area seriously damaged during the Second World War. All regressions include fixed effects for 1 km hexagons. Numbers of observations are less than 9,041 due to whether Output Areas and their buffers had pre-war built-up area and positive pre-war residents for each group of workers. Standard errors are clustered at the 1 km hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

We find that the spillover effects of wartime destruction are also negative and statistically significant for high-income workers, whereas these spillover effects are insignificant and small in absolute magnitude for low-income workers. These results are in line with the idea that the construction of council housing on bomb sites in neighboring locations reduces relative amenities in the own location for high-income workers. As in our earlier regressions for property values and socioeconomic composition, we find that these spillover effects are localized, with the spillover coefficient for high-income workers losing significance by the 300 meter buffer and falling by a factor of five between the 100 and 500 meter buffers.

Taken together, these empirical results are consistent with the mechanism in our model, in which wartime destruction changes relative amenities for the three groups of workers, which affects equilibrium patterns of spatial sorting, and hence post-war property prices and socioeconomic status. Our finding that bombing in neighboring locations affects amenities in the own location provides evidence of neighborhood effects, without taking a stand on the mechanism through which these neighborhood effects occur.

D5.2 Neighborhood Effects and Socioeconomic Composition

Motivated by our mechanisms findings in Section 4.4 in the paper, we next parameterize neighborhood effects as depending on the surrounding socioeconomic status of the population:

$$\ln B_{nt}^{o} = \eta_D^{o} D_{nt} + \eta_R^{o} \ln \mathbb{B}_{nt} + \eta_X^{o} X_{nt} + \varrho_{kt}^{o} + b_{nt}^{o}, \tag{D.10}$$

where $\mathbb{B}_{nt}$ is the surrounding socioeconomic status of the population, as defined below; ϱ_{kt}^{o} are our 1 km hexagon fixed effects; X_{nt} are controls; and b_{nt}^{o} is a stochastic error, which captures residential fundamentals (e.g., scenic views).

We allow for a direct effect of wartime destruction on amenities in bombed locations (η_D^{o}), because the destruction and reconstruction of buildings can change the amenities that residents derive from living in those buildings. But we assume that the impact of destruction in neighboring locations on residential amenities in the own location is fully summarized by the surrounding socioeconomic composition of the population (η_R^{o}).[4] The inclusion of the 1 km hexagon fixed effects ensures that the parameters are estimated from the exogenous variation in wartime destruction within these geographical grid cells and controls for other determinants of amenities that vary across these geographical grid cells.

In our general specification of amenities, and in our regressions for property values and socioeconomic composition, we find that the spillover effects from wartime destruction are close to zero and statistically insignificant by the 500-meter buffer. Therefore, we model neighborhood effects as depending on the distance-weighted average of socioeconomic status in the own location and the 100-500 meter buffers:

$$\mathbb{B}_{nt} = \sum_{\{i\,:\,\mathrm{dist}_{ni}<500\}} \frac{e^{-\delta \mathrm{dist}_{ni}}}{\sum_{\{k\,:\,\mathrm{dist}_{nk}<500\}} e^{-\delta \mathrm{dist}_{nk}}} \mathbb{S}_{it}, \tag{D.11}$$

where dist_{ni} is the distance from the outer boundary of each Output Area to the inner boundary of the buffer; and $\mathbb{S}_{it}$ is our index of socioeconomic status.

In our baseline specification in equation (D.11), we assume exponential distance decay, and calibrate the rate of decay (δ) such that the weight is close to zero by 500 meters (equal to 0.01). But we find a similar pattern of results across a range of assumptions for the rate of distance decay,

[4]Therefore, we assume either that there is no independent effect of neighboring buildings from neighboring people, or that the effect of neighboring buildings is fully summarized by the effect of neighboring people, because people and buildings are closely linked in spatial equilibrium.

including a simple step function, where either the first three or four buffers receive a weight of one and areas further away receive a weight of zero.

The main challenge in estimating neighborhood effects is that the surrounding socioeconomic composition of the population is endogenous, because workers sort endogenously across locations in response to differences in amenities, which induces a positive correlation between surrounding socioeconomic status ($\mathbb{B}_{nt}$) and the error term (b^o_{nt}). We use the exogenous variation in wartime destruction within 1 km hexagons to address this challenge. We instrument surrounding socioeconomic status ($\mathbb{B}_{nt}$) with the distance-weighted average of the share of the built-up area seriously damaged in the 100-500 meter buffers excluding the own location itself (D^{Neigh}_{nt}):

$$\ln \mathbb{B}_{nt} = \varkappa^o_D D_{nt} + \varkappa^o_N D^{\text{Neigh}}_{nt} + \varkappa^o_X X_{nt} + \varpi^o_{kt} + u^o_{nt}, \tag{D.12}$$

where the instrument (D^{Neigh}_{nt}) is defined analogously to neighborhood effects ($\mathbb{B}_{nt}$) in equation (D.11), but replacing our socioeconomic index ($\mathbb{S}_{it}$) with wartime destruction (D_{it}), including only the 100-500 meter buffers, and excluding the own location itself.

In Table 5 in the paper, we report the results of estimating the second-stage regression (D.10) for post-war amenities. In Table D.3 below, we report the corresponding results from estimating the first-stage regression (D.12). In Column (1), we instrument for post-war neighborhood effects using the distance-weighed average of overall destruction in the 100-500 meter buffers, excluding the own location. We find that overall wartime destruction in neighboring locations is a powerful instrument for surrounding socioeconomic composition, with a first-stage F-statistic of over ten. In Columns (2)-(4), we instrument for post-war neighborhood effects using the distance-weighted average of residential wartime destruction in the 100-500 meter buffers. Consistent with the mechanism in our model, in which wartime destruction of buildings changes the amenities from living in those buildings and hence affects socioeconomic composition, we find larger first-stage F-statistics using neighboring residential destruction as an instrument in Columns (2)-(4) than using neighboring overall destruction in Column (1).

In Table D.4, we report the results of estimating the reduced-form specification implied by the second and first-stage equations (D.10) and (D.12). Consistent with the first and second-stage results discussed above and in the paper, we find negative and statistically significant effects of both own destruction and neighboring destruction for high and middle-income workers, with a coefficient that is larger in absolute magnitude for high-income workers. In contrast, the effects of own destruction for low-income workers are positive and sometimes statistically significant, while the effects of neighboring destruction for low-income workers are negative and sometimes statistically significant, but smaller in absolute magnitude than for high and medium-income workers. Consistent with our mechanism based on the effects of wartime destruction on the amenities from living in buildings and the resulting changes in surrounding residential composi-

Table D.3: First-Stage Regression for Post-War Neighborhood Effects ($\ln \mathbb{B}_{n,\text{post-war}}$)

	(1)	(2)	(3)	(4)
(A) High-income	$\ln \mathbb{B}_{n,\text{post-war}}$	$\ln \mathbb{B}_{n,\text{post-war}}$	$\ln \mathbb{B}_{n,\text{post-war}}$	$\ln \mathbb{B}_{n,\text{post-war}}$
Destruction in own area	-0.038***	-0.038***	-0.037***	-0.004
	(0.007)	(0.006)	(0.006)	(0.004)
Neighboring destruction	-0.137***	-0.143***	-0.134***	-0.149***
	(0.040)	(0.033)	(0.031)	(0.028)
Pre-war neighbor effects			0.148***	0.147***
			(0.012)	(0.014)
Estimation	OLS	OLS	OLS	OLS
Observations	8771	8587	8587	8587
R-squared	0.721	0.733	0.766	0.834
First-stage F	11.56	19.19	18.72	28.08
(B) Medium-income	$\ln \mathbb{B}_{n,\text{post-war}}$	$\ln \mathbb{B}_{n,\text{post-war}}$	$\ln \mathbb{B}_{n,\text{post-war}}$	$\ln \mathbb{B}_{n,\text{post-war}}$
Destruction in own area	-0.038***	-0.038***	-0.037***	-0.004
	(0.007)	(0.006)	(0.006)	(0.004)
Neighboring destruction	-0.137***	-0.143***	-0.134***	-0.149***
	(0.040)	(0.033)	(0.031)	(0.028)
Pre-war neighbor effects			0.148***	0.147***
			(0.012)	(0.014)
Estimation	OLS	OLS	OLS	OLS
Observations	8786	8602	8602	8602
R-squared	0.721	0.733	0.766	0.835
First-stage F	11.60	19.15	18.69	28.00
	(1)	(2)	(3)	(4)
(C) Low-income	$\ln \mathbb{B}_{n,\text{post-war}}$	$\ln \mathbb{B}_{n,\text{post-war}}$	$\ln \mathbb{B}_{n,\text{post-war}}$	$\ln \mathbb{B}_{n,\text{post-war}}$
Destruction in own area	-0.038***	-0.039***	-0.038***	-0.004
	(0.007)	(0.006)	(0.006)	(0.004)
Neighboring destruction	-0.136***	-0.142***	-0.133***	-0.149***
	(0.040).	(0.033)	(0.031)	(0.028)
Pre-war neighbor effects			0.148***	0.146***
			(0.012)	(0.014)
Estimation	OLS	OLS	OLS	OLS
Observations	8767	8584	8584	8584
R-squared	0.721	0.733	0.765	0.834
First-stage F	11.42	18.84	18.43	27.81

Note: Estimates of the first-stage regression (equation (D.12)) corresponding to the estimates of the second-stage regression (equation (D.10)) reported in Table 5 in the paper; the dependent variable is post-war neighborhood effects ($\mathbb{B}_{n,\text{post-war}}$) computed as the distance-weighted average of the socioeconomic composition of the own location and the 100-500 meters buffers; $D_{n,\text{war}}$ is the fraction of the pre-war built-up area seriously damaged in each location; $D_{n,\text{war}}^{\text{neigh}}$ is the distance-weighted average of the fraction of the pre-war built-up area seriously damaged in the 100-500 meter buffers, excluding the own location itself; Column (1) uses wartime destruction of the overall built-up area; Columns (2)-(4) use wartime destruction of the residential built-up area; $\mathbb{B}_{n,\text{pre-war}}$ is pre-war neighborhood effects, computed in the same way as for post-war neighborhood effects; Column (4) excludes the own location from the measures of both post-war and pre-war neighborhood effects, such that these measures are based on the distance-weighted average of socioeconomic composition in the 100-500 meter buffers; all specifications include fixed effects for 1 kilometer hexagons; First-stage F is the first-stage F-statistic for the significance of the excluded exogenous variable ($D_{n,\text{war}}^{\text{neigh}}$); standard errors in parentheses clustered by 1 kilometer hexagons; *, ** and *** denote significance at the 10, 5 and 1 percent levels, respectively.

tion, we find stronger negative impacts of destruction on amenities for high and medium-income workers using residential destruction in Column (3) than using overall destruction in Columns (1)-(2).

Table D.4: Reduced-form Regression for Post-war Amenities, Own Destruction and Neighboring Destruction

	(1)	(2)	(3)
(A) High-income	$\ln B_n^H$	$\ln B_n^H$	$\ln B_n^H$
Destruction in own area	-0.102***	-0.080***	-0.088***
	(0.014)	(0.012)	(0.011)
Neighboring destruction		-0.133***	-0.127***
		(0.046)	(0.039)
Estimation	OLS	OLS	OLS
Observations	8779	8776	8589
R-squared	0.556	0.557	0.571
(B) Medium-income	$\ln B_n^M$	$\ln B_n^M$	$\ln B_n^M$
Destruction in own area	-0.046***	-0.035***	-0.044***
	(0.007)	(0.006)	(0.006)
Neighboring destruction		-0.066***	-0.076***
		(0.024)	(0.021)
Estimation	OLS	OLS	OLS
Observations	8794	8791	8604
R-squared	0.615	0.616	0.626
(C) Low-income	$\ln B_n^L$	$\ln B_n^L$	$\ln B_n^L$
Destruction in own area	0.014*	0.017**	0.008
	(0.008)	(0.007)	(0.006)
Neighboring destruction		-0.021	-0.046*
		(0.026)	(0.024)
Estimation	OLS	OLS	OLS
Observations	8775	8772	8586
R-squared	0.464	0.464	0.469

Note: Reduced-form regression results for post-war amenities for high, middle and low-income workers ($\ln B_n^H$, $\ln B_n^M$, $\ln B_n^L$); Destruction in own area is the fraction of the pre-war built-up area in each Output Area seriously damaged during the Second World War; Neighboring destruction is the distance-weighted average of the fraction of the pre-war built-up area in the 100-500 meter buffers seriously damaged during the Second World War; Columns (1)-(2) use wartime destruction of the overall built-up area; Column (3) uses wartime destruction of the residential built-up area; all specifications include fixed effects for 1 kilometer hexagons; standard errors in parentheses clustered by 1 kilometer hexagons; *, ** and *** denote significance at the 10, 5 and 1 percent levels, respectively.

As a Placebo specification check, Table D.5 re-estimates this reduced-form specification for pre-war amenities instead of post-war amenities. We find no relationship between pre-war amenities and the subsequent values of either own destruction or neighboring destruction. This pattern of results holds regardless of whether we use overall destruction or residential destruction. These findings provide further evidence that wartime destruction provides an exogenous source of variation within geographical grid cells. They are also consistent with the primitive bomb-aiming technology of the time, and the fact that much of the bombing occurred at night under conditions of a wartime blackout, which precluded the precise targeting of locations.

Table D.5: **Placebo** Reduced-form Regression for **Pre-war** Amenities, Own Destruction and Neighboring Destruction

	(1)	(2)	(3)
(A) High-income	$\ln B_n^H$	$\ln B_n^H$	$\ln B_n^H$
Destruction in own area	-0.006	-0.000	0.008
	(0.030)	(0.026)	(0.026)
Neighboring destruction		-0.037	-0.075
		(0.097)	(0.083)
Estimation	OLS	OLS	OLS
Observations	8717	8714	8705
R-squared	0.463	0.463	0.462
(B) Middle-income	$\ln B_n^M$	$\ln B_n^M$	$\ln B_n^M$
Destruction in own area	0.008	0.015	0.003
	(0.019)	(0.018)	(0.018)
Neighboring destruction		-0.044	-0.047
		(0.065)	(0.047)
Estimation	OLS	OLS	OLS
Observations	8718	8715	8706
R-squared	0.318	0.317	0.317
(C) Low-income	$\ln B_n^M$	$\ln B_n^M$	$\ln B_n^M$
Destruction in own area	0.019	0.021	0.006
	(0.017)	(0.016)	(0.017)
Neighboring destruction		-0.018	-0.004
		(0.061)	(0.048)
Estimation	OLS	OLS	OLS
Observations	8719	8716	8707
R-squared	0.365	0.365	0.365

Note: Placebo reduced-form regression results for pre-war amenities for high, middle and low-income workers ($\ln B_n^H$, $\ln B_n^M$, $\ln B_n^L$); Destruction in own area is the fraction of the pre-war built-up area in each Output Area seriously damaged during the Second World War; Neighboring destruction is the distance-weighted average of the fraction of the pre-war built-up area in the 100-500 meter buffers seriously damaged during the Second World War; Columns (1)-(2) use wartime destruction of the overall built-up area; Column (3) uses wartime destruction of the residential built-up area; all specifications include fixed effects for 1 kilometer hexagons; standard errors in parentheses clustered by 1 kilometer hexagons; *, ** and *** denote significance at the 10, 5 and 1 percent levels, respectively.

E Counterfactuals

In this section of the online appendix, we report further details on our counterfactuals to assess the general equilibrium implications of neighborhood effects, as discussed in Section 7 of the paper. We first examine the role of neighborhood effects in determining the impact of wartime destruction across locations. We next examine the role of these neighborhood effects in shaping observed differences in socioeconomic outcomes across locations even in counterfactual scenarios without wartime destruction. We report results for both our general specification, in which we do not take a stand on the mechanism through which neighborhood effects occur, and for our specific parametrization, in which we assume that neighborhood effects arise from preferences over the socioeconomic composition of the population.

We report the results of these counterfactuals for both closed and open-city specifications. In the closed-city specification, we hold the total population of each occupation in the LCC area constant, such that wartime destruction affects the expected utility of workers in each occupation. In the open-city specification, we hold the reservation level of utility for workers in each occupation constant, which implies that wartime destruction affects the total population of each occupation, but leaves expected utility for each occupation unchanged (after integrating across the distribution of idiosyncratic preferences). In both specifications, wartime destruction has distributional consequences for the amenity-adjusted real income of workers (without taking into account idiosyncratic preferences) and the real income of landlords across locations.

In our baseline specification, we report results for the case of exogenous productivity and perfectly inelastic supplies of commercial and residential floor space. In robustness specifications, we report results allowing for agglomeration forces (such that productivity responds to changes in employment) and an imperfectly elastic supply of residential floor space (such that the supply of residential floor space responds to changes in its price). Throughout the remainder of this section, we suppress the implicit dependence on time to reduce notational clutter.

E1 Counterfactual Equilibrium

We now discuss the system of equations that we use to solve for a counterfactual equilibrium. We follow an exact-hat algebra approach from the international trade literature, in which we use the values of the model's endogenous variables in the initial pre-war equilibrium to control for initial differences in location characteristics. We use the model's predictions for pre-war bilateral commuting flows based on our PPML gravity equation estimation, which allows for zeros and granularity for small spatial units.

We re-write the counterfactual equilibrium conditions in terms of the observed values of the endogenous variables in the initial equilibrium and the relative changes in the endogenous variables between the initial and counterfactual equilibria. We denote the value of a variable in the counterfactual equilibrium by a prime (x'_n), the value of variable in the initial equilibrium without a prime (x_n), and the relative change in a variable by a hat ($\widehat{x}_n = x'_n/x_n$).

E1.1 Baseline Closed-City Specification

We begin with our baseline closed-city specification, in which the measures of total residents from each occupation ($\overline{E}_t^L, \overline{E}_t^M, \overline{E}_t^H$) are exogenous, and the expected utilities of workers from each occupation (U_t^L, U_t^M, U_t^H) are endogenous.

System of Equations for Counterfactual Equilibrium Given an exogenous change in residential amenities ($\widehat{B}_n^o$), and initial guesses for changes in wages ($\widehat{w}_n^{o0}$) and the price of residential floor space ($\widehat{Q}_n^0$), our baseline closed-city specification solves the following system of equations:

$$\widehat{\lambda}_n^{Ro}\lambda_n^{Ro} = \frac{\sum_{\ell=1}^{N} \lambda_{n\ell}^o \left(\widehat{B}_n^o \widehat{w}_\ell^{o0}\right)^{\epsilon^o} \left(\widehat{Q}_n^0\right)^{-\epsilon^o(1-\alpha^o)}}{\sum_{k=1}^{N}\sum_{\ell=1}^{N} \lambda_{k\ell}^o \left(\widehat{B}_k^o \widehat{w}_\ell^{o0}\right)^{\epsilon^o} \left(\widehat{Q}_k^0\right)^{-\epsilon^o(1-\alpha^o)}}, \tag{E.1}$$

$$\widehat{\lambda}_n^{Eo}\lambda_n^{Eo} = \frac{\sum_{k=1}^{N} \lambda_{kn}^o \left(\widehat{B}_k^o \widehat{w}_n^{o0}\right)^{\epsilon^o} \left(\widehat{Q}_k^0\right)^{-\epsilon^o(1-\alpha^o)}}{\sum_{k=1}^{N}\sum_{\ell=1}^{N} \lambda_{k\ell}^o \left(\widehat{B}_k^o \widehat{w}_\ell^{o0}\right)^{\epsilon^o} \left(\widehat{Q}_k^0\right)^{-\epsilon^o(1-\alpha^o)}}, \tag{E.2}$$

$$\widehat{\lambda}_{ni|n}^{Ro}\lambda_{ni|n}^{Ro} = \frac{\lambda_{ni|n}^{Ro}\left(\widehat{w}_n^{o0}\right)^{\epsilon^o}}{\sum_{\ell\in\mathbb{B}} \lambda_{n\ell|n}^{Ro}\left(\widehat{w}_\ell^{o0}\right)^{\epsilon^o}}, \tag{E.3}$$

$$\widehat{R}_n^o = \widehat{\lambda}_n^{Ro}\lambda_n^{Ro}\overline{E}^o, \tag{E.4}$$

$$\widehat{E}_n^o = \widehat{\lambda}_n^{Eo}\lambda_n^{Eo}\overline{E}^o, \tag{E.5}$$

$$\widehat{q}_n = \left[s_n^L\left(\widehat{w}_n^L\right)^{1-\sigma} + s_n^M\left(\widehat{w}_n^M\right)^{1-\sigma} + s_n^H\left(\widehat{w}_n^H\right)^{1-\sigma}\right]^{-\frac{\beta}{(1-\beta)(1-\sigma)}}, \tag{E.6}$$

$$\widehat{v}_n^o v_n^o = \sum_{i\in\mathbb{N}} \widehat{\lambda}_{ni|n}\lambda_{ni|n}\widehat{w}_i^{o0} w_i^{o0}, \tag{E.7}$$

where in this baseline closed-city specification we hold constant commuting costs ($\widehat{\kappa}_{ni} = 1$), productivity ($\widehat{A}_n = 1$), the supply of residential floor space ($\widehat{H}_n^R = 1$), the supply of commercial floor space ($\widehat{H}_n^E = 1$), and total city population for each occupation ($\widehat{\overline{E}}^o = 1$).

From this system of equations, we obtain the implied changes in labor demand ($\widehat{E}_n^{oD}$) and labor supply ($\widehat{E}_n^{oS}$) for each occupation in each location:

$$\widehat{E}_n^{oD} = \frac{\widehat{q}_n}{\widehat{w}_n^{o0}}, \tag{E.8}$$

$$\widehat{E}_n^{oS} = \widehat{\lambda}_n^{Eo}, \tag{E.9}$$

and the implied changes in income from residential floor space (Ω_n^S) and expenditure (Ω_n^S) on residential floor space:

$$\widehat{\Omega}_n^S = \widehat{Q}_n^0, \tag{E.10}$$

$$\widehat{\Omega}_n^D = \sum_{o\in\{L,M,H\}} \frac{(1-\alpha^o)v_n^o R_n^o}{\sum_{m\in\{L,M,H\}}(1-\alpha^m)v_n^m H_n^m} \widehat{v}_n^o \widehat{R}_n^o, \tag{E.11}$$

where again this baseline closed-city specification holds constant the supply of residential floor space ($\widehat{H}_n^R = 1$) and the supply of commercial floor space ($\widehat{H}_n^E = 1$).

We update our initial guesses for changes in wages ($\widehat{w}_n^{o0}$) and the price of residential floor space ($\widehat{Q}_n^0$) until the changes in the demand and supply of labor are equal to one another for each occupation in equations (E.8) and (E.9), the changes in the demand and supply for residential and commercial floor space are equal to one another in equations (E.10) and (E.11), and all the other equilibrium conditions of the model are satisfied. Under the conditions for the existence of an unique equilibrium in Proposition 1 in the paper, we obtain unique counterfactual predictions for the impact of the exogenous change in residential amenities on the spatial distribution of economic activity.

In the presence of sufficiently strong neighborhood effects and agglomeration forces, there can be multiple equilibria in the model. When we solve for counterfactuals, we use the equilibrium selection rule of solving for a counterfactual equilibrium starting with initial values from the observed equilibrium in the data.

Expected Utility Given the counterfactual changes in wages for each occupation ($\widehat{w}_n^o$) and residential floor prices ($\widehat{Q}_n$) from solving the system of general equilibrium conditions above, we can compute the change in worker expected utility:

$$\widehat{U}_t^o = \left[\sum_{k\in\mathbb{N}}\sum_{\ell\in\mathbb{N}} \lambda_{nit}^o \left(\widehat{B}_{kt}^o \widehat{w}_{\ell t}^o\right)^{\epsilon^o} \left(\widehat{Q}_{kt}^{(1-\alpha^o)}\right)^{-\epsilon^o}\right]^{\frac{1}{\epsilon^o}}, \tag{E.12}$$

which takes the same value across all residence-workplace pairs for a given occupation, but differs across occupations.

Distributional Consequences Although expected utility is equalized throughout the city for workers from a given occupation, real income adjusted for amenities differs across residence-workplace pairs for that occupation. The reason is that each residence-workplace pair faces an upward-sloping supply function for commuters, such that it has to offer a higher amenity-adjusted real income ($\left(B_n^o w_i^o\right)/(\kappa_{ni}^o Q_n^{(1-\alpha^o)})$) in order to attract additional commuters, as apparent from the commuting probabilities (B.11). Therefore, changes in residential amenities have distributional consequences for workers' amenity-adjusted real income across residence-workplace pairs and occupations, with the proportional change in amenity-adjusted real income given by:

$$\widehat{B}_n^o \widehat{w}_i^o / \widehat{Q}_n^{(1-\alpha^o)}.$$

Intuitively, locations that experience large reductions in residential amenities from wartime destruction become relatively less attractive to workers in each occupation, which leads to a decline in residents. If higher-income workers are more responsive to changes in residential amenities, this decline in residents will be greater for higher-income workers. The mechanism that restores equilibrium is a fall in the price of residential floor space, but this price fall does not fully offset the decline in residential amenities, because of the upward-sloping supply function for commuters. Therefore, locations with greater reductions in residential amenities from wartime destruction experience a decline in the total number of residents, a fall in the price of residential floor space, and a shift in the composition of residents towards lower-income groups.

These changes in the price of residential floor space imply that wartime destruction also has distributional consequences for landlords across locations. In our baseline specification with an exogenous supply of residential floor space, the proportional change in landlord income is equal to the proportional change in the price of residential floor space:

$$\widehat{Q}_n.$$

In response to a reallocation of residents away from locations with greater wartime destruction that is larger for higher-income residents, other locations can experience increases in the total number of residents, the share of higher-income residents, and the price of residential floor space.

E2 Wartime Destruction

We undertake counterfactuals for wartime destruction starting at the observed pre-war equilibrium in the data. Whereas in reality many things can change between the pre and post-war periods, these counterfactuals evaluate the impact of wartime destruction holding all else constant. Although our reduced-form regressions control for many changes between these two periods by estimating separate fixed effects for 1 km hexagons for each time period, the inclusion of these fixed effects implies that these reduced-form regressions cannot capture general equilibrium effects (which are absorbed into the fixed effects). In contrast, our counterfactuals use the structure of our model to evaluate these general equilibrium effects. Although starting at the observed pre-war equilibrium is the natural choice to evaluate the general equilibrium impact of wartime destruction, we find a similar pattern of results if we instead start at the observed post-war equilibrium and remove the effects of wartime destruction.

E2.1 General Specification of Neighborhood Effects

We first undertake counterfactuals for wartime destruction using our general specification for neighborhood effects. We estimate the direct and spillover effects of wartime destruction on amenities using equation (4) in the paper, without taking a stand on whether these spillover effects occur through the surrounding composition of either people or buildings. We use the estimated direct and spillover coefficients for high, middle and low-income workers from Columns (2), (4) and (6) of Table 4 in the paper. We multiply these estimated coefficients for each occupation (η_D^o) by the share of the pre-war built-up area seriously damaged during the Second World War for each Output Area (D_n), and compute the implied exogenous changes in amenities ($\widehat{B}_n^o$):

$$\widehat{B}_n^o = e^{\eta_D^o D_n} \prod_{g=1}^{G} e^{\eta_{Bg}^o D_{ng}}, \tag{E.13}$$

where D_n is destruction in the own area for location n; η_D^o is the estimated coefficient on own destruction for occupation o; D_{ng} is destruction in buffer g for location n; and η_{Bg}^o is the estimated coefficient for occupation o on destruction in buffer g.

Given these exogenous changes in amenities ($\widehat{B}_n^o$), we solve for a counterfactual equilibrium using the system of general equilibrium conditions in Subsection E1.1 of this Online Appendix. We undertake separate counterfactuals for the direct effects of wartime destruction (setting the spillover coefficients equal to zero ($\eta_{Bg}^o = 0$ for all $g \in \{1, \ldots, G\}$)) and for the full effects of wartime destruction (allowing for both direct and spillover effects). In Section 7.1 of the paper, we report the results of these counterfactuals for the full and direct effects of wartime destruction.

E2.2 Neighborhood Effects and Socioeconomic Composition

We next undertake counterfactuals for wartime destruction using our parameterization of neighborhood effects in terms of preferences over the surrounding socioeconomic composition of the population in equation (22) in the paper. We use the estimated coefficients on wartime destruction (η_D^o) and neighborhood effects (η_R^o) from Column (5) of Table 5 in the paper. In these counterfactuals, the change in residential amenities ($\widehat{B}_n^o$) depends on the estimated coefficients on wartime destruction (η_D^o), the exogenous variation in wartime destruction (D_n), the estimated parameters for neighborhood effects (η_R^o), and the endogenous change in neighborhood effects ($\widehat{\mathbb{B}}_n$):

$$\widehat{B}_n^o = e^{\eta_D^o D_n} \widehat{\mathbb{B}}_n^{\eta_R^o}, \tag{E.14}$$

where $\mathbb{B}_n$ is the distance-weighted average of our socioeconomic index ($\mathbb{S}_n$) in the own location and the 100-500 meter buffers (as defined in equation (23) in the paper).

In these counterfactuals using our parameterization of neighborhood effects in terms of the socioeconomic composition of the population, we allow for the endogenous feedback of residential amenities to the endogenous changes in patterns of spatial sorting induced by wartime destruction. We augment the system of equations for a counterfactual equilibrium (E.1)-(E.11) with this change in residential amenities ($\widehat{B}_n^o$) from equation (E.14). In this augmented system, counterfactual changes in patterns of spatial sorting feed back to influence endogenous amenities through neighborhood effects.

We solve this augmented system of equations for a counterfactual equilibrium using the same algorithm as discussed in Subsection E1.1 above. We start with initial guesses for changes in wages ($\widehat{w}_n^{o0}$) and the price of residential floor space ($\widehat{Q}_n^0$). We update these initial guesses until the changes in the demand and supply of labor are equal to one another for each group of workers in equations (E.8) and (E.9), the changes in the demand and supply for residential and commercial floor space are equal to one another in equations (E.10) and (E.11), and all the other equilibrium conditions of the model are satisfied.

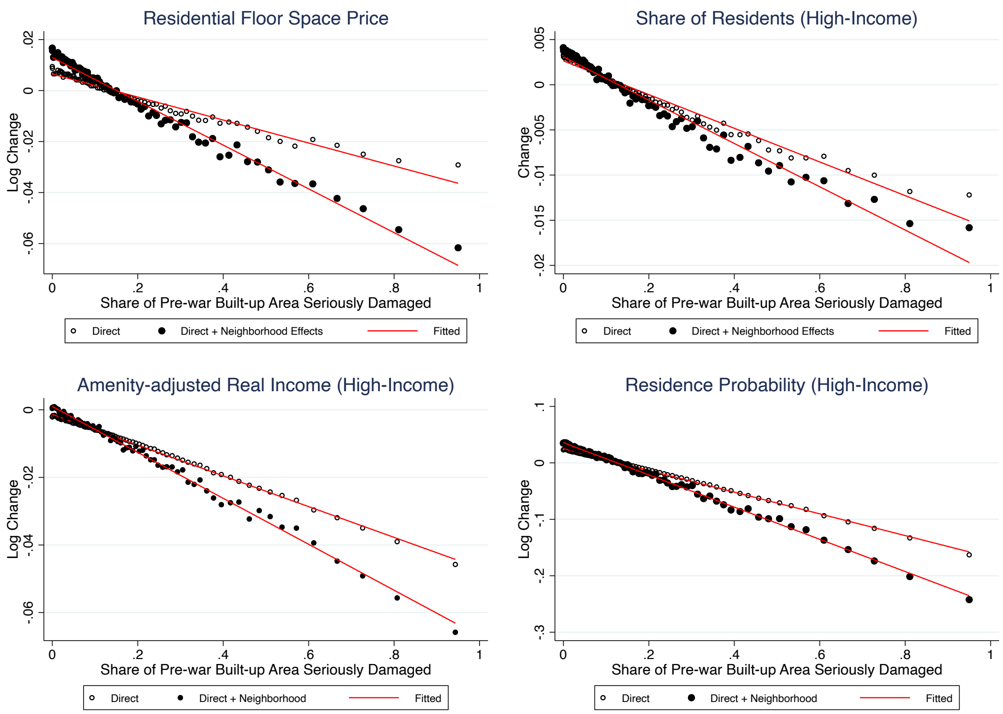
We compare the results of our counterfactuals using this parameterization of neighborhood effects (using the estimated values of η_D^o and η_R^o from Column (5) of Table 5 in the paper) to the results of counterfactuals allowing only for direct effects of wartime destruction (using the estimated value of η_D^o from Column (5) of Table 5 in the paper but setting $\eta_R^o = 0$).

In Figure E.1, we display the results of these counterfactuals for the full and direct effects of wartime destruction. This figure corresponds to Figure 4 in the paper, but uses our parameterization of neighborhood effects in terms of preferences over surrounding socioeconomic composition (Figure E.1), instead of our general specification of neighborhood effects (Figure 4 in the paper). Each panel shows a binscatter across Output Areas of the counterfactual change in a variable against the share of the pre-war built-up area seriously damaged during the Second World War. The circles correspond to percentiles of the distribution of wartime destruction and the red line represents the linear fit.

We find a similar pattern of results using this parameterization of neighborhood effects as using our general specification in the paper. As shown in Panel A, we find substantial distributional consequences of wartime destruction for landlord income. The change in the price of residential floor space ranges from a rise of 2 percentage points to a decline of over 6 percentage points. As higher-income workers move way from bombed locations, this bids up the price of residential floor space in unbombed locations. We find that the full effects of wartime destruction (solid circles) are notably larger than the direct effects (hollow circles), highlighting the quantitative relevance of neighborhood effects.

Panel B illustrates this impact of wartime destruction on equilibrium patterns of spatial sorting. The share of high-income workers declines by over 1.5 percentage points in locations that were completely destroyed and rises by around 0.5 percentage points in those with no destruction. Again the full effects of wartime destruction (solid circles) are notably larger than the direct effects (hollow circles), highlighting the role of the surrounding neighborhood in shaping resi-

Figure E.1: Counterfactuals for Wartime Destruction (Parameterizing Neighborhood Effects in Terms of Surrounding Socioeconomic Composition)



Notes: Counterfactuals using our parameterization of neighborhood effects in terms of preferences over surrounding socioeconomic composition and the estimated coefficients from Column (5) of Table 5 in the paper; circles show values by percentile of the share of the pre-war built-up area seriously damaged during the Second World War; around 40 percent of locations experience zero destruction, such that the circle for zero destruction captures the first 40 percentiles; hollow circles show counterfactuals allowing only for the direct effects of wartime destruction alone (using the estimated values of η_D^o but setting η_R^o equal to zero); solid circles show counterfactuals incorporating both direct effects of wartime destruction and neighborhood effects (using the estimated values of η_D^o and η_R^o); red lines show the linear fit; Panel A shows the log change in the price of residential floor space; Panel B shows the change (not in logs) in the share of a location's residents who are high-income ($R_n^H / \left(R_n^L + R_n^M + R_n^H\right)$); Panel C shows the log change in amenity-adjusted real income for high-income residents ($\log(B_n^H w_i^H Q_n^{-(1-\alpha^H)})$); Panel D shows the log change in the share of all high-income residents in the LCC area who live in a location ($\lambda_n^H = R_n^H / \sum_{i \in \mathbb{N}} R_i^H$).

dential choices.

As shown in Panel C, the distributional consequences of wartime destruction for the amenity-adjusted real income of workers are of a similar size to those for landlord income. The change in amenity-adjusted real income of high-income workers ranges from a decline of over 6 percentage points in locations that were completely destroyed to a small rise in locations that experienced no destruction. These patterns reflect both the direct effects of wartime destruction on amenities (more negative for high-income workers) and general equilibrium effects (as some high-income workers move out of bombed locations this reduces the price of residential floor space for those remaining). Again neighborhood effects magnify the impact of wartime destruction, with the full effects (solid circles) markedly larger than the direct effects (hollow circles).

Whereas Panel B displays the share of high-income workers in a given location in the total residents of that location ($R_n^H / \left(R_n^L + R_n^M + R_n^H\right)$), Panel D shows the log change in the share of high-income workers in a given location in the total number of high-income workers in the LCC area (the residential choice probability, $\lambda_n^H = R_n^H / \sum_{i \in \mathbb{N}} R_i^H$). The parts of London that experienced heavier wartime destruction and a decline in the share of high-income workers (often in the East) had lower initial shares of high-income workers, which results in a large percentage decline of over 20 percent in the share of high-income workers in completely-destroyed locations. Therefore, we again find that wartime destruction leads to a substantial change in the distribution of high-income workers across locations within the LCC area.

E3 Model Specification Check

In this subsection of the Online Appendix, we show that the counterfactual predictions of our model are successful in replicating the results of our reduced-form regressions for property values in Section 4 of the paper. We re-estimate our regressions for the direct and spillover effects of wartime destruction in equation (4) in the paper using our model's counterfactual predictions for post-war property prices instead of the observed data.

In Table E.1, we report the results for our general specification of neighborhood effects from Section 7.1 of the paper. Column (1) reproduces our estimates for the direct and spillover effects of wartime destruction using observed post-war property prices from Column (4) of Table 3 in the paper. Column (2) estimates the same specification using counterfactual predictions for the post-war price of residential floor space that assume no neighborhood effects for amenities (setting the spillover effects in Columns (2), (4) and (6) of Table 4 in the paper equal to zero). Column (3) estimates the same specification using counterfactual predictions for the post-war price of residential floor space that allow for neighborhood effects for amenities (using the estimated direct and spillover effects from Columns (2), (4) and (6) of Table 4 in the paper).

As shown in Column (1), we find direct and spillover effects of wartime destruction using observed post-war property prices. As shown in Column (2), when we assume no neighborhood effects for residential amenities, we find negative and statistically significant direct effects of wartime destruction using counterfactual post-war property prices. But we find no evidence of negative spillover effects: the estimated spillover coefficients are positive, small in magnitude and typically statistically insignificant. In contrast, as shown in Column (3), when we allow for neighborhood effects for residential amenities, we find negative and statistically significant direct and spillover effects of wartime destruction using counterfactual post-war property prices. The estimated spillover coefficients are comparable in magnitude to those estimated using the observed data. But since wartime destruction is the only shock in the model's counterfactuals, whereas there are many other sources of shocks in the observed data, the estimated coefficients are more precisely estimated and the regression R-squared is higher using our model's counterfactual predictions than using the observed data.

Therefore, we find that incorporating neighborhood effects for residential amenities is central

Table E.1: The Spillover Effect of Wartime Destruction (Model Versus Data, General Specification of Neighborhood Effects)

	(1) Log of Post-War Property Value (Data)	(2) Log of Post-War Property Value (Model Only Direct)	(3) Log of Post-War Property Value (Model Direct + Spillover)
Destruction in own area	-0.157***	-0.095***	-0.092***
	(0.022)	(0.004)	(0.004)
Destruction in 100m buffer	-0.080*	0.003	-0.058***
	(0.045)	(0.003)	(0.005)
Destruction in 200m buffer	-0.130**	0.001	-0.106***
	(0.059)	(0.003)	(0.005)
Destruction in 300m buffer	-0.094	0.008***	-0.072***
	(0.064)	(0.003)	(0.005)
Destruction in 400m buffer	-0.085	0.002	-0.022***
	(0.076)	(0.003)	(0.005)
Destruction in 500m buffer	-0.011	0.004	-0.046***
	(0.077)	(0.004)	(0.007)
Observations	8109	8109	8109
R-squared	0.659	0.749	0.891

Notes: The unit of observation for all regressions is an Output Area as defined in the 2001 UK Census. The dependent variable in column (1) is the logarithm of the average residential property value for output areas with at least 25 transactions over the period 1995 to 2020 and controlling for a set of observed building characteristics. The dependent variable in column (2) is the model's counterfactual prediction for the post-war price of residential floor space assuming only direct effects of wartime destruction on residential amenities (setting the spillover effects in Columns (2), (4) and (6) of Table 4 equal to zero). The dependent variable in column (3) is the model's counterfactual prediction for the post-war price of residential floor space allowing for direct and spillover effects of wartime destruction on residential amenities (using the estimated coefficients from Columns (2), (4) and (6) of Table 4). The explanatory variables are the fraction of the pre-war built-up area seriously damaged in each Output Area and five buffers of 100 meter width around each Output Area during the Second World War. All regressions include fixed effects for 1 km hexagons. Numbers of observations are less than 9041 due to the availability of modern housing transaction data and whether Output Areas and their buffers had pre-war built-up area. Standard errors are clustered at the 1 km hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

to matching the empirical finding in the data of negative spillover effects of wartime destruction for residential property prices

E4 Neighborhood Effects

We next undertake counterfactuals to evaluate the importance of neighborhood effects in shaping observed differences in socioeconomic outcomes across locations, even in counterfactual scenarios without wartime destruction. We use our parameterization of neighborhood effects in terms of the surrounding socioeconomic composition of the population. We start from the observed pre-war equilibrium in the data and set the preference parameter for socioeconomic composition in the counterfactual equilibrium equal to zero, which implies that residential fundamentals become the sole determinant of residential amenities in the counterfactual equilibrium:

$$\widehat{B}_n^o = \frac{B_n^{o\prime}}{B_n^o} = \frac{b_n^o \left(\mathbb{B}_{n,\text{pre-war}}\right)^0}{b_n^o \left(\mathbb{B}_{n,\text{pre-war}}\right)^{\eta_R^o}} = \frac{1}{\left(\mathbb{B}_{n,\text{pre-war}}\right)^{\eta_R^o}}, \tag{E.15}$$

where $\mathbb{B}_n$ is the distance-weighted average of our socioeconomic index ($\mathbb{S}_n$) in the own location and the 100-500 meter buffers (from equation (23)) and we set η_R^o in the pre-war equilibrium in the denominator equal to our estimate from Column (5) of Table 5. Since $\mathbb{S}_n$ takes values from zero to one, $\mathbb{B}_{n,\text{pre-war}}$ also lies within this interval;

This counterfactual is conceptually distinct from that in the previous subsection, in the sense that we now assess the importance of neighborhood effects for cross-sectional patterns of spatial sorting, which is a question that can be asked completely separately from wartime destruction. However, the reason that we can address this separate question is that we have estimated the model's structural parameters for neighborhood effects (η_R^o) using the exogenous variation in wartime destruction. We use the estimated coefficients from Column (5) of Table 5 in the paper, which identifies neighborhood effects using the variation in residential destruction in neighboring locations, after controlling separately for pre-war socioeconomic status. From equation (E.15), this counterfactual removes neighborhood effects ($\mathbb{B}_n$) evaluated at the observed socioeconomic composition in the initial pre-war equilibrium ($\mathbb{B}_{n,\text{pre-war}}$), which implies that the change in residential amenities ($\widehat{B}_n^o$) is exogenously determined by the pre-war data.

Substituting this exogenous change in residential amenities from equation (E.15) into the system of general equilibrium conditions (E.1)-(E.11), we solve for a counterfactual equilibrium using the same algorithm as discussed in Subsection E1.1 above. We start with initial guesses for changes in wages ($\widehat{w}_n^{o0}$) and the price of residential floor space ($\widehat{Q}_n^0$). We update these initial guesses until the changes in the demand and supply of labor are equal to one another for each group of workers in equations (E.8) and (E.9), the changes in the demand and supply for residential and commercial floor space are equal to one another in equations (E.10) and (E.11), and all the other equilibrium conditions of the model are satisfied.

E5 Robustness

In this section of the Online Appendix, we demonstrate the robustness of our counterfactual results across a wide range of specifications. Subsection E5.1 incorporates agglomeration forces. Subsection E5.2 introduces an endogenous supply of floor space. Subsection E5.3 considers an open-city specification. Subsection E5.4 undertakes counterfactuals starting from an initial post-war equilibrium and removing wartime destruction, instead of starting from an initial pre-war equilibrium and introducing wartime destruction.

E5.1 Agglomeration Forces

We first consider a generalization of our baseline closed-city specification in Subsection E1.1 to allow for agglomeration forces in production. We use the conventional specification of agglomeration forces, in which productivity is a constant elasticity function of employment density:

$$A_i = a_i \left(\frac{E_i}{K_i} \right)^{\chi_E}, \qquad \chi_E > 0, \tag{E.16}$$

where a_i are exogenous location fundamentals; $E_i = \sum_{o \in \mathbb{O}} E_i^o$ is total employment; and K_i is land area. We assume a standard value for the elasticity of productivity to employment density of $\chi_E = 0.05$, in line with existing empirical estimates, as reviewed in Rosenthal and Strange (2004).

We undertake our two sets of counterfactuals for wartime destruction (Section 7.1 of the paper) and neighborhood effects (Section 7.2 of the paper), incorporating these agglomeration forces. Productivity now responds endogenously to counterfactual changes in employment in the presence of these agglomeration forces:

$$\widehat{A}_i = \widehat{E}_i^{\chi_E}. \tag{E.17}$$

We modify the system of general equilibrium conditions for a counterfactual equilibrium (E.1)-(E.11) to incorporate this endogenous response in productivity. In particular, we replace equation (E.6) for the endogenous change in the price of commercial floor space ($\widehat{q}_n$) with the following relationship that takes account of these endogenous changes in productivity:

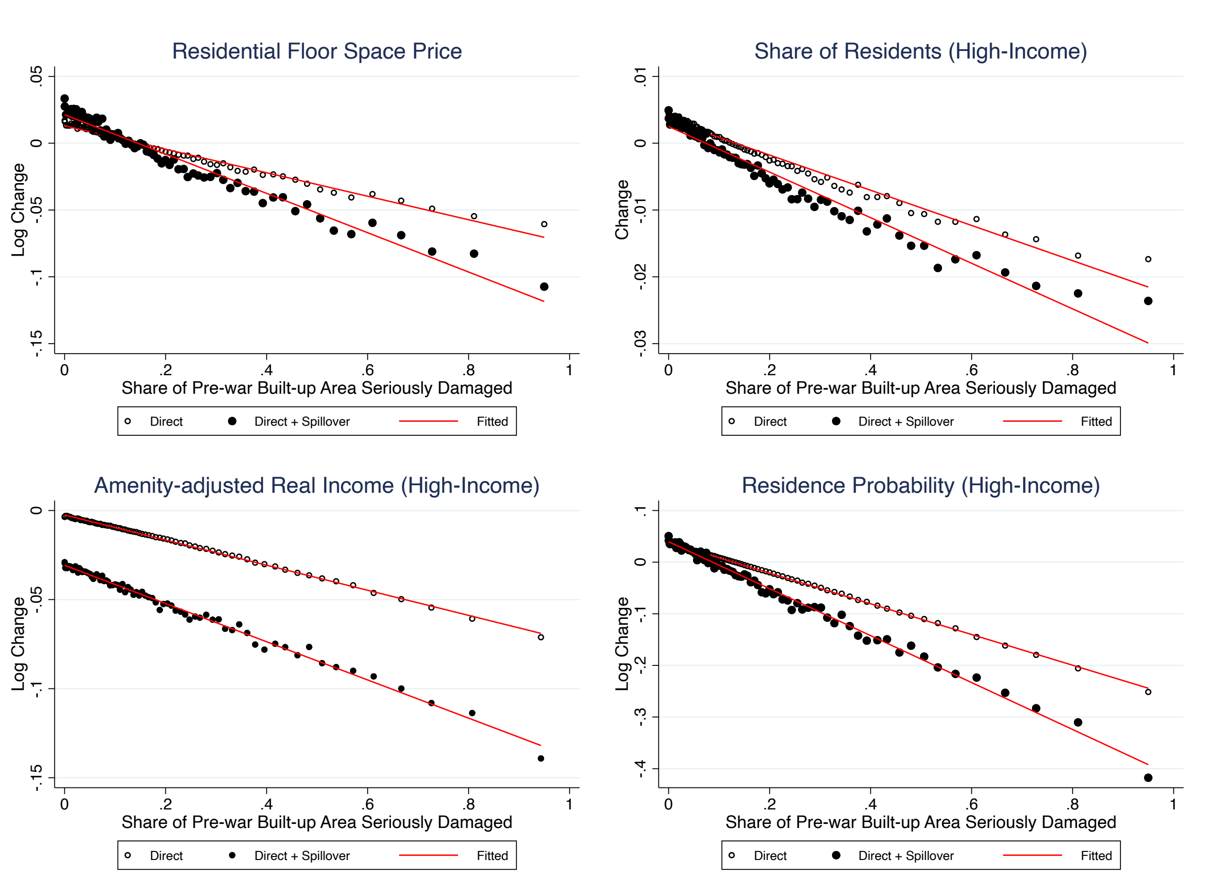
$$\widehat{q}_n = \widehat{A}_n^{\frac{1}{1-\beta}} \left[s_n^L \left(\widehat{w}_n^L \right)^{1-\sigma} + s_n^M \left(\widehat{w}_n^M \right)^{1-\sigma} + s_n^H \left(\widehat{w}_n^H \right)^{1-\sigma} \right]^{-\frac{\beta}{(1-\beta)(1-\sigma)}}. \tag{E.18}$$

All other equilibrium conditions remain the same as in equations (E.1)-(E.11) above.

We solve this augmented system of equations for a counterfactual equilibrium using the same algorithm as discussed in Subsection E1.1 above. We start with initial guesses for changes in wages ($\widehat{w}_n^{o0}$) and the price of residential floor space ($\widehat{Q}_n^0$). We update these initial guesses until the changes in the demand and supply of labor are equal to one another for each occupation in equations (E.8) and (E.9), the changes in the demand and supply for residential and commercial floor space are equal to one another in equations (E.10) and (E.11), and all the other equilibrium conditions of the model are satisfied.

Figure E.2 reports the results of our counterfactuals for the impact of wartime destruction, and corresponds to Figure 4 in the paper. Each panel shows a binscatter across Output Areas of the counterfactual change in a variable against the share of the pre-war built-up area seriously damaged during the Second World War. The circles correspond to percentiles of the distribution of wartime destruction and the red line represents the linear fit. Hollow circles show counterfactuals using the estimated direct effects of wartime destruction alone (setting the spillover coefficients to zero). Solid circles show counterfactuals incorporating both the direct and spillover effects of wartime destruction (using the estimated direct and spillover coefficients).

Figure E.2: Counterfactuals for Wartime Destruction (Agglomeration Robustness)

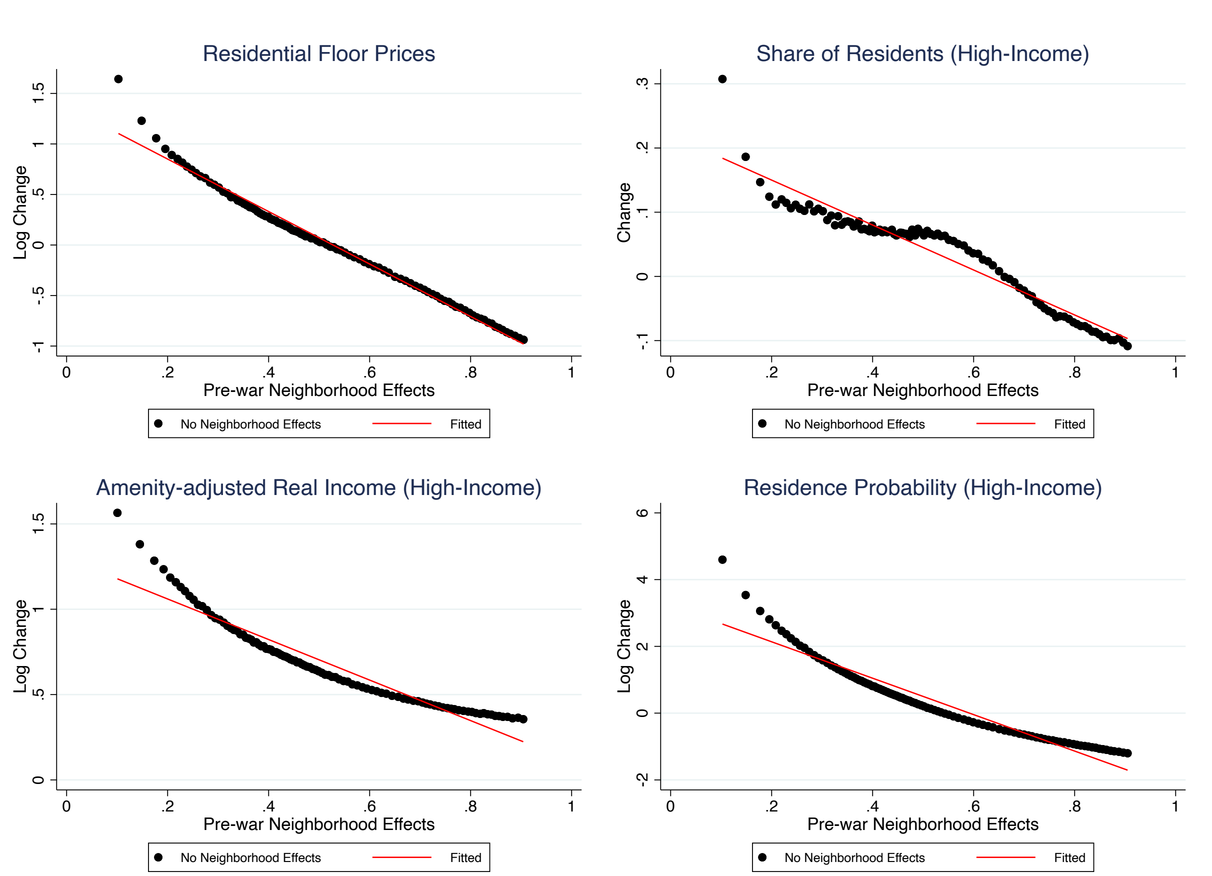


Notes: Counterfactuals using our general specification of neighborhood effects and the estimated direct and spillover effects of wartime destruction from Columns (2), (4) and (6) of Table 4 in the paper and assuming agglomeration forces with an elasticity of productivity with respect to employment density of 0.05; circles show values by percentile of the share of the pre-war built-up area seriously damaged during the Second World War; around 40 percent of locations experience zero destruction, such that the circle for zero destruction captures the first 40 percentiles; hollow circles show counterfactuals using the estimated direct coefficients alone (setting the spillover coefficients to zero); solid circles show counterfactuals using the estimated direct and spillover coefficients; red lines show the linear fit; Panel A shows the log change in the price of residential floor space; Panel B shows the change (not in logs) in the share of a location's residents who are high-income $(R_n^H / (R_n^L + R_n^M + R_n^H))$; Panel C shows the log change in amenity-adjusted real income for high-income residents $(\log(B_n^H w_i^H Q_n^{-(1-\alpha^H)}))$; Panel D shows the log change in the share of all high-income residents in the LCC area who live in a location $(\lambda_n^H = R_n^H / \sum_{i \in \mathbb{N}} R_i^H)$.

Comparing Figures 4 and E.2, we find the same pattern of results in this robustness test incorporating agglomeration forces as in our baseline specification in the paper. We find substantial distributional consequences of wartime destruction for both landlord income and the amenity-adjusted real income of workers (Panels A and C). We find that spatial sorting plays an important role in shaping these distributional consequences (Panels B and D). In both cases, we find that neighborhood effects magnify the impact of wartime destruction, with the full effects (solid circles) notably larger than the direct effects (hollow circles).

Figure E.3 reports the results of our counterfactuals for neighborhood effects, and corresponds to Figure 5 in the paper. Each panel shows a binscatter across Output Areas of the counterfactual change in a variable against pre-war neighborhood effects ($\mathbb{B}_{n,\text{pre-war}}$). The solid circles correspond to percentiles of the distribution of pre-war neighborhood effects and the red line represents the linear fit.

Figure E.3: Counterfactuals Removing Neighborhood Effects (Agglomeration Robustness)



Notes: Counterfactuals for removing neighborhood effects ($\widehat{B}^o_n = 1/\left(\mathbb{B}_{n,\text{pre-war}}\right)^{\eta^o_R}$) based on the estimated coefficients from Column (5) of Table 5 in the paper and assuming agglomeration forces with an elasticity of productivity with respect to employment density of 0.05; solid circles show values by percentile of pre-war neighborhood effects ($\mathbb{B}_{n,\text{pre-war}}$); red line shows the linear fit; Panel A shows the log change in the price of residential floor space; Panel B shows the change (not in logs) in the share of a location's residents who are high-income ($R^H_n / \left(R^L_n + R^M_n + R^H_n\right)$); Panel C shows the log change in amenity-adjusted real income for high-income residents ($\log(B^H_n w^H_i Q_n^{-(1-\alpha^H)})$); Panel D shows the log change in the share of all high-income residents in the LCC area who live in a location ($\lambda^H_n = R^H_n / \sum_{i \in \mathbb{N}} R^H_i$).

Comparing Figures 5 and E.3, we again find the same pattern of results in this robustness test incorporating agglomeration forces as in our baseline specification in the paper. We find that neighborhood effects make a substantial contribution to the observed differences in property prices and socioeconomic composition across locations (Panels A-D). When high-income workers no longer value surrounding socioeconomic composition, they are no longer willing to pay the high prices of residential floor space to live in initially-more-exclusive neighborhoods, and instead find it attractive to move to initially-less-exclusive neighborhoods. As high-income workers reallocate across neighborhoods, this bids down the price of residential floor space in initially-more-exclusive neighborhoods, and bids up the price of residential floor space in initially-less-exclusive neighborhoods. We find that this reallocation has substantial distributional consequences for both landlord income (Panel A) and the amenity-adjusted real income of workers (Panel C).

E5.2 Endogenous Supply of Floor Space

We next consider another generalization of our baseline closed-city specification to allow for an endogenous supply of floor space. We assume a constant elasticity specification for the supply of floor space following Saiz (2010), Epple et al. (2010) and Combes et al. (2019).

In particular, we assume that the quantity of residential floor space (H_{nt}^{R}) depends on inputs of land (K_n) and capital (M_{nt}) according to the following Cobb-Douglas construction technology:

$$H_{nt}^{R} = M_{nt}^{\mu}\left(\left(1-\theta_{nt}\right)K_n\right)^{1-\mu}, \qquad 0<\mu<1, \tag{E.19}$$

where $(1-\theta_{nt})$ is the fraction of land allocated to commercial use, which we assume to be exogenously determined in our empirical setting, and capital is assumed to be in perfectly elastic supply from the wider economy at a constant price of R_t.

From cost minimization and zero profits in construction, equilibrium payments for capital are a constant share of payments for residential floor space ($Q_{nt}H_{nt}^{R}$):

$$R_t M_{nt} = \mu Q_{nt} H_{nt}^{R}. \tag{E.20}$$

Using this equilibrium condition (E.20) to substitute for capital (M_{nt}) in the construction technology (E.19), we obtain the following constant elasticity supply function for residential floor space:

$$H_{nt}^{R} = \left(\frac{\mu}{R_t}\right)^{\frac{\mu}{1-\mu}} Q_{nt}^{\frac{\mu}{1-\mu}}\left(1-\theta_{nt}\right)K_n, \tag{E.21}$$

where $\mu/(1-\mu)$ corresponds to the elasticity of the supply of floor space with respect to its price.

We undertake our two sets of counterfactuals for wartime destruction and neighborhood effects, allowing for this endogenous response in the supply of residential floor space:

$$\widehat{H}_{n}^{R} = \widehat{Q}_{n}^{\frac{\mu}{1-\mu}}. \tag{E.22}$$

In particular, we modify the system of general equilibrium conditions for a counterfactual equilibrium (E.1)-(E.11) to take into account this endogenous response in the supply of residential floor space by replacing equation (E.10) with the following equality between income from residential floor space and expenditure on residential floor space:

$$\widehat{\Omega}_{n}^{S} = \left(\widehat{Q}_{n}^{0}\right)^{\frac{1}{1-\mu}}. \tag{E.23}$$

We use the estimated floor space supply elasticity based on the construction of London's 19th-century railway network in Heblich et al. (2020) of $\frac{\mu}{1-\mu} = 1.83$.

We solve this modified system of equations for a counterfactual equilibrium using the same algorithm as discussed in Subsection E1.1 above. We start with initial guesses for changes in wages ($\widehat{w}_n^{o0}$) and the price of residential floor space ($\widehat{Q}_n^0$). We update these initial guesses until the changes in the demand and supply of labor are equal to one another for each occupation in equations (E.8) and (E.9), the changes in the demand and supply for residential and commercial floor space are equal to one another in equations (E.23) and (E.11), and all the other equilibrium conditions of the model are satisfied.

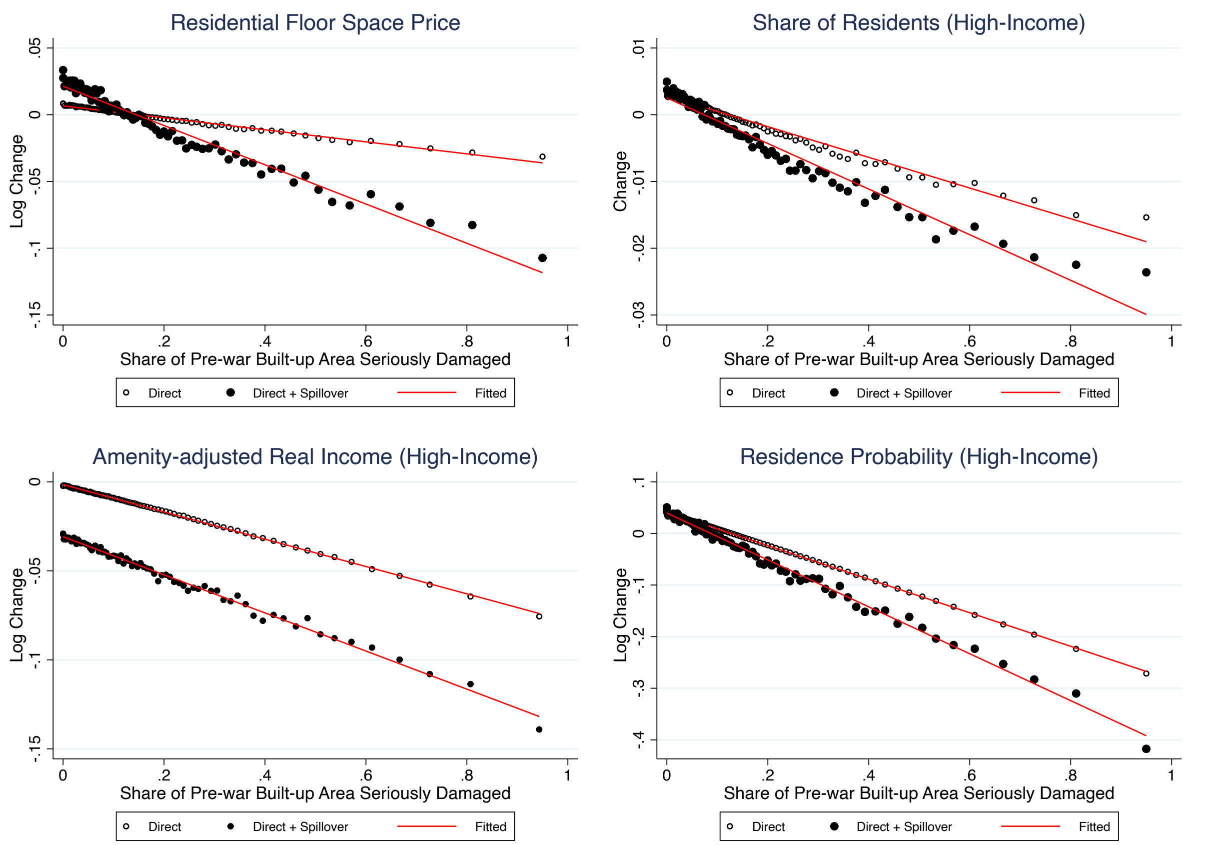
Figure E.4 reports the results of our counterfactuals for the impact of wartime destruction, and corresponds to Figure 4 in the paper. Each panel shows a binscatter across Output Areas of the counterfactual change in a variable against the share of the pre-war built-up area seriously damaged during the Second World War. The circles correspond to percentiles of the distribution of wartime destruction and the red line represents the linear fit. Hollow circles show counterfactuals using the estimated direct effects of wartime destruction alone (setting the spillover coefficients to zero). Solid circles show counterfactuals incorporating both the direct and spillover effects of wartime destruction (using the estimated direct and spillover coefficients).

Comparing Figures 4 and E.4, we find the same pattern of results in this robustness test incorporating an endogenous response in the supply of floor space as in our baseline specification in the paper. We find substantial distributional consequences of wartime destruction for both landlord income and the amenity-adjusted real income of workers (Panels A and C). We find that spatial sorting plays an important role in shaping these distributional consequences (Panels B and D). In both cases, we find that neighborhood effects magnify the impact of wartime destruction, with the full effects (solid circles) notably larger than the direct effects (hollow circles).

Figure E.5 reports the results of our counterfactuals for neighborhood effects, and corresponds to Figure 5 in the paper. Each panel shows a binscatter across Output Areas of the counterfactual change in a variable against pre-war neighborhood effects ($\mathbb{B}_{n,\text{pre-war}}$). The solid circles correspond to percentiles of the distribution of pre-war neighborhood effects and the red line represents the linear fit.

Comparing Figures 5 and E.5, we again find the same pattern of results in this robustness test incorporating an endogenous response in the supply of floor space as in our baseline specification in the paper. We find that neighborhood effects make a substantial contribution to the observed differences in property prices and socioeconomic composition across locations (Panels A-D). When high-income workers no longer value surrounding socioeconomic composition, they are no longer willing to pay the high prices of residential floor space to live in initially-more-exclusive neighborhoods, and instead find it attractive to move to initially-less-exclusive

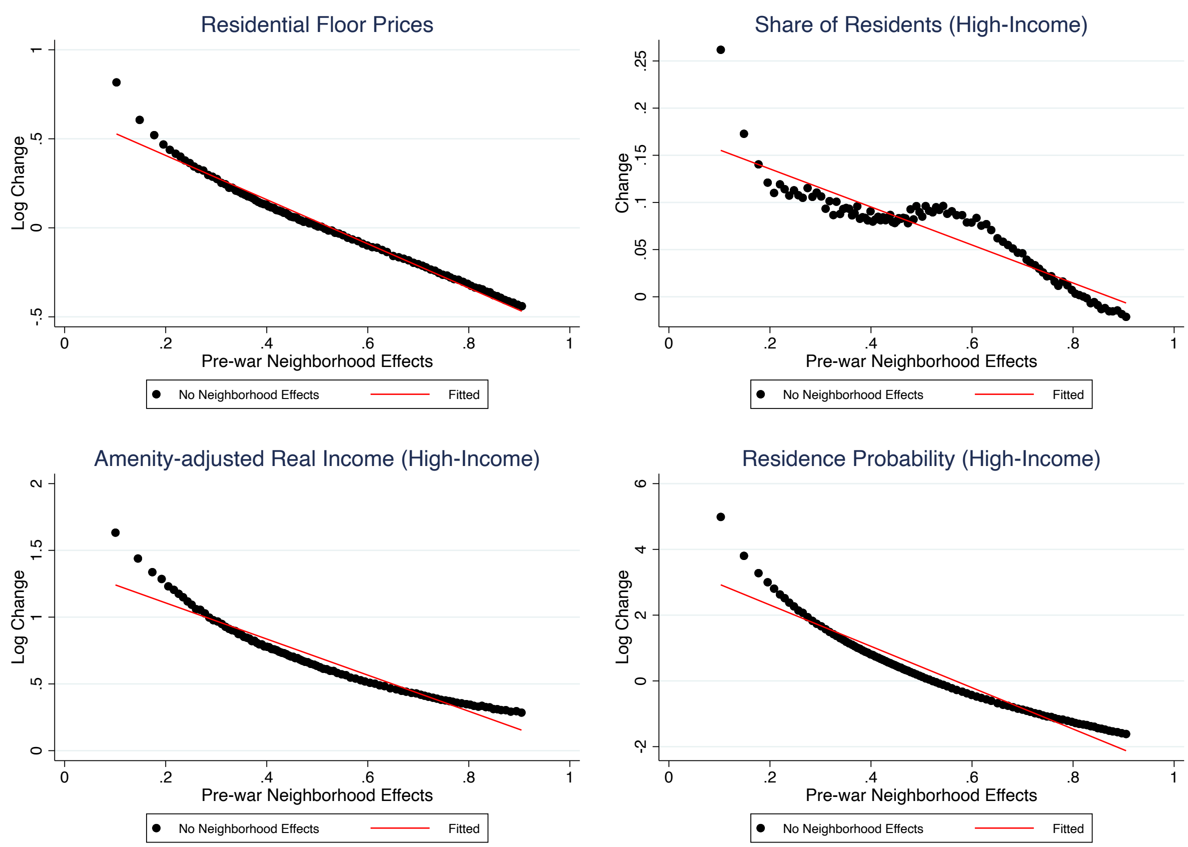
Figure E.4: Counterfactuals for Wartime Destruction (Endogenous Floor Space Robustness)



Notes: Counterfactuals using our general specification of neighborhood effects and the estimated direct and spillover effects of wartime destruction from Columns (2), (4) and (6) of Table 4 in the paper and assuming an elasticity of the supply of residential floor space with respect to its price of $\mu/(1-\mu) = 1.83$; circles show values by percentile of the share of the pre-war built-up area seriously damaged during the Second World War; around 40 percent of locations experience zero destruction, such that the circle for zero destruction captures the first 40 percentiles; hollow circles show counterfactuals using the estimated direct coefficients alone (setting the spillover coefficients to zero); solid circles show counterfactuals using the estimated direct and spillover coefficients; red lines show the linear fit; Panel A shows the log change in the price of residential floor space; Panel B shows the change (not in logs) in the share of a location's residents who are high-income ($R_n^H / \left(R_n^L + R_n^M + R_n^H\right)$); Panel C shows the log change in amenity-adjusted real income for high-income residents ($\log(B_n^H w_i^H Q_n^{-(1-\alpha^H)})$); Panel D shows the log change in the share of all high-income residents in the LCC area who live in a location ($\lambda_n^H = R_n^H / \sum_{i \in \mathbb{N}} R_i^H$).

neighborhoods. As high-income workers reallocate across neighborhoods, this bids down the price of residential floor space in initially-more-exclusive neighborhoods, and bids up the price of residential floor space in initially-less-exclusive neighborhoods. We find that this reallocation has substantial distributional consequences for both landlord income (Panel A) and the amenity-adjusted real income of workers (Panel C).

Figure E.5: Counterfactuals Removing Neighborhood Effects (Endogenous Floor Space Robustness)



Notes: Note: Counterfactuals for removing neighborhood effects ($\widehat{B}^o_n = 1/\left(\mathbb{B}_n^{\text{pre-war}}\right)^{\eta_R^o}$) based on the estimated coefficients from Column (5) of Table 5 in the paper and assuming an elasticity of the supply of residential floor space with respect to its price of $\mu/(1-\mu) = 1.83$; solid circles show values by percentile of pre-war neighborhood effects ($\mathbb{B}_{n,\text{pre-war}}$); red line shows the linear fit; Panel A shows the log change in the price of residential floor space; Panel B shows the change (not in logs) in the share of a location's residents who are high-income ($R_n^H / \left(R_n^L + R_n^M + R_n^H\right)$); Panel C shows the log change in amenity-adjusted real income for high-income residents ($\log(B_n^H w_i^H Q_n^{-(1-\alpha^H)})$); Panel D shows the log change in the share of all high-income residents in the LCC area who live in a location ($\lambda_n^H = R_n^H / \sum_{i \in \mathbb{N}} R_i^H$).

E5.3 Open-city Specification

We next consider an open-city specification, in which the measure of total residents from each occupation ($\overline{E}_t^L, \overline{E}_t^M, \overline{E}_t^H$) is endogenous, and the expected utility of workers from each occupation (U_t^L, U_t^M, U_t^H) is pinned down by the reservation level of utility for each occupation in the wider economy ($\widehat{U}_t^o = 1$ for all o). As in our baseline closed-city specification, we hold productivity and the supplies of commercial and residential floor space constant.

We undertake our two sets of counterfactuals for wartime destruction and neighborhood effects, allowing for these endogenous changes in the total residents from each occupation. We solve for a counterfactual equilibrium using the same system of equations as for our baseline closed-city specification in Section E1.1 above, augmented with the following population mobility condition for each occupation that determines the total residents from each occupation:

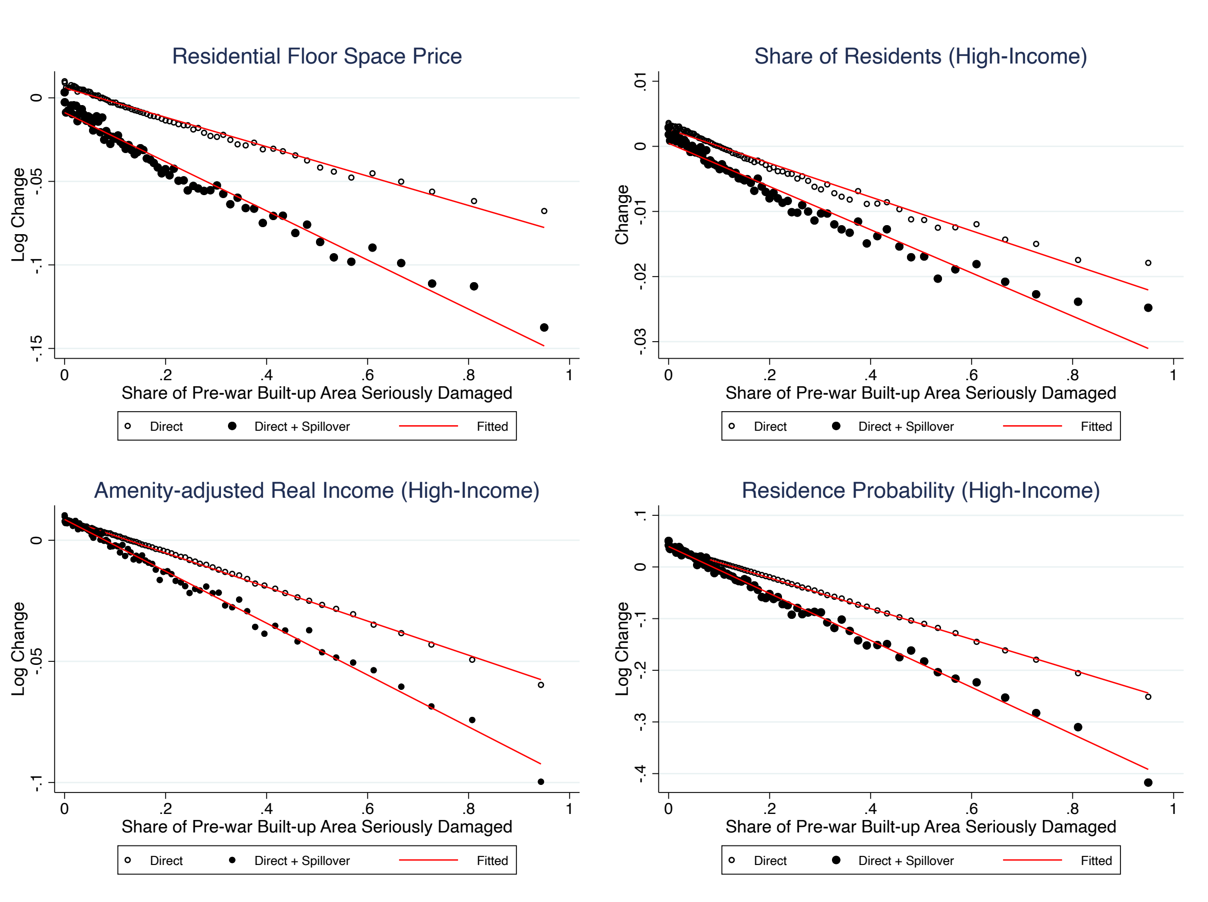
$$\widehat{U}_t^o = \left[\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \lambda_{nit}^o \left(\widehat{B}_{kt}^o \widehat{w}_{\ell t}^o \right)^{\epsilon^o} \left(\widehat{Q}_{kt}^{(1-\alpha^o)} \right)^{-\epsilon^o} \right]^{\frac{1}{\epsilon^o}} = 1. \tag{E.24}$$

We solve this augmented system of equations for a counterfactual equilibrium using a similar algorithm as discussed in Subsection E1.1 above. We start with initial guesses for changes in wages ($\widehat{w}_n^{o0}$), the price of residential floor space ($\widehat{Q}_n^0$) and the measure of total residents from each occupation ($\widehat{E}_t^{o0}$). We update these initial guesses until the changes in the demand and supply of labor are equal to one another for each occupation in equations (E.8) and (E.9), the changes in the demand and supply for residential and commercial floor space are equal to one another in equations (E.23) and (E.11), the population mobility condition holds for each occupation in equation (E.24), and all the other equilibrium conditions of the model are satisfied.

Intuitively, in the initial equilibrium observed in the data, expected utility for each occupation is equal to the reservation level of utility for that occupation in the wider economy. The direct effect of reductions in residential amenities in bombed locations is to depress expected utility for a given occupation below its reservation level of utility in the wider economy. The mechanism that restores equilibrium is an outflow of residents from that occupation to the wider economy, which bids up wages for that occupation and bids down the price of residential floor space, until expected utility for that occupation is again equal to its reservation level of utility in the wider economy. If higher-income workers are more responsive to changes in residential amenities, this outflow of residents to the wider economy will be greater for higher-income workers.

Figure E.6 reports the results of our counterfactuals for the impact of wartime destruction, and corresponds to Figure 4 in the paper. Each panel shows a binscatter across Output Areas of the counterfactual change in a variable against the share of the pre-war built-up area seriously damaged during the Second World War. The circles correspond to percentiles of the distribution of wartime destruction and the red line represents the linear fit. Hollow circles show counterfactuals using the estimated direct effects of wartime destruction alone (setting the spillover coefficients to zero). Solid circles show counterfactuals incorporating both the direct and spillover effects of wartime destruction (using the estimated direct and spillover coefficients).

Figure E.6: Counterfactuals for Wartime Destruction (Open-City Robustness)

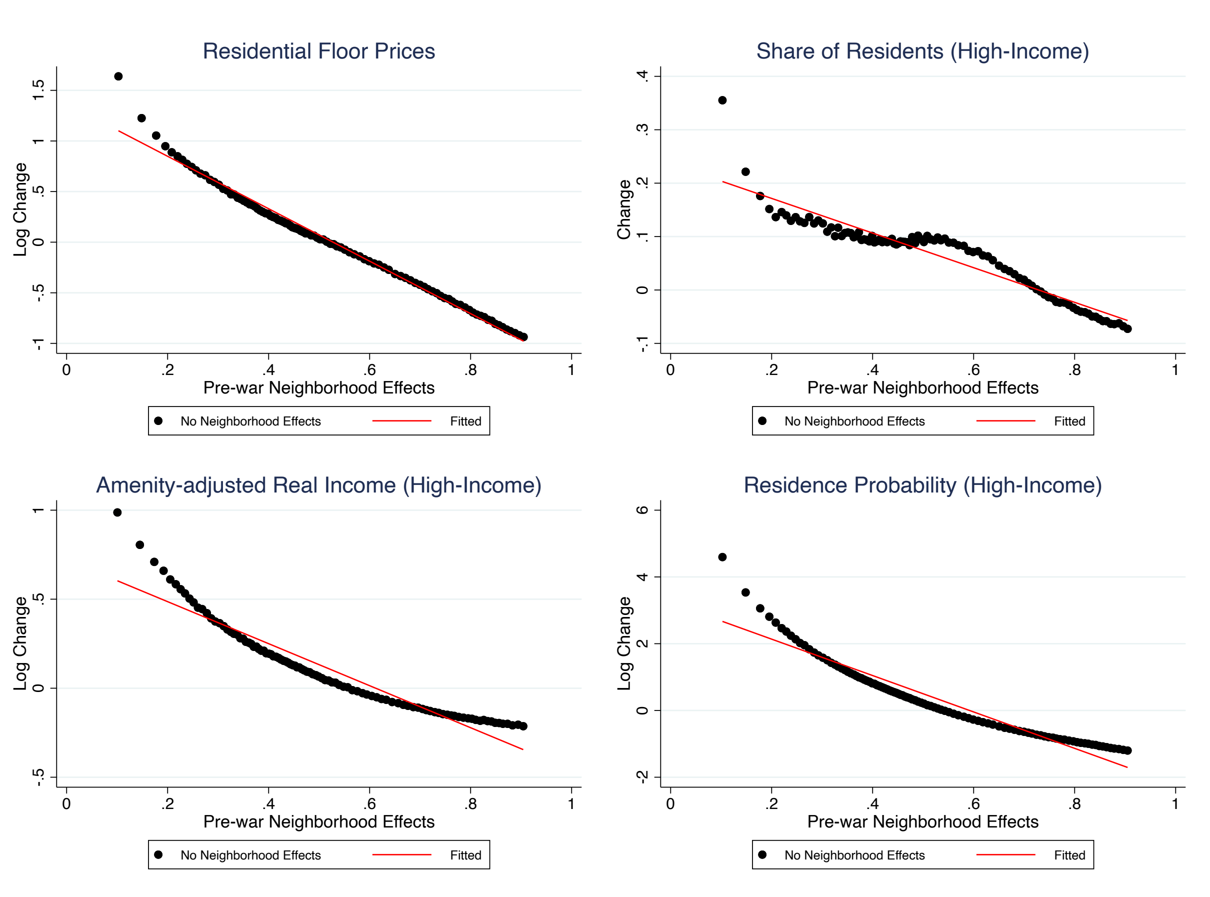


Notes: Counterfactuals using our general specification of neighborhood effects and the estimated direct and spillover effects of wartime destruction from Columns (2), (4) and (6) of Table 4 in the paper and assuming an open city; circles show values by percentile of the share of the pre-war built-up area seriously damaged during the Second World War; around 40 percent of locations experience zero destruction, such that the circle for zero destruction captures the first 40 percentiles; hollow circles show counterfactuals using the estimated direct coefficients alone (setting the spillover coefficients to zero); solid circles show counterfactuals using the estimated direct and spillover coefficients; red lines show the linear fit; Panel A shows the log change in the price of residential floor space; Panel B shows the change (not in logs) in the share of a location's residents who are high-income $(R_n^H / \left(R_n^L + R_n^M + R_n^H\right))$; Panel C shows the log change in amenity-adjusted real income for high-income residents $(\log(B_n^H w_i^H Q_n^{-(1-\alpha^H)}))$; Panel D shows the log change in the share of all high-income residents in the LCC area who live in a location $(\lambda_n^H = R_n^H / \sum_{i \in \mathbb{N}} R_i^H)$.

Comparing Figures 4 and E.6, we find the same pattern of results in this robustness test for an open city as in our baseline specification with a closed city in the paper. We find substantial distributional consequences of wartime destruction for both landlord income and the amenity-adjusted real income of workers (Panels A and C). We find that spatial sorting plays an important role in shaping these distributional consequences (Panels B and D). In both cases, we find that neighborhood effects magnify the impact of wartime destruction, with the full effects (solid circles) notably larger than the direct effects (hollow circles).

Figure E.7 reports the results of our counterfactuals for neighborhood effects, and corresponds to Figure 5 in the paper. Each panel shows a binscatter across Output Areas of the counterfactual change in a variable against pre-war neighborhood effects ($\mathbb{B}_{n,\text{pre-war}}$). The solid circles correspond to percentiles of the distribution of pre-war neighborhood effects and the red line represents the linear fit.

Figure E.7: Counterfactuals Removing Neighborhood Effects (Open-City Robustness)



Notes: Counterfactuals for removing neighborhood effects ($\widehat{B}_n^o = 1/\left(\mathbb{B}_{n,\text{pre-war}}\right)^{\eta_R^o}$) based on the estimated coefficients from Column (5) of Table 5 in the paper and assuming an open city; solid circles show values by percentile of pre-war neighborhood effects ($\mathbb{B}_{n,\text{pre-war}}$); red line shows the linear fit; Panel A shows the log change in the price of residential floor space; Panel B shows the change (not in logs) in the share of a location's residents who are high-income ($R_n^H/\left(R_n^L + R_n^M + R_n^H\right)$); Panel C shows the log change in amenity-adjusted real income for high-income residents ($\log(B_n^H w_i^H Q_n^{-(1-\alpha^H)})$); Panel D shows the log change in the share of all high-income residents in the LCC area who live in a location ($\lambda_n^H = R_n^H / \sum_{i \in \mathbb{N}} R_i^H$).

Comparing Figures 5 and E.7, we again find the same pattern of results in this robustness test for an open city as in our baseline specification for a closed city in the paper. We find that neighborhood effects make a substantial contribution to the observed differences in property prices and socioeconomic composition across locations (Panels A-D). When high-income workers no longer value surrounding socioeconomic composition, they are no longer willing to pay the high prices of residential floor space to live in initially-more-exclusive neighborhoods, and instead find it attractive to move to initially-less-exclusive neighborhoods. As high-income workers reallocate across neighborhoods, this bids down the price of residential floor space in initially-more-exclusive neighborhoods, and bids up the price of residential floor space in initially-less-exclusive neighborhoods. We find that this reallocation has substantial distributional consequences for both landlord income (Panel A) and the amenity-adjusted real income of workers (Panel C).

E5.4 Counterfactuals Starting from the Post-War Equilibrium

In our baseline specification, we undertake counterfactuals for wartime destruction starting at the observed pre-war equilibrium in the data. Although starting at the observed pre-war equilibrium is the natural choice to evaluate the general equilibrium impact of wartime destruction, in this section of the Online Appendix we show that we find a similar pattern of results if we instead start of the observed post-war equilibrium and remove the effects of wartime destruction.

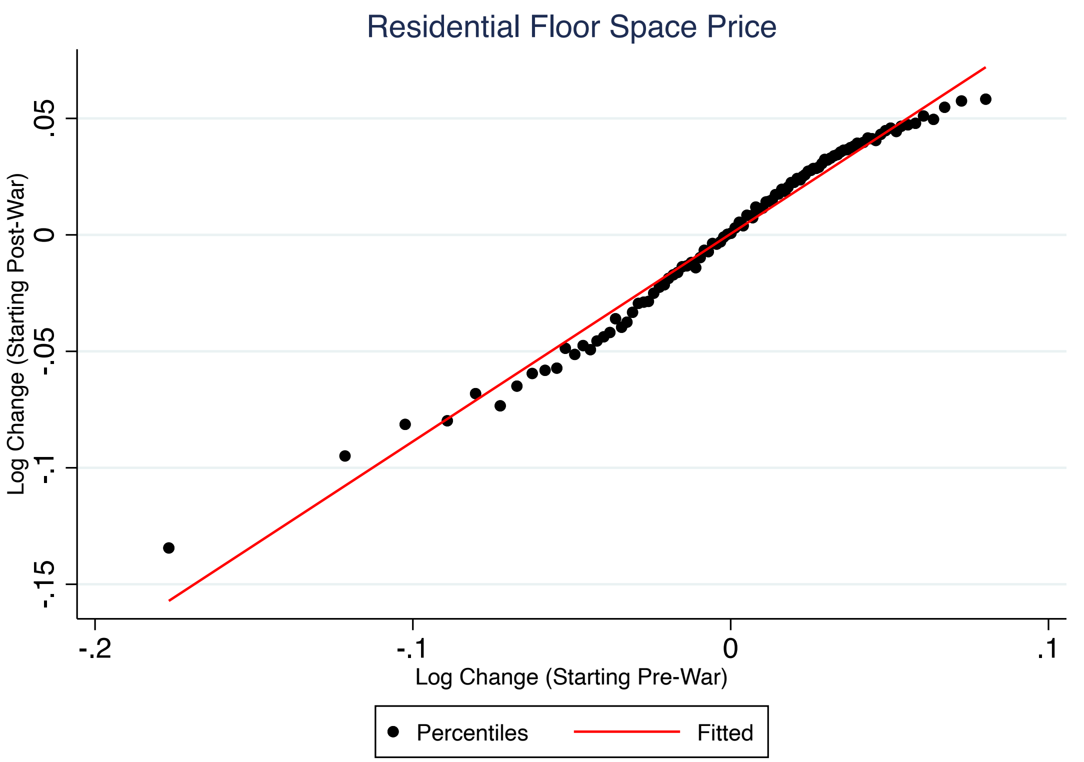
We undertake this robustness test for our general specification of neighborhood effects, using the estimated direct and spillover coefficients for high, middle and low-income workers from Table 4 in the paper. We start at the observed values of the endogenous variables in the post-war equilibrium and remove the impact of wartime destruction by multiplying the estimated coefficients on own and neighboring destruction by minus one. We compute the implied exogenous changes in amenities ($\widehat{B}_n^o$) removing the effects of wartime destruction as:

$$\widehat{B}_n^o = e^{-\eta_D^o D_n} \prod_{g=1}^{G} e^{-\eta_{Bg}^o D_{ng}}, \tag{E.25}$$

where D_n is destruction in the own area for location n; η_D^o is the estimated coefficient on own destruction for occupation o; D_{ng} is destruction in buffer g for location n; and η_{Bg}^o is the estimated coefficient for occupation o on destruction in buffer g.

Given these exogenous changes in amenities, we solve for a counterfactual equilibrium for wartime destruction using the same system of general equilibrium conditions as in Subsection E1.1 of this Online Appendix. Figure E.8 reports the results of these counterfactuals for the full effects of wartime destruction (incorporating both direct and spillover effects). The vertical axis shows log changes in the price of residential floor space for counterfactuals starting from the initial post-war equilibrium. The horizontal axis shows log changes in the price of residential floor space for counterfactuals starting from the initial pre-war equilibrium. In both cases, the figure displays the log property price in an equilibrium with wartime destruction minus the log property price in an equilibrium without wartime destruction, so that the two sets of counterfactual predictions for the impact of wartime destruction are comparable to one another. The solid circles correspond to percentiles of the distribution of log changes in the price of residential floor space starting from the pre-war equilibrium and the red line represents the linear fit. As shown

Figure E.8: Counterfactuals for Wartime Destruction (Starting from the Initial Post-War Equilibrium Versus the Initial Pre-War Equilibrium)



Notes: Counterfactuals using our general specification of neighborhood effects and the estimated direct and spillover effects of wartime destruction from Columns (2), (4) and (6) of Table 4 in the paper; horizontal axis shows counterfactual log changes in the price of residential floor space starting from the initial pre-war equilibrium; vertical axis shows counterfactual log changes in the price of residential floor space starting from the initial post-war equilibrium; in both cases, we display the log price of residential floor space in an equilibrium with wartime destruction minus the log price of residential floor space in an equilibrium without wartime destruction, so that the two sets of counterfactual predictions are comparable to one another; solid circles show the full effects of wartime destruction (including both direct and spillover effects) by percentile of the counterfactual log change in the price of residential floor space starting from the initial pre-war equilibrium.

in the figure, we find a similar pattern of counterfactual predictions whether we start from the initial pre-war equilibrium or the initial post-war equilibrium.

F Additional Empirical Results

This section of the online appendix reports additional reduced-form empirical results for Section 4 of the paper. Section F2 reports our robustness check using our bomb damage index. Section F1 provides evidence on potential heterogeneity in the treatment effect of wartime destruction. Section F3 re-estimates our reduced-form specifications from Section 4 of the paper using Conley (1999) standard errors with a 1 km distance threshold, instead of clustering the standard errors on 1 km hexagons. Section F4 re-estimates the impact of wartime destruction on post-war socioeconomic outcomes using data from the 2011 UK population census, instead of using data from the 2001 UK population census, as in the main paper. Section F5 re-estimates the impact of wartime destruction on post-war property values for sub-periods, instead of the full sample period from 1995 to 2020, as in the paper. Section F6 re-estimates the impact of wartime destruction on post-war socioeconomic outcomes excluding observations from the Cities of London and Westminsters, as the two most central and non-residential boroughs of the LCC area. Section F7 provides further evidence on the mechanisms through which the direct and spillover effects of wartime destruction occur, augmenting the discussion in Section 4.4 of the paper.

F1 Treatment Heterogeneity

In this section of the Online Appendix, we provide evidence on potential heterogeneity in the treatment effect of wartime destruction, supplementing the discussion in Section 4.2 of the paper. We augment our baseline specification in equation (3) in the paper with interaction terms between wartime destruction and indicator variables for quintiles of pre-war socioeconomic status:

$$Y_{i,\text{post-war}} = \beta D_{i,\text{war}} + \sum_{g=1}^{5} \delta_g \left(\mathbb{I}_{ig,\text{pre-war}} \times D_{i,\text{war}} \right) + \sum_{g=1}^{5} \theta_g \left(\mathbb{I}_{ig,\text{pre-war}} \right) + \varrho_k + u_i \tag{F.1}$$

where i indexes Output Areas, g indexes quintiles of pre-war socioeconomic status, and k indexes hexagonal grid cells; $Y_{i,\text{post-war}}$ is a socioeconomic outcome after the end of the Second World War; $D_{i,\text{war}}$ is our baseline measure of wartime destruction based on the fraction of the built-up area seriously damaged; $\mathbb{I}_{ig,\text{pre-war}}$ is an indicator variable that is one if Output Area i is in quintile g of the distribution of pre-war socioeconomic status; ϱ_k are fixed effects for 1 km hexagonal grid cells; u_i is stochastic error; and we report standard errors clustered by 1 km hexagons.

Table F.1: The Direct Effect of Wartime Destruction: Heterogeneity

	(1) Fraction High Status	(2) Fraction Middle Status	(3) Fraction Low Status	(4) Socio-Economic Index	(5) Log of Property Value	(6) Log of Property Value
Destruction in Own Area	-0.049^{***}	-0.019^{**}	0.069^{***}	-0.059^{***}	-0.187^{***}	-0.106^{***}
	(0.011)	(0.009)	(0.016)	(0.013)	(0.042)	(0.036)
Bombing * 2nd Quintile SES	0.010	-0.020^{*}	0.010	0.000	-0.003	0.003
	(0.016)	(0.011)	(0.021)	(0.018)	(0.059)	(0.051)
Bombing * 3rd Quintile SES	0.007	-0.015	0.008	-0.001	0.004	-0.001
	(0.018)	(0.014)	(0.024)	(0.020)	(0.062)	(0.053)
Bombing * 4th Quintile SES	0.009	0.001	-0.010	0.010	-0.006	-0.014
	(0.016)	(0.013)	(0.021)	(0.017)	(0.064)	(0.053)
Bombing * 5th Quintile SES	0.026^{*}	0.020	-0.045^{**}	0.036^{**}	0.074	0.000
	(0.015)	(0.012)	(0.020)	(0.017)	(0.063)	(0.051)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8720	8720	8720	8720	7953	7953
R-squared	0.549	0.298	0.496	0.538	0.681	0.822

Notes: This table re-estimates the direct impact regressions of Table 2 in the main paper allowing for heterogeneity by quintiles of the pre-war socio-economic status (SES). The first quintile is the excluded group of Output Areas in the bottom twenty percent of the distribution of pre-war SES. The main explanatory variable is the fraction of the pre-war built-up area seriously damaged in each Output Area during the Second World War. We also control for the main effects of the quintile dummies, which are not reported in the table. All regressions include fixed effects for 1 km hexagons. Standard errors are clustered at the 1 kilometer hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

In this augmented specification, the main effect of wartime destruction (β) captures the treatment effect for the first quintile, while the sum of the main effect (β) and the coefficient on the interaction term (δ_g) captures the treatment effect for the other quintiles. Table F.1 reports the estimation results. The structure of the table is the same as in Table 2 in the paper. Each column corresponds to a different post-war outcome. Columns (1)-(3) use the fraction of the population with high, middle and low socioeconomic status, respectively; Column (4) uses our index of socioeconomic status; Column (5) uses the unconditional average property value; Column (6) uses the average property value conditional on observed property characteristics.

Across all of these specifications, we find the same pattern of estimated coefficients for the main effect effect of wartime destruction (β) as in our baseline specification in Table 2 in the paper. We find coefficients on the interaction terms (δ_g) that are substantially smaller than the main effect and in most cases statistically insignificant, which implies a relatively homogenous treatment effect of wartime destruction across quintiles of pre-war socioeconomic status. This pattern of results is consistent with the idea that the negative effect of wartime destruction is mainly driven by the post-war construction of council housing, rather than by the characteristics of the pre-war houses that were destroyed.

F2 Bomb Damage Index

In this section of the Online Appendix, we report our robustness check using our bomb damage index, instead of our baseline measure of wartime destruction that uses the share of the pre-war built-up area that experienced serious damage or worse. To construct this bomb damage index, we first score levels of damage to each building from 0 to 6 (from no to total destruction). We next compute our bomb damage index for each Output Area as the weighted average of these scores, using the shares of its pre-war built-up area with each level of destruction.

Table F.2 re-estimates Table 1 in the paper for the randomness of Second World War destruction; Table F.3 re-estimates Table 2 in the paper for the direct effects of wartime destruction; Table F.4 re-estimates Table 3 in the paper for the spillover effects of wartime destruction; Table F.5 re-estimates Table F.17 from Online Appendix F7 for the mechanisms for wartime destruction.

In all four tables, we find the same pattern of results with this alternative measure of wartime destruction as in our baseline specification in Section 4 of the paper. After controlling for fixed effects for 1 km hexagons, we continue to find that wartime destruction is entirely uncorrelated with pre-war socioeconomic outcomes, providing empirical support for its use as an exogenous source of variation. Even after controlling for these 1km fixed effects, we find negative and statistically significant direct effects of wartime destruction on post-war outcomes in the own location, and negative and statistically significant effects spillover effects on post-war outcomes in neighboring locations. Finally, after controlling for these 1km fixed effects, we find direct effects of wartime destruction on the type of buildings, but no evidence of spillover effects of wartime destruction on the type of buildings.

Table F.2: Randomness of Wartime Destruction: Bomb Damage Index

Fixed Effects	(1) Fraction High Status	(2) Fraction Middle Status	(3) Fraction Low Status	(4) Socio-Economic Index	(5) Log of Property Value	(6) Log of Property Value
None	-0.056^{***}	0.012^{**}	0.044^{***}	-0.050^{***}	-0.091^{***}	-0.100^{***}
	(0.007)	(0.005)	(0.005)	(0.006)	(0.012)	(0.013)
4 km Hexagons	-0.012^{***}	0.005	0.007^{*}	-0.009^{**}	-0.008	-0.015^{*}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.009)	(0.008)
1 km Hexagons	0.000	0.000	0.000	0.000	0.011	0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.007)	(0.007)

Notes: This table re-estimates the specifications in Table 1 in the main paper using a bomb damage index based on the built-up area as described in the main text as the explanatory variable in each panel. Standard errors are clustered at the 1 kilometer hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

Table F.3: The Direct Effect of Wartime Destruction: Bomb Damage Index

	(1) Fraction High Status	(2) Fraction Middle Status	(3) Fraction Low Status	(4) Socio-Economic Index	(5) Log of Property Value	(6) Log of Property Value
Destruction in Own Area	−0.008***	−0.004***	0.012***	−0.010***	−0.035***	−0.022***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.005)	(0.004)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8912	8912	8912	8912	8111	8111
R-squared	0.505	0.279	0.438	0.483	0.658	0.800

Notes: This table re-estimates the specifications in Table 2 in the main paper using a bomb damage index based on the built-up area as described in the main text as the explanatory variable. Standard errors are clustered at the 1 kilometer hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

Table F.4: The Spillover Effect of Wartime Destruction: Bomb Damage Index

	Socio-Economic Index		Log of Property Value		Log of Property Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Destruction in Own Area	−0.010***	−0.009***	−0.035***	−0.030***	−0.022***	−0.018***
	(0.002)	(0.001)	(0.005)	(0.005)	(0.004)	(0.004)
Destruction in 100m Buffer		−0.005		−0.012		−0.012
		(0.003)		(0.010)		(0.009)
Destruction in 200m Buffer		−0.004		−0.021*		−0.017*
		(0.004)		(0.012)		(0.010)
Destruction in 300m Buffer		−0.005		−0.013		−0.019
		(0.004)		(0.014)		(0.012)
Destruction in 400m Buffer		0.003		−0.004		−0.001
		(0.005)		(0.016)		(0.012)
Destruction in 500m Buffer		−0.002		−0.005		−0.009
		(0.005)		(0.016)		(0.014)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8912	8909	8111	8108	8111	8108
R-squared	0.483	0.483	0.658	0.659	0.800	0.800

Notes: This table re-estimates the specifications in Table 3 in the main paper using a bomb damage index based on the built-up area as described in the main text as the explanatory variable in each column. Standard errors are clustered at the 1 kilometer hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

Table F.5: The Mechanisms of Wartime Destruction: Bomb Damage Index

	(1) Fraction Buildings Surviving	(2) Log of Height of Buildings	(3) Fraction Land Area Built-Up	(4) Fraction Council in 2001	(5) Fraction Council in 1981	(6) Log of Empl. Density
Destruction in Own Area	−0.058***	0.015***	−0.009***	0.030***	0.046***	−0.053***
	(0.005)	(0.004)	(0.001)	(0.004)	(0.006)	(0.014)
Destruction in 100m Buffer	0.009	−0.012	0.001	0.004	0.005	−0.012
	(0.009)	(0.008)	(0.003)	(0.007)	(0.012)	(0.029)
Destruction in 200m Buffer	0.015	0.010	0.001	0.003	0.009	0.036
	(0.011)	(0.011)	(0.003)	(0.009)	(0.015)	(0.035)
Destruction in 300m Buffer	−0.009	−0.025**	0.007*	−0.002	−0.008	0.006
	(0.012)	(0.012)	(0.003)	(0.009)	(0.016)	(0.037)
Destruction in 400m Buffer	0.035**	−0.035***	0.001	−0.017	−0.041**	−0.004
	(0.015)	(0.013)	(0.004)	(0.010)	(0.018)	(0.042)
Destruction in 500m Buffer	−0.021	0.014	0.004	0.005	0.003	0.043
	(0.015)	(0.016)	(0.004)	(0.012)	(0.020)	(0.045)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8909	8909	8909	8909	6697	8909
R-squared	0.403	0.473	0.465	0.396	0.442	0.479

Notes: This table re-estimates the specifications in Table F.17 in this Online Appendix using a bomb damage index based on the pre-war built-up area as described in the main text as the explanatory variable in each column. Observations differ in column (5) due to the different spatial units used in the 1981 census. Standard errors are clustered at the 1 kilometer hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

F3 Conley (1999) Standard Errors

In our baseline specification in Section 4 of the paper, we report standard errors clustered on 1 km hexagons.[5] In this section of the Online Appendix, we demonstrate the robustness of our results to using Heteroskedasticity and Autocorrelation Consistent (HAC) standard errors following Conley (1999).

Table F.6 re-estimates Table 1 in the main paper for the randomness of Second World War destruction; Table F.7 re-estimates Table 2 in the main paper for the direct effects of wartime destruction; Table F.8 re-estimates Table 3 in the main paper for the spillover effects of wartime destruction; Table F.9 re-estimates Table F.17 from Online Appendix F7 for the mechanisms for wartime destruction.

In all four tables, we find the same pattern of results using this alternative method of computing standard errors as in our baseline specification in Section 4 of the paper. After controlling for fixed effects for 1 km hexagons, we find that wartime destruction is entirely uncorrelated with

[5]Bertrand et al. (2004) examine several approaches to control for serial correlation. They show that clustering the standard errors performs well in settings with at least 50 clusters as in our application.

pre-war socioeconomic outcomes, providing empirical support for its use as an exogenous source of variation. Even after controlling for these 1km fixed effects, we find negative and statistically significant direct effects of wartime destruction on post-war socioeconomic outcomes in the own location, and negative and statistically significant effects spillover effects on post-war socioeconomic outcomes in neighboring locations. Finally, after controlling for these 1km fixed effects, we find direct effects of wartime destruction on the type of buildings, but no evidence of spillover effects of wartime destruction on the type of buildings.

Table F.6: Randomness of Wartime Destruction: Conley (1999) Standard Errors

Fixed Effects	(1) Fraction High Status	(2) Fraction Middle Status	(3) Fraction Low Status	(4) Socio-Economic Index	(5) Log of Property Value	(6) Log of Property Value
None	-0.235^{***}	0.039^{*}	0.196^{***}	-0.215^{***}	-0.478^{***}	-0.517^{***}
	(0.032)	(0.022)	(0.025)	(0.026)	(0.056)	(0.061)
4km Hexagons	-0.061^{***}	0.020	0.042^{***}	-0.051^{***}	-0.099^{***}	-0.124^{***}
	(0.020)	(0.017)	(0.016)	(0.016)	(0.037)	(0.035)
1km Hexagons	−0.007	−0.004	0.011	−0.009	−0.003	−0.026
	(0.014)	(0.013)	(0.011)	(0.011)	(0.027)	(0.025)

Notes: This table re-estimates Table 1 in the main paper using Conley (1999) standard errors with a distance cutoff of 1 kilometer instead of standard errors that are clustered at the 1 km hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

Table F.7: The Direct Effect of Wartime Destruction: Conley (1999) Standard Errors

	(1) Fraction High Status	(2) Fraction Middle Status	(3) Fraction Low Status	(4) Socio-Economic Index	(5) Log of Property Value	(6) Log of Property Value
Destruction in Own Area	-0.039^{***}	-0.023^{***}	0.062^{***}	-0.051^{***}	-0.174^{***}	-0.112^{***}
	(0.006)	(0.004)	(0.008)	(0.007)	(0.020)	(0.016)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8912	8912	8912	8912	8111	8111
R-squared	0.505	0.280	0.439	0.483	0.658	0.799

Notes: This table re-estimates Table 2 in the main paper using Conley (1999) standard errors with a distance cutoff of 1 kilometer instead of standard errors that are clustered at the 1 km hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

Table F.8: The Spillover Effect of Wartime Destruction: Conley (1999) Standard Errors

	Socio-Economic Index		Log of Property Value		Log of Property Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Destruction in Own Area	−0.051***	−0.042***	−0.174***	−0.157***	−0.112***	−0.097***
	(0.007)	(0.006)	(0.020)	(0.020)	(0.016)	(0.016)
Destruction in 100m Buffer		−0.030**		−0.078*		−0.069*
		(0.012)		(0.043)		(0.037)
Destruction in 200m Buffer		−0.026		−0.129**		−0.110**
		(0.016)		(0.056)		(0.045)
Destruction in 300m Buffer		−0.026		−0.093		−0.100**
		(0.018)		(0.059)		(0.050)
Destruction in 400m Buffer		0.004		−0.084		−0.063
		(0.022)		(0.070)		(0.057)
Destruction in 500m Buffer		0.001		−0.011		−0.024
		(0.021)		(0.071)		(0.066)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8912	8909	8111	8108	8111	8108
R-squared	0.483	0.485	0.659	0.660	0.800	0.800

Notes: This table re-estimates Table 3 in the main paper using Conley (1999) standard errors with a distance cutoff of 1 kilometer instead of standard errors that are clustered at the 1 km hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

Table F.9: The Mechanisms of Wartime Destruction: Conley (1999) Standard Errors

	(1) Fraction Buildings Surviving	(2) Log of Height of Buildings	(3) Fraction Land Area Built-Up	(4) Fraction Council in 2001	(5) Fraction Council in 1981	(6) Log of Empl. Density
Destruction in Own Area	−0.301***	0.079***	−0.041***	0.135***	0.246***	−0.209***
	(0.022)	(0.019)	(0.005)	(0.016)	(0.028)	(0.063)
Destruction in 100m Buffer	0.018	−0.017	0.005	0.028	0.035	−0.004
	(0.039)	(0.038)	(0.012)	(0.030)	(0.054)	(0.126)
Destruction in 200m Buffer	0.052	0.037	0.007	0.031	0.072	0.156
	(0.051)	(0.052)	(0.014)	(0.041)	(0.070)	(0.150)
Destruction in 300m Buffer	−0.044	−0.070	0.020	−0.016	0.035	0.098
	(0.059)	(0.052)	(0.016)	(0.045)	(0.074)	(0.169)
Destruction in 400m Buffer	0.163***	−0.164***	0.015	−0.075	−0.204**	0.032
	(0.062)	(0.064)	(0.017)	(0.049)	(0.083)	(0.193)
Destruction in 500m Buffer	−0.077	0.082	0.017	0.014	0.003	0.236
	(0.063)	(0.065)	(0.019)	(0.053)	(0.085)	(0.199)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8909	8909	8909	8909	6697	8909
R-squared	0.407	0.473	0.464	0.396	0.444	0.479

Notes: This table re-estimates Table F.17 in this Online Appendix using Conley (1999) standard errors with a distance cutoff of 1 kilometer instead of standard errors that are clustered at the 1 km hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

F4 2011 Population Census

In our baseline specification in Section 4 of the paper, we report results using the 2001 population census, because it is the first census after the Second World War to report representative data on socioeconomic composition at the Output Area level, and it plausibly allows us to capture the long-run adjustment of spatial sorting to the shock of wartime destruction. In this section of the Online Appendix, we confirm that our results are capturing long-run effects, by showing that we find a similar pattern of results using the 2011 population census. The Output Areas used to report the results of the 2011 census are similar to those in the 2001 census but not identical, resulting in somewhat different numbers of observations compared to the results for 2001.

Table F.10 re-estimates Table 2 in the paper for the direct effects of wartime destruction; Table F.11 re-estimates Table 3 in the paper for the spillover effects of wartime destruction; Table F.12 re-estimates Table F.17 from Online Appendix F7 for the mechanisms for wartime destruction. In all three tables, we find the same pattern of results as in our baseline specification in Section 4 of the paper. Even after controlling for fixed effects for 1 km hexagons, we find negative and statistically significant direct effects of wartime destruction on post-war socioeconomic outcomes in the own location, and negative and statistically significant effects spillover effects on post-war socioeconomic outcomes in neighboring locations. After controlling for these 1 km fixed effects, we also find direct effects of wartime destruction on the type of buildings, but no evidence of spillover effects of wartime destruction on the type of buildings.

Therefore, we find similar results using either the 2001 or 2011 population census, consistent with our findings capturing the long-run adjustment of spatial sorting to wartime destruction.

Table F.10: The Direct Effect of Wartime Destruction: 2011 Population Census

	(1) Fraction High Status	(2) Fraction Middle Status	(3) Fraction Low Status	(4) Socio-Economic Index	(5) Log of Property Value	(6) Log of Property Value
Destruction in Own Area	-0.040^{***}	-0.030^{***}	0.070^{***}	-0.056^{***}	-0.169^{***}	-0.106^{***}
	(0.007)	(0.005)	(0.010)	(0.008)	(0.021)	(0.018)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	9402	9402	9402	9402	8505	8505
R-squared	0.483	0.290	0.409	0.445	0.653	0.792

Notes: This table re-estimates Table 2 in the main paper using socioeconomic data from the 2011 UK census for the results in columns (1) to (4). Observations differ from the table in the main paper across all columns as the Output Areas in the 2011 census differ somewhat from those used in 2001. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

Table F.11: The Spillover Effect of Wartime Destruction: 2011 Population Census

	Socio-Economic Index		Log of Property Value		Log of Property Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Destruction in Own Area	-0.056^{***}	-0.046^{***}	-0.169^{***}	-0.147^{***}	-0.106^{***}	-0.088^{***}
	(0.008)	(0.008)	(0.021)	(0.021)	(0.018)	(0.017)
Destruction in 100m Buffer		-0.034^{**}		-0.091^{**}		-0.076^{**}
		(0.015)		(0.042)		(0.036)
Destruction in 200m Buffer		-0.030		-0.135^{**}		-0.101^{**}
		(0.019)		(0.058)		(0.048)
Destruction in 300m Buffer		-0.009		-0.067		-0.068
		(0.022)		(0.062)		(0.053)
Destruction in 400m Buffer		0.006		-0.068		-0.041
		(0.025)		(0.073)		(0.061)
Destruction in 500m Buffer		0.011		0.003		-0.005
		(0.027)		(0.073)		(0.068)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	9402	9399	8505	8502	8505	8502
R-squared	0.445	0.446	0.653	0.654	0.792	0.793

Notes: This table re-estimates Table 3 in the main paper using socioeconomic data from the 2011 UK census for the results in columns (1) and (2). Observations differ from the table in the main paper across all columns as the Output Areas in the 2011 census differ somewhat from those used in 2001. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

Table F.12: The Mechanisms of Wartime Destruction: 2011 Population Census

	(1) Fraction Buildings Surviving	(2) Log of Height of Buildings	(3) Fraction Land Area Built-Up	(4) Fraction Council in 2011	(5) Fraction Council in 1981	(6) Log of Empl. Density
Destruction in Own Area	-0.295^{***}	0.077^{***}	-0.039^{***}	0.120^{***}	0.246^{***}	-0.227^{***}
	(0.022)	(0.020)	(0.006)	(0.015)	(0.030)	(0.053)
Destruction in 100m Buffer	0.032	-0.033	-0.010	0.029	0.035	-0.004
	(0.041)	(0.039)	(0.013)	(0.033)	(0.059)	(0.105)
Destruction in 200m Buffer	0.041	0.040	0.008	0.017	0.072	0.126
	(0.052)	(0.055)	(0.014)	(0.041)	(0.077)	(0.130)
Destruction in 300m Buffer	-0.034	-0.043	0.009	-0.028	0.035	0.136
	(0.057)	(0.055)	(0.017)	(0.043)	(0.080)	(0.158)
Destruction in 400m Buffer	0.143^{**}	-0.195^{***}	0.012	-0.051	-0.204^{**}	-0.105
	(0.064)	(0.070)	(0.019)	(0.047)	(0.088)	(0.177)
Destruction in 500m Buffer	-0.072	0.128^{*}	0.001	0.025	0.003	0.280
	(0.069)	(0.076)	(0.021)	(0.052)	(0.101)	(0.193)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	9399	9399	9399	9399	6697	9399
R-squared	0.412	0.467	0.441	0.303	0.444	0.465

Notes: This table re-estimates Table F.17 in this Online Appendix using socioeconomic data from the 2011 UK census in all columns other than column (5), which just reports the results from the 1981 census already in Table F.17 for completeness. Observations differ from the table in the main paper in all columns other than column (5) as the Output Areas in the 2011 census differ somewhat from those used in 2001. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

F5 Post-War Property Prices for Sub-Periods

In our baseline specification in Section 4 of the paper, we report results for post-war property prices for our entire sample period from 1995-2020. In this section of the Online Appendix, we report estimates in which we break out these post-war property prices into sub-periods, to demonstrate that our results are capturing persistent long-run impacts.

Table F.13 re-estimates the property prices specifications from columns (5) and (6) of Table 2 in the main paper for sub-periods. Table F.14 re-estimates the property price specifications from columns (4) and (6) of Table 3 in the main paper for sub-periods. Both tables repeat the specifications from the main paper for the entire 1995 to 2020 period in columns (1) and (4) for ease of reference. In both tables, we find a similar pattern of estimated coefficients across each of the sub-periods, which provides further support for the idea that our estimates are capturing the long-run adjustment of patterns of spatial sorting to wartime destruction.

Table F.13: The Direct Effect of Wartime Destruction: Estimates for Sub-Periods

	(1) Log of Property Value 95-20	(2) Log of Property Value 95-07	(3) Log of Property Value 08-20	(4) Log of Property Value 95-20	(5) Log of Property Value 95-07	(6) Log of Property Value 08-20
Destruction in Own Area	-0.174^{***}	-0.150^{***}	-0.137^{***}	-0.112^{***}	-0.097^{***}	-0.086^{***}
	(0.022)	(0.023)	(0.025)	(0.018)	(0.020)	(0.020)
Land Registry Controls	No	No	No	Yes	Yes	Yes
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8111	7289	6393	8111	7289	6393
R-squared	0.659	0.667	0.695	0.800	0.805	0.836

Notes: This table re-estimates the specification from columns (5) and (6) of Table 2 in the main paper for different time periods. Columns (1) and (4) of the table repeat the specification for the full 1995 to 2020 time period from the main paper for ease of reference. As in the main table, we compute the logarithm of the average residential property value only for Output Areas with at least 25 transactions, which results in somewhat different numbers of observations for the early and late sub-period. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

Table F.14: The Spillover Effect of Wartime Destruction: Estimates for Sub-Periods

	(1) Log of Property Value 95-20	(2) Log of Property Value 95-07	(3) Log of Property Value 08-20	(4) Log of Property Value 95-20	(5) Log of Property Value 95-07	(6) Log of Property Value 08-20
Destruction in Own Area	-0.157^{***}	-0.130^{***}	-0.135^{***}	-0.097^{***}	-0.080^{***}	-0.086^{***}
	(0.022)	(0.024)	(0.025)	(0.017)	(0.020)	(0.019)
Destruction in 100m Buffer	-0.078^{*}	-0.093^{*}	-0.029	-0.069^{*}	-0.085^{**}	-0.019
	(0.045)	(0.048)	(0.051)	(0.038)	(0.042)	(0.043)
Destruction in 200m Buffer	-0.129^{**}	-0.118^{*}	-0.077	-0.110^{**}	-0.096^{*}	-0.063
	(0.059)	(0.060)	(0.065)	(0.047)	(0.049)	(0.050)
Destruction in 300m Buffer	-0.093	-0.091	-0.067	-0.100^{*}	-0.118^{**}	-0.094
	(0.064)	(0.066)	(0.072)	(0.054)	(0.056)	(0.060)
Destruction in 400m Buffer	-0.084	-0.133^{*}	-0.162^{*}	-0.063	-0.091	-0.088
	(0.076)	(0.074)	(0.084)	(0.061)	(0.059)	(0.064)
Destruction in 500m Buffer	-0.011	0.023	0.002	-0.024	-0.005	-0.048
	(0.076)	(0.073)	(0.080)	(0.070)	(0.067)	(0.070)
Land Registry Controls	No	No	No	Yes	Yes	Yes
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8108	7286	6392	8108	7286	6392
R-squared	0.660	0.668	0.696	0.800	0.806	0.836

Notes: This table re-estimates the specification in columns (4) and (6) of Table 3 in the main paper for different time periods. Columns (1) and (4) of the table repeat the specification for the full 1995 to 2020 time period from the main paper for ease of reference. As in the main table, we compute the logarithm of the average residential property value only for Output Areas with at least 25 transactions, which results in somewhat different numbers of observations for the early and late sub-period. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

F6 Excluding Central Boroughs

In our baseline specification in Section 4 of the paper, we report estimation results using Output Areas from all 29 boroughs in the LCC area. In this section of the Online Appendix, we report a robustness test excluding Output Areas from the City of London and the City of Westminster, as the two most central and non-residential boroughs of the LCC.

Table F.15 re-estimates Table 2 in the main paper for the direct effects of wartime destruction, while Table F.16 re-estimates Table 3 in the main paper for the spillover effects of wartime destruction. In both tables, we find the same pattern of results as in our baseline specification including all boroughs in Section 4 of the paper. Even after controlling for fixed effects for 1 km hexagons, we find negative and statistically significant direct effects of wartime destruction on post-war socioeconomic outcomes in the own location, and negative and statistically significant effects spillover effects on post-war socioeconomic outcomes in neighboring locations.

Therefore, we find a similar pattern of results when we exclude Output Areas from the City of London and the City of Westminster, which is consistent with our estimates capturing effects through residential activity.

Table F.15: The Direct Effect of Wartime Destruction: Excluding Central Boroughs

	(1)	(2)	(3)	(4)	(5)	(6)
	Excluding the City of London			Also Excluding Westminster		
	Socio-Economic Index	Log of Property Value	Log of Property Value	Socio-Economic Index	Log of Property Value	Log of Property Value
Destruction in Own Area	−0.053***	−0.175***	−0.113***	−0.051***	−0.165***	−0.102***
	(0.007)	(0.022)	(0.018)	(0.007)	(0.022)	(0.018)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8876	8076	8076	8148	7388	7388
R-squared	0.481	0.659	0.800	0.479	0.656	0.800

Notes: This table re-estimates the specifications in columns (4), (5), and (6) of Table 2 in the main paper. The first three columns of this table exclude Output Areas in the City of London while the final three columns also exclude Output Areas in Westminster. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

Table F.16: The Spillover Effect of Wartime Destruction: Excluding Central Boroughs

	(1)	(2)	(3)	(4)	(5)	(6)
	Excluding the City of London			Also Excluding Westminster		
	Socio-Economic Index	Log of Property Value	Log of Property Value	Socio-Economic Index	Log of Property Value	Log of Property Value
Destruction in Own Area	−0.045***	−0.158***	−0.098***	−0.043***	−0.148***	−0.088***
	(0.007)	(0.022)	(0.018)	(0.007)	(0.022)	(0.017)
Destruction in 100m Buffer	−0.034***	−0.087*	−0.078**	−0.035***	−0.080*	−0.070*
	(0.013)	(0.045)	(0.038)	(0.013)	(0.045)	(0.037)
Destruction in 200m Buffer	−0.026	−0.117**	−0.099**	−0.021	−0.104*	−0.089*
	(0.018)	(0.059)	(0.048)	(0.019)	(0.059)	(0.047)
Destruction in 300m Buffer	−0.033*	−0.097	−0.103*	−0.037*	−0.068	−0.074
	(0.019)	(0.065)	(0.055)	(0.019)	(0.064)	(0.054)
Destruction in 400m Buffer	0.000	−0.086	−0.064	−0.003	−0.051	−0.040
	(0.023)	(0.077)	(0.062)	(0.023)	(0.078)	(0.062)
Destruction in 500m Buffer	−0.004	−0.007	−0.022	−0.008	−0.046	−0.056
	(0.023)	(0.078)	(0.071)	(0.023)	(0.076)	(0.070)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8873	8073	8073	8145	7385	7385
R-squared	0.482	0.660	0.801	0.481	0.656	0.801

Notes: This table re-estimates the specifications in Table F.15 in this Online Appendix also including destruction in our five buffers. The first three columns of this table exclude Output Areas in the City of London while the final three columns also exclude Output Areas in Westminster. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

F7 Mechanisms for Second World War Destruction

In this section of the Online Appendix, we provide further evidence on the mechanisms through which the direct and spillover effects of wartime destruction occur.

F7.1 Building Types

First, we show that the direct effects of wartime destruction operate through changes in the types of buildings, whereas the spillover effects of wartime destruction do not.

We re-estimate our spillover regression from equation (4) in the paper using a number of different measures of building types. Column (1) of Table F.17 uses the share of buildings within the pre-war building footprint; Column (2) considers the height of modern buildings; Column (3) examines the share of the land area that is built up. As reported in the first row, we find substantial direct effects of wartime destruction on each of these building outcomes. Wartime destruction substantially reduces the probability that buildings lie within the pre-war building footprint. Wartime destruction also increases the height of buildings by around 7.9 percent, and reduces the share of the land area that is built up by 4.1 percent, which is in line with post-war architectural trends towards high-rise tower blocks surrounded by open areas. As reported in the second to sixth rows, we find no systematic evidence of spillover effects of wartime destruction on these building outcomes. Therefore, destruction in neighboring locations does not change the types of buildings in the own location.

In Table F.9 of Online Appendix F3 we show that these results are robust to using HAC standard errors. In Table F.12 of Online Appendix F4, we show that they are also robust to using data from the 2011 census instead of the 2001 census.

Taken together, these results suggest that the spillover effects are not capturing the demolition of undamaged buildings in neighboring areas in response to wartime destruction. Instead these results are consistent with the idea that the spillover effects reflect changes in surrounding neighborhood characteristics.

F7.2 Share of Households Living in Council Housing

Second, we show that the direct effects of wartime destruction involve significant changes in the share of households living in council housing, whereas the spillover effects do not.

Column (4) of Table F.17 re-estimates our spillovers specification using the share of households living in council housing in the 2001 population census. We find that areas that experienced more own destruction have higher council housing shares in 2001. This pattern is consistent with the space created by wartime destruction being used to accommodate the large post-war expansion in council housing. As shown in Figure G.24 in Online Appendix G8, over 80 percent of all

Table F.17: Mechanisms for Wartime Destruction

	(1) Fraction Buildings Surviving	(2) Log of Height of Buildings	(3) Fraction Land Area Built-Up	(4) Fraction Council in 2001	(5) Fraction Council in 1981	(6) Log of Empl. Density
Destruction in Own Area	-0.301^{***}	0.079^{***}	-0.041^{***}	0.135^{***}	0.246^{***}	-0.209^{***}
	(0.023)	(0.019)	(0.005)	(0.017)	(0.030)	(0.064)
Destruction in 100m Buffer	0.018	−0.017	0.005	0.028	0.035	−0.004
	(0.043)	(0.039)	(0.012)	(0.034)	(0.059)	(0.129)
Destruction in 200m Buffer	0.052	0.037	0.007	0.031	0.072	0.156
	(0.053)	(0.053)	(0.015)	(0.047)	(0.077)	(0.160)
Destruction in 300m Buffer	−0.044	−0.070	0.020	−0.016	0.035	0.098
	(0.059)	(0.052)	(0.017)	(0.045)	(0.080)	(0.172)
Destruction in 400m Buffer	0.163^{**}	-0.164^{**}	0.015	−0.075	-0.204^{**}	0.032
	(0.065)	(0.066)	(0.018)	(0.050)	(0.088)	(0.202)
Destruction in 500m Buffer	−0.077	0.082	0.017	0.014	0.003	0.236
	(0.068)	(0.071)	(0.020)	(0.057)	(0.101)	(0.210)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8909	8909	8909	8909	6697	8909
R-squared	0.407	0.473	0.464	0.396	0.444	0.479

Notes: Each column reports the results of a separate regression. The unit of observation are Output Areas from the 2001 UK Census except column (5), which uses enumeration districts from the 1981 UK Census. The dependent variable in column (1) is the fraction of pre-war buildings that exist in the same footprint in 2014; in column (2) the logarithm of the average building height; in column (3) the fraction of total area of an Output Area that is built up; in columns (4) and (5) the fraction of households that reside in council housing in 2001 and 1981 respectively and in column (6) the logarithm of employment density. The explanatory variables in all columns are the fraction of the built-up area seriously damaged within the unit of observation as well as five buffers of 100 meter width around each unit of observation. Observations differ in column (5) due the different spatial units used in the 1981 census. All regressions include fixed effects for 1 km hexagons. Standard errors are clustered at the 1 km hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

housing units constructed in the LCC area from 1945-80 were council housing. In 1980, Margaret Thatcher's Housing Act gave council tenants the "right to buy" their properties at considerable discounts on the market price, which led to a large large-scale transfer to private ownership. To capture the impact of wartime destruction on council housing before this large-scale transfer, Column (5) repeats the same regression using data from the 1981 census, and finds an even larger effect on the share of households living in council housing.[6] In contrast, in both Columns (4) and (5), we find no systematic evidence of spillover effects of wartime destruction on the share of households living in council housing.

In Table F.9 of Online Appendix F3 we show that these results are robust to using HAC standard errors. In Table F.12 of Online Appendix F4, we show that they are also robust to using data

[6] Although "right to buy" sales transferred ownership to the private sector, demolitions were relatively rare, with only 140,000 council units demolished in England since 1997, when official data on demolitions were first published. The 1988 Housing Act reduced the barriers for local authorities to transfer council housing units to housing associations and led to a number of such transfers. To control for this policy change, we re-estimated Columns (4) and (5) in Table F.17 counting housing association tenants as council tenants, and find similar results.

from the 2011 census instead of the 2001 census.

Overall, these results provide further evidence that spillover effects are not occurring through the demolition of undamaged buildings in neighboring areas in response to wartime destruction and are instead operating through changes in surrounding neighborhood characteristics.

F7.3 Commercial and Residential Activity

Third, we provide further evidence that the impact of wartime destruction is operating through changes in the residential attractiveness of locations.

Employment Density Column (6) of Table F.17 re-estimates our spillovers specification using using the log of employment density (employment by workplace per land area). We find a negative and statistically significant direct effect of wartime destruction, with no evidence of statistically significant spillover effects.

In Table F.9 of Online Appendix F3 we show that these results are robust to using HAC standard errors. In Table F.12 of Online Appendix F4, we show that they are also robust to using data from the 2011 census instead of the 2001 census.

Therefore these results suggest that if anything wartime destruction shifted economic activity towards residential use, which is consistent with our focus in the model on residential choices.

Residential and Commercial Destruction We next use residential and commercial destruction as separate sources of variation to provide further evidence that wartime destruction is operating through changes in the residential attractiveness of locations.

Randomness of Residential and Commercial Destruction: Table F.18 re-estimates our randomization specification from equation (2) in the main paper. Panel A replicates our results for overall destruction from Table 1 of the main paper. Panel B reports results for residential destruction. Panel C reports results for commercial destruction.

We find the same pattern of results whether we use overall damage, residential damage or commercial damage. In the specification with no fixed effects in the top row of each panel, we find a correlation between pre-war socioeconomic outcomes and subsequent wartime destruction. Output Areas that had larger pre-war shares of the population with lower socioeconomic status and lower pre-war property values experienced more destruction during the Second World War. Once we include fixed effects for 4 km hexagons in the middle row of each panel, much of this correlation goes away. Nevertheless, 4 km hexagons still cover a relatively large geographical area, and are still likely affected by the West-East gradients noted above. Once we include fixed effects for 1 km hexagons in the bottom row of each panel, the coefficients fall close to zero and are entirely statistically insignificant.

Table F.18: Randomness of Wartime Destruction (Overall, Residential and Commercial)

Fixed Effects	(1) Fraction High Status	(2) Fraction Middle Status	(3) Fraction Low Status	(4) Socio-Economic Index	(5) Log of Property Value	(6) Log of Property Value
			Panel A - All Destruction			
None	-0.235^{***}	0.039^{*}	0.196^{***}	-0.215^{***}	-0.478^{***}	-0.517^{***}
	(0.031)	(0.023)	(0.025)	(0.026)	(0.056)	(0.061)
4 km Hexagons	-0.061^{***}	0.020	0.042^{**}	-0.051^{***}	-0.099^{***}	-0.124^{***}
	(0.020)	(0.018)	(0.017)	(0.016)	(0.037)	(0.036)
1 km Hexagons	−0.007	−0.004	0.011	−0.009	−0.003	−0.026
	(0.014)	(0.013)	(0.012)	(0.012)	(0.029)	(0.028)
			Panel B - Residential Destruction			
None	-0.207^{***}	0.023	0.184^{***}	-0.195^{***}	-0.456^{***}	-0.497^{***}
	(0.028)	(0.021)	(0.023)	(0.023)	(0.052)	(0.057)
4 km Hexagons	-0.051^{***}	0.005	0.046^{***}	-0.049^{***}	-0.096^{***}	-0.121^{***}
	(0.018)	(0.016)	(0.015)	(0.014)	(0.034)	(0.033)
1 km Hexagons	−0.003	−0.015	0.017	−0.010	−0.003	−0.026
	(0.013)	(0.012)	(0.011)	(0.010)	(0.027)	(0.025)
			Panel C - Commercial Destruction			
None	-0.170^{***}	0.033^{**}	0.136^{***}	-0.152^{***}	-0.313^{***}	-0.317^{***}
	(0.023)	(0.017)	(0.019)	(0.019)	(0.041)	(0.043)
4 km Hexagons	-0.046^{***}	0.018	0.028^{**}	-0.037^{***}	-0.083^{***}	-0.088^{***}
	(0.015)	(0.013)	(0.013)	(0.012)	(0.029)	(0.028)
1 km Hexagons	−0.016	0.011	0.006	−0.011	−0.029	-0.036^{*}
	(0.011)	(0.010)	(0.009)	(0.008)	(0.022)	(0.021)

Notes: This table re-estimates the specifications in Table 1 in the main paper disaggregating destruction into residential and commercial bomb damage. Panel A repeats the specifications for all war-time destruction from the main paper for ease of reference, Panels B and C breaks down overall destruction into residential and commercial destruction respectively. We refer to the non-residential built-up area for simplicity as the commercial built-up area. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

Therefore, once we focus on variation within narrow geographical grid cells, wartime damage is again entirely unrelated to pre-war socioeconomic status and property values. This pattern of results is consistent with the primitive bomb-aiming technology and night-time bombing, which precluded the precise targeting of locations. Overall, these results provide further support for the idea that wartime destruction within narrow geographical areas is as good as randomly assigned and provides an exogenous source of variation.

Table F.19: The Direct Effect of Wartime Destruction (Overall, Residential and Commercial)

	(1) Fraction High Status	(2) Fraction Middle Status	(3) Fraction Low Status	(4) Socio-Economic Index	(5) Log of Property Value	(6) Log of Property Value
	Panel A - All Destruction					
All Destruction	-0.039^{***}	-0.023^{***}	0.062^{***}	-0.051^{***}	-0.174^{***}	-0.112^{***}
	(0.006)	(0.005)	(0.009)	(0.007)	(0.022)	(0.018)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8912	8912	8912	8912	8111	8111
R-squared	0.505	0.280	0.439	0.483	0.659	0.800
	Panel B - Residential and Commercial Destruction					
Residential Destruction	-0.038^{***}	-0.024^{***}	0.062^{***}	-0.050^{***}	-0.185^{***}	-0.108^{***}
	(0.006)	(0.004)	(0.008)	(0.007)	(0.025)	(0.020)
Commercial Destruction	-0.003	0.001	0.003	-0.003	0.006	-0.004
	(0.005)	(0.004)	(0.006)	(0.005)	(0.020)	(0.017)
Hexagon Fixed Effects	1 km	1 km	1 km	1 km	1 km	1 km
Observations	8511	8511	8511	8511	7793	7793
R-squared	0.518	0.290	0.455	0.498	0.665	0.804

Notes: This table re-estimates the direct impact specifications in Table 2 in the main paper. Panel A repeats the specifications in Table 2 in the main paper for ease of comparison. Panel B breaks down destruction into destruction of residential and commercial buildings. Numbers of observations vary slightly across specifications due to the availability of modern housing transaction prices and whether Output Areas had both commercial and residential built-up area pre-war. Standard errors are clustered at the 1 km hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

Direct Effects of Residential and Commercial Destruction: Table F.19 re-estimates our direct effects specification from equation (3) in the paper. Panel A replicates our results for overall destruction from Table 2 of the paper. Panel B reports results for residential destruction. Panel C reports results for commercial destruction. In Panel B, we show that the estimated effects of wartime destruction on post-war outcomes are entirely driven by damage to residential buildings. When we include both residential and commercial damage separately, we find coefficients on residential damage that are statistically significant and close in magnitude to those for overall damage above. In contrast, we find coefficients on commercial damage in Panel C that are entirely statistically insignificant and close to zero in magnitude. This pattern of results is consistent with the mechanism in our model, in which wartime destruction changes the residential amenities from in a location, and hence in turn leads to a change in socioeconomic composition.

Spillover Effects of Residential and Commercial Destruction: Table F.20 re-estimates our spillover effects specification from equation (4) in the main paper. Columns (1)-(2) measure post-war outcomes using our index of socioeconomic composition, while Columns (3)-(4) use property values controlling for observed property characteristics. Columns (1) and (3) use overall damage, while

Columns (2) and (4) include separate measures of residential and commercial damage. To conserve space, we display the estimated coefficients on residential and commercial damage next to one another underneath the labels for Columns (2) and (4), even though these coefficients are estimated from a single regression. We find that our estimated spillover effects are entirely driven by destruction of the residential built-up area, with estimated coefficients for destruction of the commercial built-up area that are close to zero and almost always statistically insignificant. This pattern of results again provides further support for the mechanism in our model, in which these spillover effects operate through changes in the residential attractiveness of locations.

Table F.20: The Spillover Effect of Wartime Destruction (Overall, Residential and Commercial)

	Socio-Economic Index			Log of Property Value		
	(1)	(2)		(3)	(4)	
Destruction in Own Area	-0.042^{***}	-0.045^{***}	0.000	-0.097^{***}	-0.095^{***}	-0.001
	(0.007)	(0.007)	(0.005)	(0.017)	(0.021)	(0.016)
Destruction in 100m Buffer	-0.030^{**}	-0.025^{*}	-0.008	-0.069^{*}	-0.090^{**}	0.018
	(0.013)	(0.013)	(0.010)	(0.038)	(0.039)	(0.031)
Destruction in 200m Buffer	-0.026	-0.024	-0.003	-0.110^{**}	-0.118^{**}	-0.004
	(0.018)	(0.019)	(0.013)	(0.047)	(0.055)	(0.043)
Destruction in 300m Buffer	-0.026	-0.029	0.000	-0.100^{*}	-0.090	0.010
	(0.019)	(0.018)	(0.014)	(0.054)	(0.055)	(0.047)
Destruction in 400m Buffer	0.004	-0.020	0.026^{*}	-0.063	-0.098^{*}	0.034
	(0.023)	(0.019)	(0.014)	(0.061)	(0.054)	(0.046)
Destruction in 500m Buffer	0.001	-0.016	0.014	-0.024	-0.026	-0.001
	(0.023)	(0.021)	(0.016)	(0.070)	(0.063)	(0.048)
Damage Type	Total	Res	Com	Total	Res	Com
Hexagon Fixed Effects	1 km	1 km		1 km	1 km	
Observations	8909	8496		8108	7779	
R-squared	0.485	0.500		0.800	0.805	

Notes: This table re-estimates the spillover specifications in Table 3 in the main paper for total, residential and commercial destruction. The dependent variable in columns (1) and (2) is an index of socioeconomic status and in columns (3) and (4) the logarithm of the average property transaction price conditional on a set of property characteristics. Columns (2) and (4) break down overall damage into damage to residential and commercial buildings. The left hand side of the column reports the estimates for residential destruction (Res) while the right hand side reports the estimates for commercial destruction (Com). Numbers of observations vary slightly across specifications due to the availability of modern housing transaction prices and whether Output Areas have both commercial and residential built-up area before the Second World War. Standard errors are clustered at the 1 km hexagon level. * denotes significance at 10% level; ** denotes significance at 5% level; *** denotes significance at 1% level.

G Data Appendix

This section of the online appendix reports additional information about the data sources and definitions, supplementing the discussion in Section 3 of the paper. Section G1 discusses the spatial units used for our empirical analysis. Section G2 provides further detail on the construction of our wartime destruction data.

Section G3 considers our pre-war data on socioeconomic status. Section G4 examines our pre-war population data. Section G5 presents additional information about our pre-war property values data. Section G6 discusses our post-war population census data. Section G7 summarizes our post-war property values data.

Section G8 introduces our post-war data on private and council housing construction. Section G9 considers our post-war data on the built-up area and building height. Section G10 presents additional information on our pre-war commuting data.

G1 Spatial Units

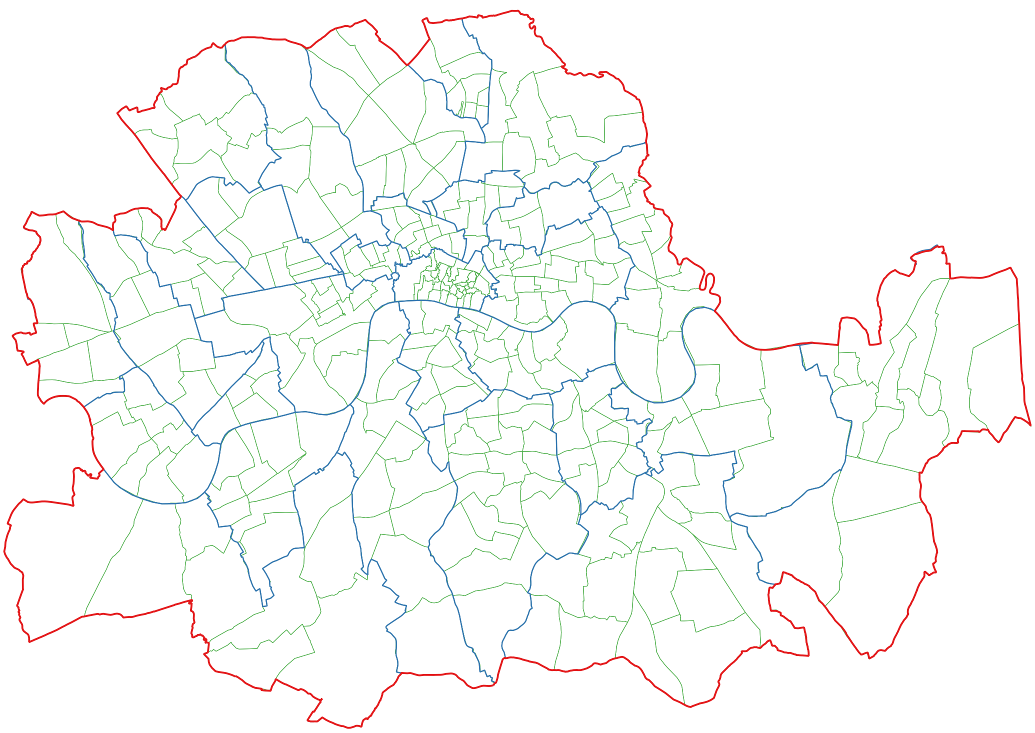
Geographical Area Our geographic area of analysis is the area administered by London County Council (LCC), which was the principal local government body for London from 1889-1965. We focus on the LCC area, because it is the largest geographical area that is fully covered by the bomb damage maps, and comprises the entire of the Central Area and the inner suburbs of London. At the time of the 1931 census, the closest census to the beginning of the Second World War in 1939, the LCC area covered 312 km^2 with a population of 4.4 million. The LCC area was subdivided into 29 Metropolitan Boroughs for administrative purposes, as listed in Table G.1 below. Each of these Metropolitan Boroughs was further split into wards, as shown in the map in Figure G.1. In 1965 the 29 historical boroughs of the LCC area were amalgamated into 13 modern boroughs and become part of the Greater London Council (GLC), which also included areas outside the LCC area. The GLC was dissolved by Margaret Thatcher's administration in 1986 and was replaced in 2000 by the Greater London Authority (GLA).

Table G.1: Metropolitan Boroughs of the London County Council (LCC) Area

	Borough Name	Postal Area		Borough Name	Postal Area
1	Battersea	SW	16	Kensington	W
2	Bermondsey	SE	17	Lambeth	SW
3	Bethnal Green	E	18	Lewisham	SE
4	Camberwell	SE	19	Paddington	W
5	Chelsea	SW	20	Poplar	E
6	City of London	EC	21	Shoreditch	EC
7	Deptford	SE	22	Southwark	SE
8	Finsbury	NW	23	St Marylebone	W
9	Fulham	SW	24	St Pancras	N
10	Greenwich	SE	25	Stepney	E
11	Hackney	E	26	Stoke Newington	N
12	Hammersmith	W	27	Wandsworth	SW
13	Hampstead	NW	28	Westminster	WC
14	Holborn	WC	29	Woolwich	SE
15	Islington	N			

Note: 29 Metropolitan Boroughs of the London County Council (LCC) Area and the first two-digits of the historical postal codes; EC is East Central; WC is West Central; E is East; W is West; N is North; S is South; SW is South West; SE is South East; NW is North West; NE is North East.

Figure G.1: 1931 Administrative Boundaries of the London County Council



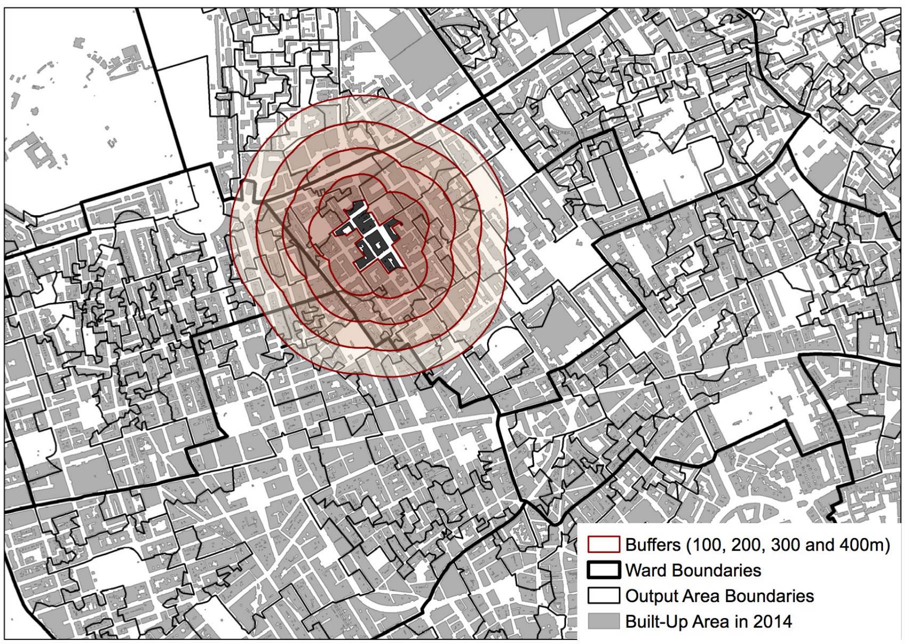
Note: Map of the London County Council (LCC) area, showing its boundary in red, 1931 Metropolitan Boroughs in blue, and 1931 wards in green.

Output Areas Our main units of statistical analysis are Output Areas (OAs), as defined in the 2001 United Kingdom (UK) Census. Output Areas are built from postcode units specifically for statistical purposes and provide the lowest geographical level at which census data are published in the UK. Output Areas were first created as part of the 2001 Census and were targeted to include roughly 125 resident households. Output Areas are a partition of the surface area of the UK. We focus on the 9,041 Output Areas that have the majority of their area inside the LCC boundary.

100m Buffers In our analysis of spillovers between locations, we also create buffer regions around each output area in order to define its "neighbors." We construct four buffers for each output area, 100m in width, such that its surrounding region up to 400m is sorted into bands of increasing distance. As Output Areas are a partition of geographic space, they can include substantial areas that are not built up, including parks and rivers. Therefore, we measure distance to neighbors from the building footprint within each Output Area, rather than the geographical boundary of the Output Area itself. We use the union of the pre-war and post-war building footprints, in order to account for empty areas that existed pre-war but have subsequently been developed, as well as the creation of new parks where there were once buildings.

We create 100m buffers by expanding the footprint of every building within the output area by 100m, 200m, 300m and 400m, and subtracting the shape of the next smallest buffer, making them "hollow." This creates rings around the output areas, with the 400 meter buffer, for example, containing areas between 300 and 400 meters from the output area. The smallest buffer (100m) has the shape of the output area itself subtracted from it, so that there is no overlap between own and neighboring areas. An illustration of an output area and its buffers is shown in Figure G.2. Finally, we require the buffers of each Output Area to lie on the same side of the river Thames as the Output Area itself, by removing any buffers that cross to the other side of the river Thames, which excludes neighbors that are geographically close but not accessible.

Figure G.2: Example of 100 meter Buffers around an Output Area



Note: Example of 100 meter buffers around an output area in the borough of Holborn in Central London.

Fixed Effects and Clustering of Standard Errors To include fixed effects and cluster standard errors in our estimation, we group Output Areas by geographical location. One challenge in doing so is that the larger administrative units of wards vary substantially in their geographical area. As apparent from Figure G.1, the largest wards are hundreds of times bigger than the smallest ones, with the smaller wards generally found in the centre of London, and the larger wards located in the suburbs. To overcome this challenge, we construct two regular grids of larger spatial units, by defining grids of hexagons with sides of length 1 km and 4 km. The size of these hexagons is chosen to roughly mimic the number of wards (1 km hexagons) and boroughs (4 km hexagons). Table G.2 reports summary statistics on these larger spatial units and the number of Output Areas contained within them.

Table G.2: Summary Statistics on Larger Spatial Units

		Output Areas per Spatial Unit				
	Number of Observations	Mean	Median	Min	Max	Standard Deviation
1 km Hexagon	386	23.4	23	2	67	13.4
Wards 1931	316	30.3	23	1	236	27.8
Wards 2001	237	38.1	41	1	54	10.9
4 km Hexagon	35	258.3	257	2	648	194.7
Borough 1931	29	311.8	259	28	1018	211.2
Borough 2001	18	502.3	645	1	946	371.9

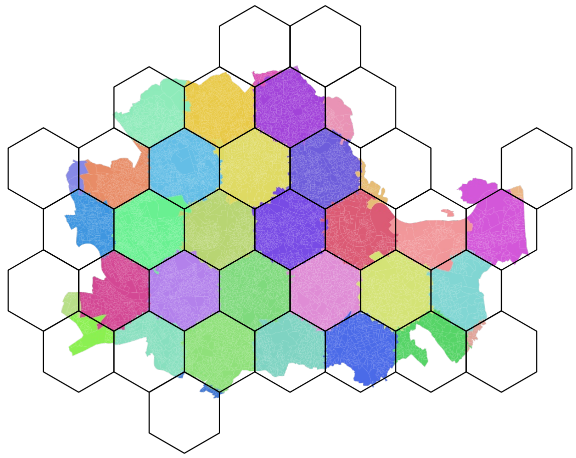
Note: Summary statistics on larger spatial units for both administrative areas and our 1 km and 4 km hexagonal grids, and the number of Output Areas within them for the London County Council (LCC) area. While there are 316 wards in the LCC area in the 1931 Census several wards are very small. For these very small wards no 2001 Output Area lies with the majority of its area in such small wards and these wards are therefore counted as having zero Output Areas. In the City of London, there are also some small 2001 wards that only contain one Output Area.

Figure G.3 overlays these hexagonal grids on top of the Output Areas within the administrative boundaries of the London County Council (LCC) area. Output Areas are colored to indicate those associated with each hexagon[7]. Figure G.3a shows 4 km hexagons, while Figure G.3b shows 1 km hexagons. Figure G.3c zooms in to the hexagons shown in Figures G.3a and G.3b, in order to provide a more detailed picture of the relationship between Output Areas (shown by the faint gray lines) and the different hexagon sizes (shown by the black lines).

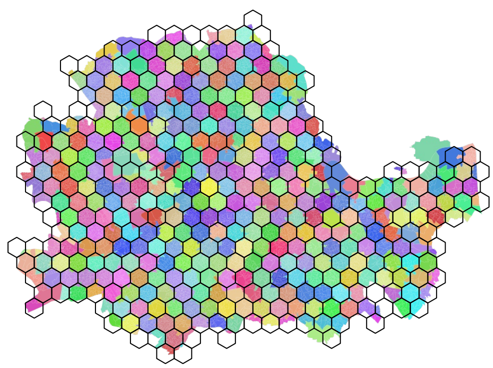
[7]Output areas are allocated to the hexagon which covers the largest share of their area. In the event that there is only one output area associated with a hexagon, it is moved to the closest other hexagon based on distance from its centroid, which can happen on the fringes of the LCC area.

Figure G.3: Hexagonal Grids of Larger Spatial Units for the London County Council (LCC) Area

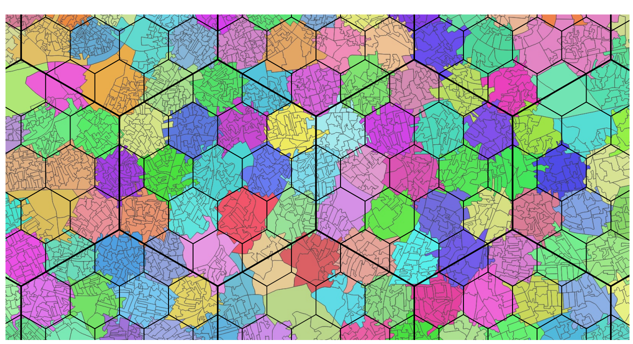
(a) 4 km Hexagons Covering the LCC Area



(b) 1 km Hexagons Covering the LCC Area



(c) Zooming in to the Hexagons Shown in (a) and (b)



G2 Second World War Bomb Damage Data

In this section of the Online Appendix, we provide further information on the construction of our data on Second World War destruction from 1939-45 at the level of individual buildings within the London County Council (LCC) area. Section G2.1 provides additional historical background on the German bombing of London during the Second World War. Section G2.2 discuss the creation of the LCC bomb damage maps. Section G2.3 discusses the georeferencing and geoprocessing of these maps to construct our data on bomb damage to individual buildings.

G2.1 German Bombing of London

Although the Second World War started in September 1939 in Poland, it was not until 10 May 1940 that German forces invaded Belgium and France. Their main armored thrust through the Ardennes cut off and surrounded the British Expeditionary Force (BEF) and led to its evacuation from Dunkirk in late May and early June. As further German armored penetration outflanked the main French forces that were organized around the defensive Maginot Line, the French government signed an armistice with Germany on 22 June 1940.[8]

Following the fall of France, German military planning turned to the invasion of Britain codenamed Operation Sea Lion ("Seelöwe"). In preparation, the German air force ("Luftwaffe") began a series of concentrated aerial attacks in August 1940 (referred to as the Battle of Britain), aimed at the destruction of the Royal Air Force (RAF) and establishing air superiority over Britain.[9] Initially, these attacks were concentrated on RAF airfields and infrastructure. However, on 24 August 1940 night bombers aiming for RAF airfields drifted off course and accidentally destroyed several London homes and killed a number of civilians. After Winston Churchill ordered the immediate bombing of Berlin in response the following night, Adolf Hitler responded with a strategic bombing campaign of British cities concentrated on London.[10]

Although London experienced a few bombing raids by German Zeppelins and light aircraft during the First World War, these raids were small in number, and the extent of destruction was minuscule, because of the primitive aircraft technology at the time.[11] In the face of the dramatic improvements in aircraft and bomb technology in the years leading up to the Second World War, a UK Cabinet Committee in 1937 estimated that an attacker dropping 600 tons of bombs each day could cause 200,000 casualties a week, of which 66,000 would be killed. In reality, levels of destruction and casualties from wartime bombing were far smaller than anticipated. According to the official post-war report of the UK government, 60,595 civilians were killed in total from

[8] For the history of the Fall of France, see Jackson (2004).

[9] For more details on Sea Lion and the Battle of Britain, see Mckinstry (2014) and Holland (2012), respectively.

[10] In the face of the continuing resistance of the RAF and the shift towards a strategic bombing campaign, Operation Sea Lion was postponed indefinitely on 17 September 1940.

[11] See for example White (2014).

enemy action over the entire course of the war, with roughly half of these deaths occurring in the London Civil Defense region (which included the LLC area and the surrounding region), compared to an estimated population in the LCC area and the United Kingdom in 1931 of 4.4 and 46 million, respectively.

Although less than anticipated before the war, the extent of destruction experienced was nevertheless substantial. The intense bombing of London (the "Blitz") lasted from 7 September 1940 to 21 May 1941. Starting on 7 September 1940, London was bombed for 57 consecutive nights. Between the months of September and November alone, almost 30,000 bombs were dropped. These included both high-explosive bombs (which directly damaged buildings) and incendiaries (which created fires that damaged buildings). Heavy day-time aircraft losses led to a concentration on night-bombing from October 1940 onwards. Of the 127 major Luftwaffe attacks on British cities that involved more than 100 tons of bombs during this period, 71 were targeted on London, with an estimated total of 18,291 tons of bombs dropped on London, making up around 60 percent of the total for all British cities, as summarized in Table G.3 below.

Table G.3: Major German Bombing Raids from 7 September 1940 to 21 May 1941

Target	Number of Attacks	Tons of Bombs Dropped
London	71	18,291
Merseyside	8	1,957
Birmingham	8	1,852
Plymouth	8	1,228
Bristol	6	919
Clydeside	5	1,329
Southampton	4	647
Portsmouth	3	687
Hull	3	593
Manchester	3	578
Coventry	2	818
Belfast	2	440
Sheffield	1	355
Tyneside	1	152
Nottingham	1	137
Cardiff	1	115

Note: Data from Appendix B, major Luftwaffe attacks involving more than 100 tons of bombs , 7 Sept 1940 - 16 May 1941, Ray (2004).

With the start of preparations for the German invasion of the USSR in June 1941 ("Barbarossa"), conventional air attacks on London were greatly reduced. However, as the Second World War progressed and in part responding to the Allied bombing of German cities, the German airforce and army ("Wehrmacht") developed long-range retaliatory weapons ("Vergeltungswaffen") for strategic bombing purposes. The first of these weapons, the V-1, was a pulsejet predecessor of the cruise missile (commonly referred to as a "Buzz bomb" or "Doodlebug" because of its characteristic noise). The second, the V-2, was the first long-range ballistic missile.[12]

Following the Allied landings in Normandy on 6 June 1944, the first V-1 was launched at London on 13 June 1944. The V-1 had a range of 250km, carried an 850 kg warhead, and flew at a speed of 640 km/hr. It was launched from an fixed starting ramp and guided by a gyrocompass that controlled altitude and direction. An odometer driven by a vane anemometer on the V-1's nose determined when the target had been reached. Before launch, the counter was set to a value that would reach zero upon arrival at the target (Tower Bridge for London) in the prevailing wind conditions. When the count reached zero, two detonating bots were fired, which put the V-1 into a steep dive on to the target. Overall, only about 25 per cent of the V-1s are estimated to have hit their target area, with the majority being lost because of a combination of defensive measures, mechanical unreliability or guidance errors.[13] As a result of these factors, and fluctuations in prevailing winds and atmospheric conditions, V-1 impacts are scattered throughout London and surrounding South-East England, as shown below.[14] Using data for an area of 144 square kilometers in South London, Clarke (1946) is unable to reject the null hypothesis that V-1 missiles fell randomly according to an independent Poisson distribution.

Development of the V-2 lagged somewhat behind that of the V-1 and it was not until September 1944 that the first missile was launched against London. The V-2 had a range of 320 km, carried a 1,000 kg warhead, travelled at up to 5 times the speed of sound during its powered phase, and fell to earth from an altitude of 100km at nearly 3 times the speed of sound. Power was supplied by a liquid-propellant rocket engine and the V-2 was launched from a mobile launcher. Distance and angle to the target (again Tower Bridge for London) were set at the launch site. The guidance system consisted of two gyroscopes (a horizontal and a vertical) to stabilize the rocket and an accelerometer to control engine cutoff at a specified velocity. Once the engine cut out, the missile continued to follow its ballistic trajectory on to the target. The speed and trajectory of the V-2 made it practically invulnerable to anti-aircraft guns and fighter interception. Nonetheless variation in manufacturing quality and technical malfunctions resulted in considerable inaccuracy.[15] Using data for both V-1 and V-2 missiles, Shaw and Shaw (2019) is again unable to reject the null hypothesis that they fell randomly according to an independent Poisson distribution.

Even individual V-1 and V-2 impacts could create considerable destruction. For example, on

[12]For the history of the development of the V-1 and V-2, see Johnson (1981) and Campbell (2012).

[13]Defensive measures included barrage balloons, anti-aircraft guns South of London, and fighter interception.

[14]British Intelligence leaked false information to the Germans implying that the rockets were overshooting their London targets, which is believed to have shifted V1-impacts towards less-populated areas South of London.

[15]V-2 rockets were produced in the Mittelwerk using forced labor from the Mittelbau-Dora concentration camp, with documented heroic acts of sabotage to manufacturing components.

14th January 1945, twenty houses in South London were demolished by a single missile, with another 50 suffering serious damage. Impact craters could be more than thirty feet wide and damage might extend for up to one quarter of a mile. According to official estimates, some 32,000 V-1s were manufactured. Around 10,492 were launched at London, of which 2,419 reached the target area, killing 6,184 and injuring 17,981. In comparison, approximately 6,000 V-2s were manufactured, of which a little more than half were fired operationally. As a result of the 1,358 V-2 that landed in the London civil defense region (which extended substantially beyond the LCC area), 2,754 civilians were killed and 6,500 injured.

Despite the technological advances of the V-1 and V-2 missiles, most wartime destruction and casualties were caused by bombing from conventional aircraft, as summarized for Britain as a whole in Tables G.4 and G.5 below.

Table G.4: Tons of Bombs dropped by Conventional Aircraft, V1s and V2s in Britain from 1940-5

Date	Bombs (Excluding I.B.s & A.P.s)	Flying-Bombs V1 (War-head)	Long-Range Rockets V2 (War-head)	Total
Sept-Dec 1940	34,970	–	–	34,970
1941	22,176	–	–	22,176
1942	3,039	–	–	3,039
1943	2,232	–	–	2,232
1944	1,960	5,731	390	8,081
Jan-May 1945	16	92	664	772
Total	64,393	5,823	1,054	71,270

Note: Taken from Appendix III of O'Brien (1955), part of the Official History of the Second World War, United Kingdom Civil Series; columns report metric tons of bombs; I.B. refers to incendiary bomb; A.P. refers to armor piercing bomb.

Table G.5: Casualties from the German Bombing of Britain during 1940-5

Weapon	Killed	Seriously Injured	Total
Bombs	51,509	61,423	112,932
Flying Bombs (V1)	6,184	17,981	24,165
Long-range Rockets (V2)	2,754	6,523	9,277
Cross-channel Bombardment	148	25	403
Total	60,595	86,182	146,777

Note: Taken from Appendix II of O'Brien (1955), part of the Official History of the Second World War, United Kingdom Civil Series.

G2.2 LCC Bomb Damage Map

To keep a record of the destruction of the built-up area, London County Council (LCC) used detailed pre-war Ordinance Survey (OS) maps at a scale of 1:2,500. There are 110 of these map sheets for the LCC area, which were recently re-published in Ward (2016). Each map sheet shows the outlines of individual buildings, which were color coded according to the level of destruction. If a building was damaged more than once to different degrees of destruction, the color coding was updated. Six categories of destruction were recorded, ranging from minor bast damage (yellow) to total destruction (black), as summarized in the first two columns of Table G.6 below.

Table G.6: Categories of Destruction in the LCC Bomb Damage Maps

Category of Destruction	Color	Destruction Variable	Destruction Index
Total Destruction	Black	Serious Damage	6
Damage Beyond Repair	Purple	Serious Damage	5
Serious Damage - Doubtful if Repairable	Dark Red	Serious Damage	4
Serious Damage - Repairable at Cost	Light Red	Serious Damage	3
General Blast Damage - Not Structural	Orange	–	2
Blast Damage - Minor in Nature	Yellow	–	1
Clearance Areas	Green	–	–

Note: First two columns summarize color-coded levels of destruction for each building in the LCC Bomb Damage Maps (BDMs); third column summarizes our baseline definition of serious damage; fourth column summarizes our index of destruction; we exclude the small number of clearance areas (green) that were areas assigned for post-war redevelopment (1.3 percent of the pre-war builtup area), which typically included both bombed areas and nearby areas with no destruction, because the choice to label parts of the city as clearance areas is endogenous.

The maps also indicated the point of impact of each V-1 and V-2 missile, with a V-1 strike denoted by a large black circle, and a V-2 strike shown by a smaller black circle. Inside these circles for V-1 and V-2 strikes, buildings were color coded with their level of destruction, regardless of whether this was caused by the missiles themselves or bombing by conventional aircraft. Later a number of green "clearance areas" were added to the maps, which were areas set aside for post-war redevelopment, and included both damaged and undamaged buildings.

We use "Serious Damage - Repairable at Cost" or worse as our baseline measure of wartime destruction, as summarized in the third column of Table G.6. We do not distinguish between the "Serious Damage - Repairable at Cost" and "Serious Damage - Doubtful if Repairable" categories, because the decision to repair building structures could be endogenous to economic considerations. We exclude blast damage from our baseline measure, because such blast damage is unlikely to permanently affect building structures. We also exclude the small number of clearance areas (green) that were areas assigned for post-war redevelopment (1.3 percent of the pre-war builtup area), which typically included both bombed areas and nearby areas with no destruction, because the choice to label parts of the city as clearance areas is endogenous. Nevertheless, we record the levels of color-coded wartime destruction beneath the green shading wherever possible.

In addition to our baseline measure, we also compute an index of wartime destruction, in which we assign numerical scores that are increasing in the severity of destruction, as summarized in the fourth column of Table G.6. "Blast Damage - Minor in Nature" is scored one; "General Blast Damage - Not Structural" is scored two; "Serious Damage - Repairable at Cost" is scored three; "Serious Damage - Doubtful if Repairable" is scored four; "Damage Beyond Repair" is scored five; and "Total Destruction" is scored six.

In Figures G.4 and G.5, we provide some photographic examples of destruction during the Second World War. Destruction extended throughout the LCC area and exhibited substantial idiosyncratic variation. Figure G.4 shows an area close to St. Pauls Cathedral in Central London, in which substantial parts of the surrounding area have been destroyed. Figure G.5 displays the suburban street of Needham Road in Notting Hill in West London. Although the building in the center of the photograph has been almost completely destroyed, other surrounding buildings are less heavily damaged.

Figure G.4: Photograph of Bomb Damage during the Second World War
to Paternoster Square Near St. Paul's Cathedral in Central London

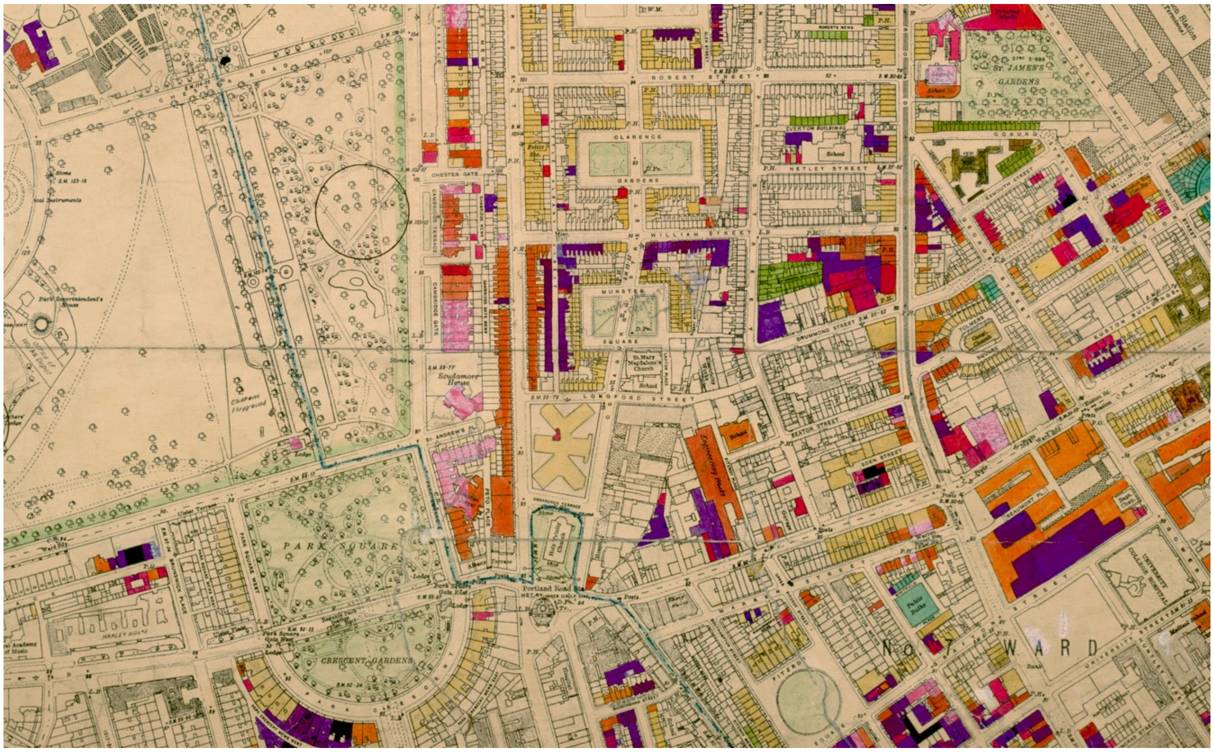
Source: Contemporary photograph.

Figure G.5: Photograph of Bomb Damage during the Second World War to the Suburban Street of Needham Road in Notting Hill in West London

Source: Contemporary photograph.

In Figure 1 in the paper, we show a small extract from one of the LCC bomb damage maps sheets for an area around Regent's Park in Central London, as reproduced in Figure G.6 below. Individual buildings and streets are discernible. Visible inside Regent's Park is one of the black circles for a V-1 impact. Even within this small extract from one map sheet, we observe the full range of levels of destruction ranging from yellow to black. Individual buildings can experience complete destruction, while surrounding buildings experience only light blast damage, or survive unscathed. This idiosyncratic variation in levels of destruction is consistent with the primitive bomb aiming technology at the time, the fact that much of the bombing occurred at night under conditions of a wartime blackout, and the fact that some of this damage was the result of fire caused by incendiary bombs, which spread in unpredictable directions.

Figure G.6: Excerpt of London County Council (LCC) Bomb Damage Map for the Area Around Regent's Park in Central London



Note: Excerpt from London Sheet V.5 of the LCC Bomb Damage Maps. Buildings color-coded by level of bomb damage: minor blast damage (yellow); general blast damage (orange); seriously damaged but repairable at cost (light red); seriously damaged and doubtful if repairable (dark red); damaged beyond repair (purple); and total destruction (black). Large black circle in Regent's Park shows a V-1 missile impact.

G2.3 Geolocating and Geoprocessing Wartime Destruction

We constructed our data on bomb damage to individual buildings in a sequence of steps according to the following procedure.

1. We obtained high-resolution digital scans of the 110 original LCC Bomb Damage Maps (BDMs) from the London Metropolitan Archives (LMA).

2. We georeferenced each of the 110 LCC BDMs, using the detailed 1:2,500 scale of the map sheets to precisely locate them in latitude and longitude space.

3. We compare the georeferrenced LCC BDMs to the 2014 Ordnance Survey MasterMap Topography Layer, downloaded from Edina Digimap in November 2014. This is a polygon shapefile with one polygon for the footprint of every building in modern-day London, which we use to measure the post-war built-up area.

4. By overlaying the modern built-up area at partial opacity over the georeferenced bomb damage maps, buildings which differed from modern London could be identified, deleted and manually redrawn to match pre-war London. In the final digitized bomb map, around 43% of the pre-war built-up area is the same as the modern built-up area.

5. We color coded the polygons for individual buildings in the pre-war built-up area with the corresponding color for the level of wartime destruction from the BDMs, as summarized in Table G.6 above.

6. We assign total built-up area and the amount of the built-up area with each level of destruction to each Output Area, based on the geographical boundaries of the Output Area in which the building footprint lies.

7. For each Output Area, we compute its total geographical land area, its total pre-war built-up area, and its pre-war built-up area with each level of destruction.

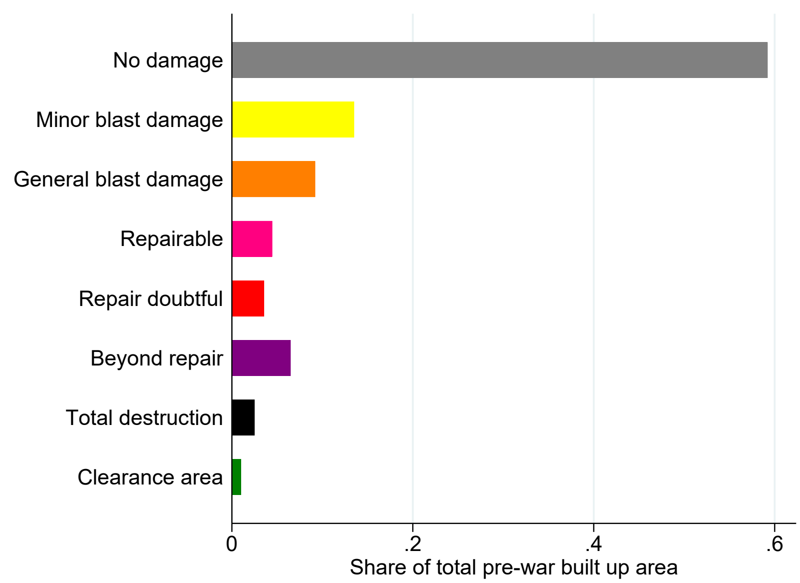
8. For each Output Area, we compute our baseline measure of *Serious Damage* as the fraction of its pre-war built-up area that experienced "Serious Damage - Repairable at Cost" or worse, as defined in the third column of Table G.6 above.

9. For each Output Area, we compute our *Destruction Index* as the weighted average of the scores for each level of destruction in the fourth column of Table G.6 above, using the shares of its pre-war built-up area as weights. Therefore, this destruction index is bounded between 0 (if none of the pre-war buildings in an Output Area were destroyed) and 6 (if all of the pre-war buildings in an Output Area were totally destroyed).

G2.4 Extent of Wartime Destruction

We find substantial destruction to the pre-war built-up area from the German bombing of London during the Second World War. In Figure G.7 below, we show the share of the pre-war built-up area within the LCC's boundaries with each level of destruction from the LCC Bomb Damage maps. More than 40 percent of the pre-war built-up area experienced some level of wartime damage. Around 10 percent of this pre-war built-up area was either totally destroyed (black) or destroyed beyond repair (purple). Approximately 17 percent of this built-up area experiences "Serious Damage - Repairable at Cost" or worse (pink or darker) according to our baseline measure of destruction. The green clearance areas set aside for post-war redevelopment make up around only 1.3 percent of the pre-war built-up area. We record the color-coded bomb damage beneath the green shading where we are able to do so.

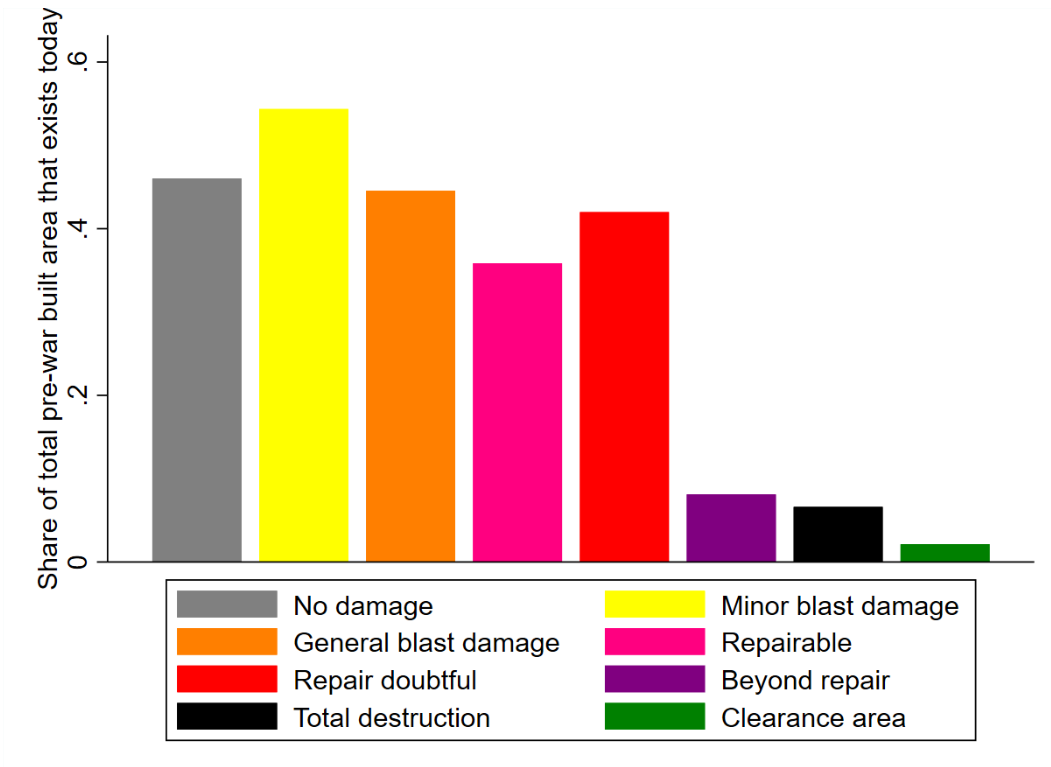
Figure G.7: Shares of Pre-war Built-up Area with Each Level of Wartime Destruction



Note: Shares of the pre-war built-up area of the LCC's total land area with each level of wartime destruction from the LCC Bomb Damage Maps; color codes for each level of wartime destruction are summarized in Table G.6 above.

In Figure G.8, we provide evidence that wartime destruction had long lasting effects on building structures. For the parts of the LCC's pre-war built-up area with each level of wartime destruction, we display the fraction of that pre-war built-up area that lies within the modern-day (2014) building footprint. Around 40 percent of the undamaged pre-war built-up area lies within the modern-day building footprint, consistent with the idea that there is substantial persistence in durable building structures over time. We find a similar level of persistence for Minor Blast Damage and General Blast Damage, as well as for the two lowest levels of destruction included in our baseline serious damage measure, namely Repairable and Repair Doubtful Serious Damage, as summarized in Table G.6 above. In contrast, we observe a sharp drop to around 10 percent or less in the share of the pre-war built-up area within the modern-day building footprint for the two highest levels of destruction of Beyond Repair and Total Destruction. This pattern of results is in line with the idea that wartime destruction had a substantial impact on durable building structures, including through the construction of council housing, which often replaced pre-war detached, semi-detached or terraced houses with post-war tower blocks.

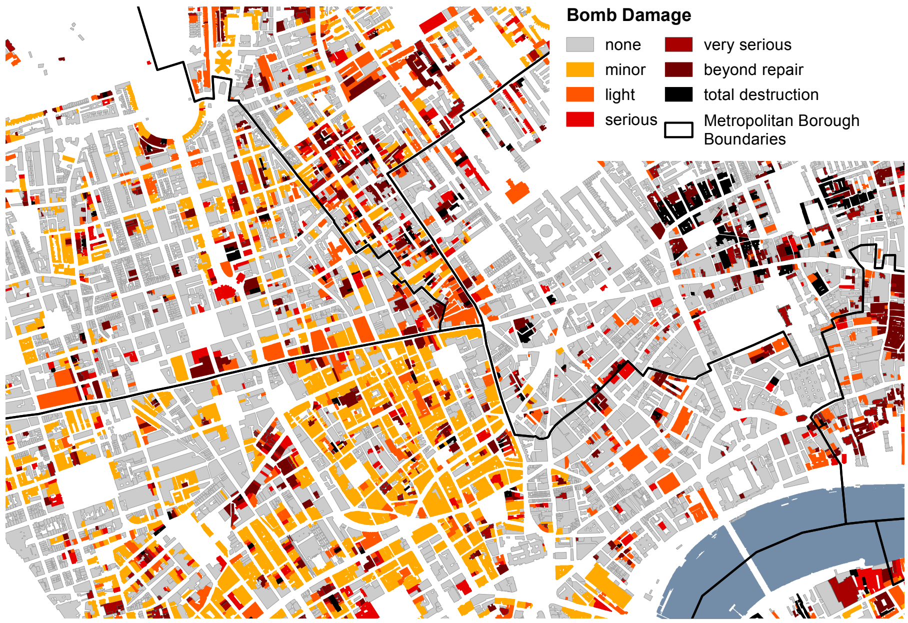
Figure G.8: Shares of Pre-war Built-up Area Within the Modern-Day Building Footprint by Level of Wartime Destruction



Note: Shares of the LCC's pre-war built-up area that lies within the modern-day building footprint by level of wartime destruction; colors corresponds to the color codes for each level of wartime destruction, as summarized in Table G.6 above.

In Figure G.9, we provide an example of a small extract from the shapefile constructed using the LCC bomb damage maps. This extract corresponds to the area around Regent's Park in Central London shown in Figure G.6 above. The white areas correspond to roads, railways or open spaces; the grey polygons show the pre-war built-up area and the colored polygons show different levels of wartime destruction from the LCC Bomb Damage Maps, as summarized in Table G.6 above; the blue area is the River Thames. As apparent from the figure, we obtain a detailed characterization of the extent of damage to the pre-war built-up area at a fine spatial scale.

Figure G.9: Extract of Digitized Bomb Damage Map for the Area Around Regent's Park



Note: White areas show roads, railways or open spaces; grey polygons show the pre-war built-up area and the colored polygons show different levels of wartime destruction from the LCC Bomb Damage Maps; colors correspond to the color codes for each level of wartime destruction, as summarized in Table G.6 above; blue area is the River Thames.

In Figure 3 in the paper, reproduced in Figure G.10 below, we show our geoprocessed data on wartime destruction for the LCC area. A high share of Output Areas experience some wartime destruction: 87 percent of the Output Areas containing pre-war buildings experience minor blast damage or worse, while 61 percent of them contain buildings that experience serious damage or worse (according to our baseline measure). The degree of destruction also varies substantially across these Output Areas: The share of the built-up area experiencing serious damage or worse ranges from 0 to 100 percent.

There is a clear geographical gradient in wartime destruction, which is greater in the East than in the West. This gradient is consistent with the fact that London's docks were located in the East and were targeted in the early stages of the Blitz. It is also consistent with German aircraft sometimes approaching from the East, using the Thames River to try to overcome the challenges of navigation at night. Nevertheless, the extent of idiosyncratic variation within narrow geographic areas is striking, with substantial destruction throughout the LCC area.

This idiosyncratic variation is consistent with the primitive bomb-aiming technology at the time, the fact that much of the bombing occurred at night under conditions of a wartime blackout, and the fact that some of this damage was the result of fire caused by incendiary bombs, which spread in unpredictable directions. It is also consistent with our finding that wartime destruction is uncorrelated with pre-war economic characteristics within 1 km hexagons.

Figure G.10: Second World War Destruction in the London County Council (LCC) Area

Note: The map shows the bomb damage for each building in the LCC area using the color scheme used by the original bomb damage maps: minor blast damage (yellow); general blast damage (orange); seriously damaged but repairable at cost (light red); seriously damaged and doubtful if repairable (dark red); damaged beyond repair (purple); and total destruction (black). Buildings that suffered no damage are shown in grey and clearance areas (1.3 percent of the pre-war built-up area) are in green.

G2.5 Comparison with Bomb Sight Data

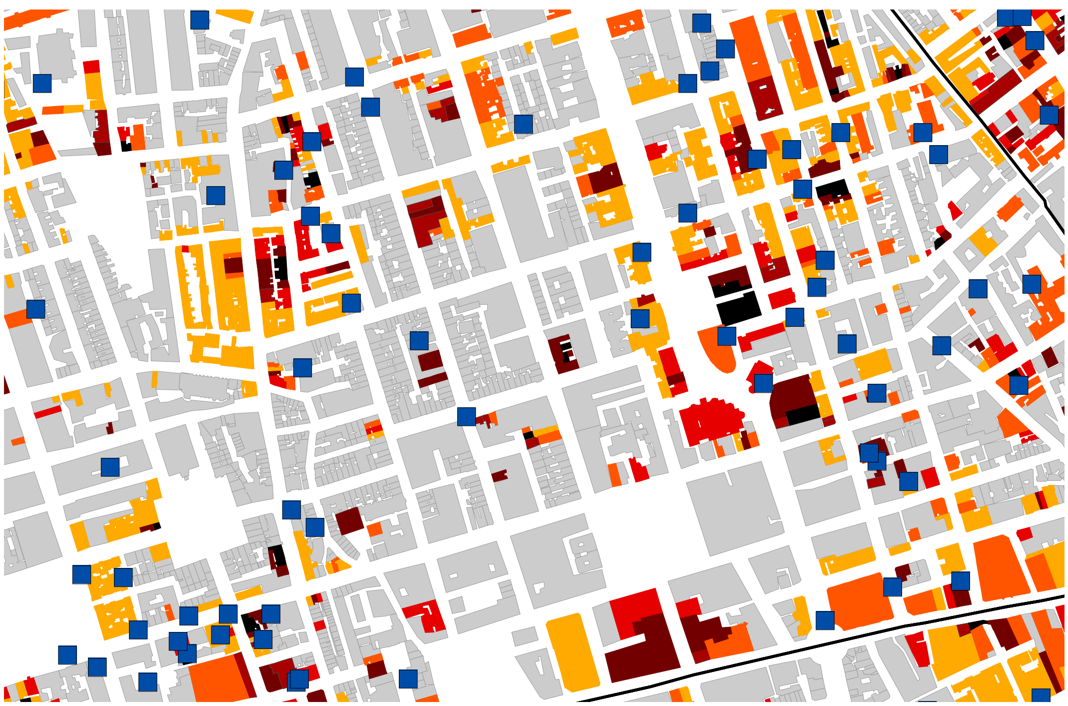
Our measure of wartime destruction is substantially more accurate than other prior sources of data on wartime destruction in London. We employed research assistants to precisely georeference the individual 1:2,500 scale LCC Bomb Damage Maps. Combining these georeferrenced map sheets to the 2014 Ordnance Survey MasterMap Topography Layer, which contains polygons for the footprint of every building in modern-day London, we construct the pre-war built-up area and wartime destruction to each pre-war building. In contrast, Fetzer (2023) applies automated color-recognition algorithms to crowd-sourced digital versions of these maps to construct an instrument for building energy efficiency based on wartime destruction.

Our measures of wartime destruction from the administrative LCC Bomb Damage Maps also differ substantially from the crowd-sourced BombSight data used in Dericks and Koster (2021), which claims to record the locations where German bombs landed. A major limitation of the BombSight data is that it does not record building damage. Therefore it does not distinguish between unexploded bombs, minor blast damage and total destruction. This limitation is especially problematic, because much of the wartime destruction in London was caused by fires that were started by incendiary bombs and then spread to neighboring buildings, which is not directly captured by the locations where German bombs landed.

We find many cases in our bomb damage data where destruction occurred, but no bomb impacts are recorded in the Bombsight data. In Figure G.11, we provide an example of such a comparison for the area around Regent's Park in Central London. As apparent from the figure, we find substantial inaccuracies in the Bomb Sight data: some buildings that were in fact totally destroyed are recorded as having no points of bomb impact in their vicinity; other buildings that in fact experienced no damage are recorded as points of bomb impact.

Therefore, our data from the LCC bomb damage maps are a substantial improvement over these other publicly-available data. Given the fine spatial scale over which neighborhood effects operate, accurately measuring the location of bomb damage is particularly important for our empirical application.

Figure G.11: Comparison with Data from the Bomb Sight Website



Note: Grey polygons show the pre-war built-up area and the colored polygons show different levels of wartime destruction from the LCC Bomb Damage Maps; colors correspond to the color codes for each level of wartime destruction, as summarized in Table G.6 above; blue squares show points of bomb impacts according to the Bomb Sight website.

G3 Pre-War Socioeconomic Status Data

Our pre-war data on socioeconomic status are taken from the *New Survey of Life and Labor in London* (henceforth NSOL), as published in Llewellyn Smith (1930). NSOL was a study of the socioeconomic conditions of the people of London that was undertaken from 1928-30 at the London School of Economics, directed by Sir Hubert Llewellyn Smith. This study itself was an update after 40 years of the earlier socioeconomic study of the *Life and Labor of the People in London*, as published in Booth (1889). We focus on NSOL because it provides data on socioeconomic status shortly before the German bombing of London during the Second World War.

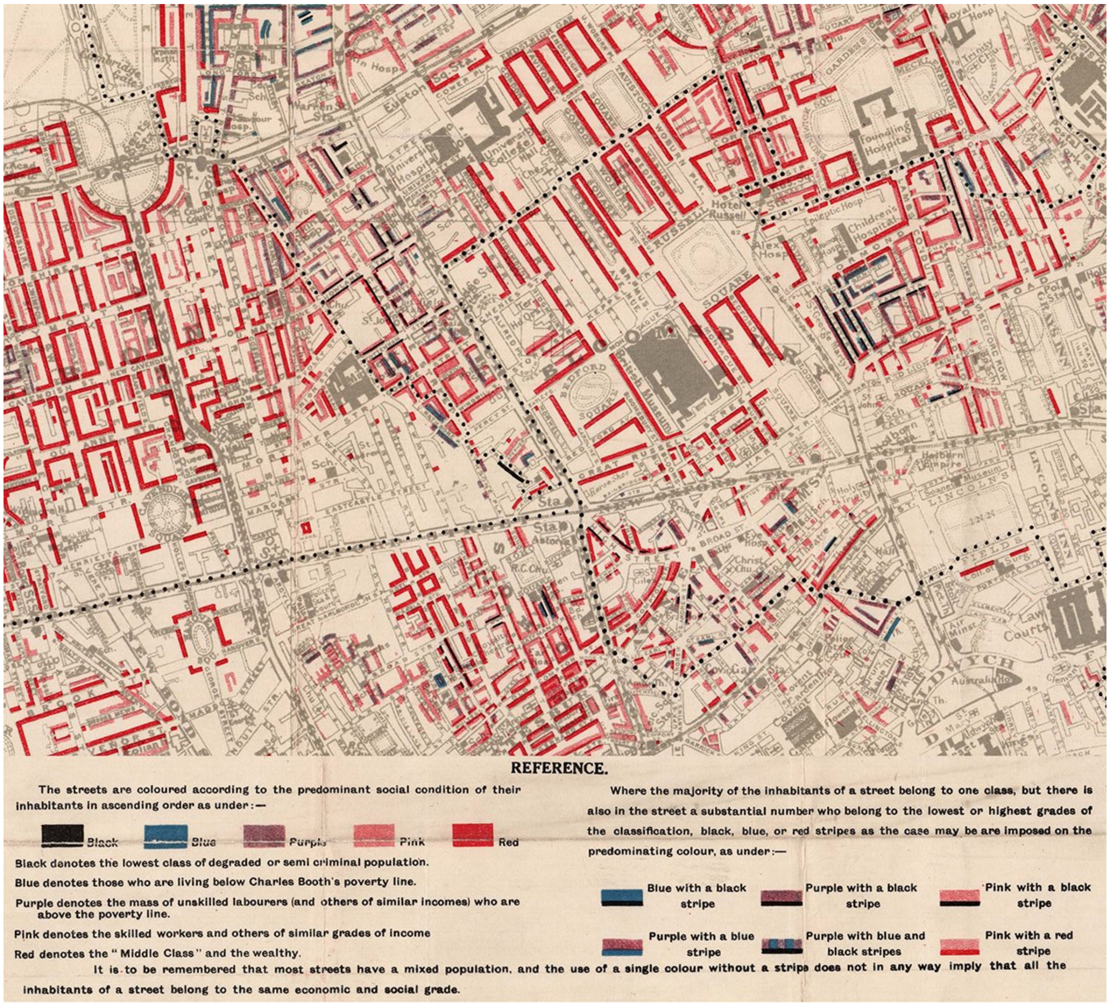
NSOL Data NSOL consisted of nine published volumes, which were concerned with London's industries, the social conditions in its neighborhoods, and the life and leisure of its inhabitants. Two volumes consisted of street maps, which were based on detailed pre-war Ordinance Survey maps, and showed individual buildings. On these street maps, each street segment was color-coded with the socioeconomic status of its inhabitants, as illustrated for the area around Regent's Park in Central London in Figure G.12 below.

Five main categories of socioeconomic status were distinguished: black ("lowest class of degraded or semi-criminal population"); blue ("below the poverty line"); purple ("unskilled laborers or others of similar income who are above the poverty line"); pink ("skilled workers and others of similar grades of income"); and red ("middle class and wealthy"). In addition, to these five main categories, several intermediate classifications were distinguished, where street segments contained a mixture of inhabitants in these different categories (e.g., pink with a black strip).

This socioeconomic classification was based on the income of the primary earner of the household, and interviews with 326 school attendance officers of the LCC, who made home visits to families with elementary school-age children. Additional information used in this classification of socioeconomic status included the recipients of benefits and job assistance from the Boards of Guardians and the Ministry of Labour, and additional information on criminality was obtained from the Criminal Investigation Department and Uniformed Branch of the Police Authorities.

The income of the primary earner was used to sort the residents of each street segment into the following four categories: (i) poor (blue), (ii) unskilled (purple), (iii) skilled (pink), and (iv) middle class (red). These four categories were based on the following "weekly income limits for a family of ordinary size": (i) £2 and under, (ii) £2 to £3, (iii) £3 to £5, and (iv) over £5, respectively. The presence of "moral sub-normality," such as drunkenness and gambling, was used to determine the fifth black category of "lowest class of degraded or semi-criminal population" in areas below the poverty line.

Figure G.12: Excerpt from New London Survey (1930) Poverty Maps
for the Area Around Regent's Park



Notes: Excerpt from Sheet 8 (Inner North West) of the NSOL (1930) Poverty Maps. Streets color coded by socioeconomic status, as indicated in the legend at the bottom of the map.

Measuring Pre-war Socioeconomic Status We measure the pre-war socioeconomic status of each residential building in the LCC area by combining these NSOL data on socioeconomic status by street segment, our data on the pre-war built-up area, and our 1936 rateable values data, using the following procedure.

1. We georeferenced the NSOL map sheets by overlaying them with the Ordnance Survey MasterMap Topography Layer from 2011.

2. We digitized the NSOL data by drawing lines in front of each street segment shown on the NSOL map sheets to create a line shapefile, in which each line is color coded with the corresponding NSOL socioeconomic category.

3. We measure the pre-war built-up area using the LCC Bomb Damage maps, as discussed in Section G2 above.

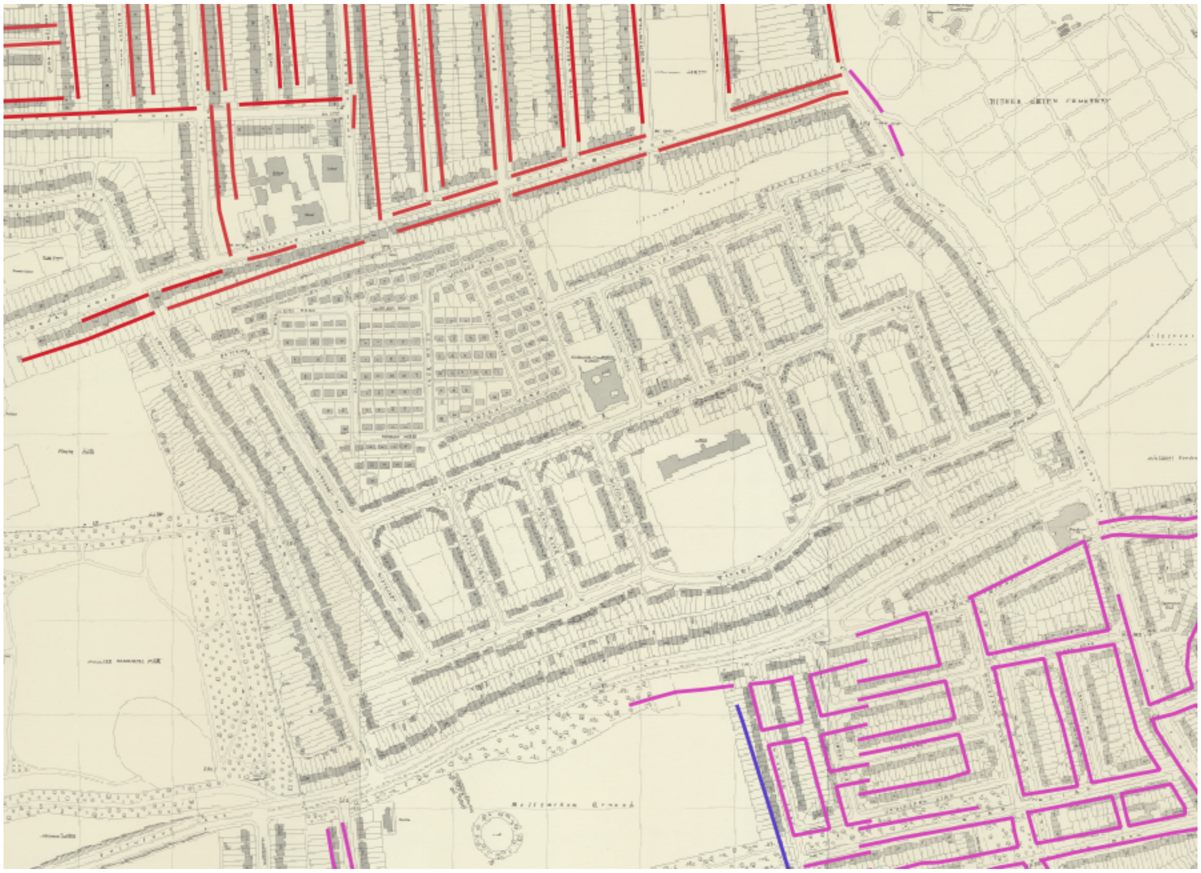
4. We determine whether each building in the pre-war built-up area is residential or not using our geocoded 1936 rateable value data, as discussed in Section G5 of this Online Appendix. For buildings of mixed use (e.g., flats above shops), we divide the building's total built-up area into commercial and residential components, using the corresponding shares of commercial and residential rateable in the total rateable value of the building.

5. We assign a NSOL socioeconomic status to each residential building by intersecting it with NSOL street segment lines. To account for small irregularities created during the digitization of the NSOL street segments, and small inconsistencies between the georeferencing of the NSOL and the LCC bomb damage maps, we use spatial intersections within a 5-meter buffer. In cases where one residential building is intersected by more than one line, we compare the distance between the polygon centroid and the centroid of each intersection, and assign a socioeconomic status to the residential building based on the closest NSOL street segment.

6. We thus obtain a single pre-war socioeconomic status for each residential building. Section G4 describes how we use this socioeconomic status for each residential building together with the 1931 population census data to determine the number of residents of each 2001 Output Area in each socioeconomic category.

Out of all the buildings that we determine to be residential (using the procedure described Section G5 of this Online Appendix), 43 percent are not intersected by any NSOL line within our buffer of 5 meters. Using the procedure outlined in Section G4 of this Online Appendix, we estimate that these non-intersected residential buildings house 31 percent of our 1931 population. We assign a socioeconomic status to each of these non-intersected residential buildings using the nearest NSOL line by minimum edge to edge distance, based on the assumption that nearby areas are likely to have similar population demographics. Of these non-intersected residential buildings, 80 percent lie within 15 meters of the nearest NSOL line. These non-intersected residential buildings are often in areas where the georeferencing of the NSOL maps and LCC bomb damage maps became misaligned. Around 5 percent are located in the City of London, which was not surveyed by the NSOL due to its low residential population. The rest are accounted for by developments on derelict land or around the outskirts of London which were built after the NSOL survey was conducted, an example of which is shown in Figure G.13 below.

Figure G.13: Example of Buildings Constructed after the NSOL Survey



Note: Map showing buildings constructed after the completion of the NSOL survey in grey, with NSOL survey result lines covering buildings to the north and south. Red NSOL lines show "Middle Class" streets; pink NSOL lines show "Skilled Workers;" and purple NSOL lines show "Poor." Our solution of assigning non-intersected buildings to the nearest NSOL line results in the northern half of these buildings being classified as red, while the southern half of these buildings is classified as pink.

Low, Middle and High-Income To construct consistent measures of socioeconomic status before and after the Second World War, we aggregate both our pre-war and post-war data on socioeconomic composition into the three groups of low, middle and high-income, as summarized in Table G.7 below, where our post-war data are discussed further in Section G6 of this online appendix. We find a relatively similar distribution of the share of the population across these three categories before and after the Second World War, which is consistent the fact that the occupation classification in the population census was heavily influenced by the Booth and NSOL studies of socioeconomic status. During the pre-war period, the low and high-income categories make up 24 and 28 percent of the population, respectively, which compares with 22 and 20 percent of the population, respectively, during the post-war period.

Table G.7: Pre-War and Post-War Socioeconomic Status

Socioeconomic Status	NSOL		Census 2001	
	Groups	Share	Groups	Share
Low	Extreme Poverty Below Poverty Line Unskilled Workers	0.24	Long-Term Unemployed Routine and Semi-Routine	0.22
Middle	Skilled Workers	0.47	Lower-Managerial Intermediate Occupations Own Account Workers Technical Occupations	0.58
High	Middle-Class and Wealthy	0.28	Higher-Managerial Higher-Professional	0.20

Notes: The table shows how we aggregate the different socioeconomic groups in the NSOL (Smith 1930) and the 2001 UK population census into our three socioeconomic groups. The shares do not always add up to one due to rounding errors.

Socioeconomic Index We construct an index of socioeconomic status at the Output Area level following Orford et al. (2022). We first assign a score to each socioeconomic group (low, middle and high), which equals the mid-point of the cumulative distribution of residents for the LCC area as a whole:

$$\$^{\mathrm{L}} = 0.5 \times \frac{R^L}{R}, \qquad \$^{\mathrm{M}} = \frac{R^L}{R} + 0.5 \times \frac{R^M}{R}, \qquad \$^{\mathrm{H}} = R^L + R^M + 0.5 \times \frac{R^H}{R},$$

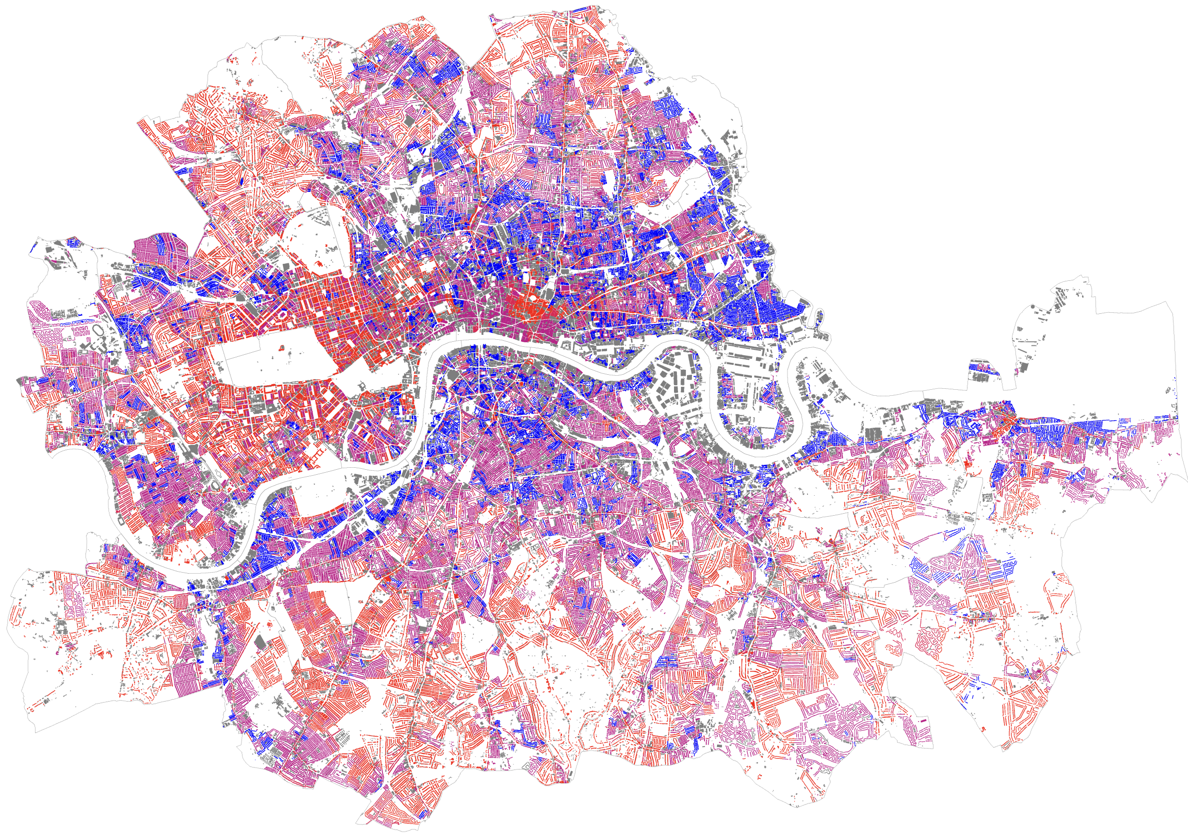
where R^L, R^M, R^H and $R = R^L + R^M + R^H$ are the number of low-income, middle-income, high-income and total residents for the LCC area as a whole. We next calculate the socioeconomic status ($\$_i$) of each Output Area i as the weighted average of these scores, using the shares of residents in each group for each Output Area as weights (R_i^o / R_i for $o \in \{L, M, H\}$):

$$\$_i = \left(\frac{R_i^L}{R_i} \times \$^L \right) + \left(\frac{R_i^M}{R_i} \times \$^M \right) + \left(\frac{R_i^H}{R_i} \times \$^H \right). \tag{G.1}$$

Finally, we rescale this socioeconomic index such that it varies between zero (all residents are low income) to one (all residents are high income).

In Figure G.14, we show the distribution of this index of socioeconomic status in the LCC area. We find a strong pattern of spatial sorting, with the areas characterized by higher property values in Figure 2 in the paper typically exhibiting higher socioeconomic status in Figure G.14. As a result, we find a clear East-West gradient in socioeconomic status, with higher values in the West End than in the East End. Nevertheless, we again observe substantial variation in socioeconomic status even within narrow geographical areas.

Figure G.14: Pre-War Index of Socioeconomic Status by Building in the LCC Area



Notes: Socioeconomic status by building in the LCC area based on the New Survey of London Life and Labor 1928-31. The color of each building corresponds to the socioeconomic index of the residents of the building with red denoting high and blue low socioeconomic status. Non-residential buildings such as factories or churches are shown in gray.

G4 Pre-War Population Data

Our pre-war population data is taken from the 1931 census ("Census of England and Wales 1931: County of London", H.M. Stationary Office 1932). The 1931 census counted 4,397,003 people as living in the London County Council (LCC) area. The smallest spatial units for which population is reported in the 1931 Census are the 316 Wards of the LCC area. In Subsection G4.1, we discuss our procedure to estimate total population and the population of each socioeconomic group (low, middle, and high-income) in 1931 for each 2001 Output Area. In Subsection G4.2, we report a number of specification checks on this estimation procedure.

G4.1 Population and Socioeconomic Status in 1931 by Output Area

Mapping 1931 Ward Boundaries We begin by mapping the boundaries of 1931 wards. No comprehensive map of 1931 ward boundaries for the entire LCC area appears to have survived. Digital scans of Ordinance Survey maps that show the ward boundaries in 1947 are available from the National Library of Scotland (maps.nls.uk). In the vast majority of cases, the ward names that are used in the 1931 census also appear on the 1947 maps, and we assume that the wards in the 1931 census correspond to the boundaries shown on the 1947 maps. For the boroughs where such a one-to-one correspondence did not exist, we searched for maps of pre-war wards in the archives of the boroughs, and were able to find pre-war boundaries maps from years close to 1931 that showed a set of wards that corresponded to the 1931 census returns. We used these maps to draw a shapefile of the 1931 ward boundaries.

Measuring 1931 Population and Socioeconomic Status for 2001 Output Areas We allocate 1931 population and 1931 population by socioeconomic group (low, middle, and high-income) to 2001 Output areas using our data on the pre-war built-up residential area, our shapefile of the 1931 ward boundaries, and our data on pre-war socioeconomic status from the New Survey of London (NSOL), according to the following step-by-step procedure:

1. We measure the pre-war built-up area using the LCC Bomb Damage maps, as discussed in Section G2 of this Online Appendix.

2. We determine whether each building in the pre-war built-up area is residential or not using our geocoded 1936 rateable value data, as discussed in Section G5 of this Online Appendix. For buildings of mixed use (e.g., flats above shops), we estimate the fraction of a building used residentially and commercially using the descriptions of buildings in the rateable value data as described in more detail in Section G5 of this Online Appendix.

3. We assign buildings to 1931 wards by intersecting the pre-war built-up area from the LCC Bomb Damage maps with our shapefile for 1931 ward boundaries.

4. We allocate the 1931 ward population across residential buildings in a ward using the shares of these residential buildings in the total residential built-up area within the ward. By construction, the sum of population across residential buildings within each ward equals the 1931 ward population reported in the 1931 population census.

5. We determine the NSOL socioeconomic status of the residents of each building using our geocoded data from the New Survey of London (NSOL), as discussed in Section G3 of this Online Appendix. NSOL distinguishes five main socioeconomic categories (black, blue, purple, pink, red) and intermediate mixed classifications (e.g., pink with a black stripe). We split the residents of buildings with these intermediate mixed classifications equally between categories. Therefore, in a building marked pink with a black strip, half of the residents are assigned to pink and the remaining half are assigned to black.

6. We compute the 1931 population of each 2001 Output Area by summing the population of all residential buildings within its boundaries. For residential buildings that straddle the boundary between Output Areas, we allocate the population of the residential building to each Output Area based on the share of the building's built-up area in each Output Area.

7. We compute 1931 employment by residence for each 2001 Output Area by rescaling its 1931 population by the ratio of employment by residence to population at the borough level from the 1921 population census.

8. We compute the 1931 population of each NSOL socioeconomic group for each 2001 Output Area by summing the population of the residential buildings assigned to each NSOL socioeconomic category, as discussed in Section G3 of this Online Appendix. For residential buildings that straddle the boundary between Output Areas, we again allocate the population of the residential building to each Output Area based on the share of the building's built-up area in each Output Area.

9. We aggregate the NSOL socioeconomic categories to our three socioeconomic categories of low (black, blue, purple), middle (pink) and high (red) for each 2001 Output Area.

10. We compute 1931 employment by residence for each socioeconomic group (low, middle, high) for each 2001 Output Area by rescaling the 1931 population for each socioeconomics group by the ratio of employment by residence to population at the borough level from the 1921 population census.

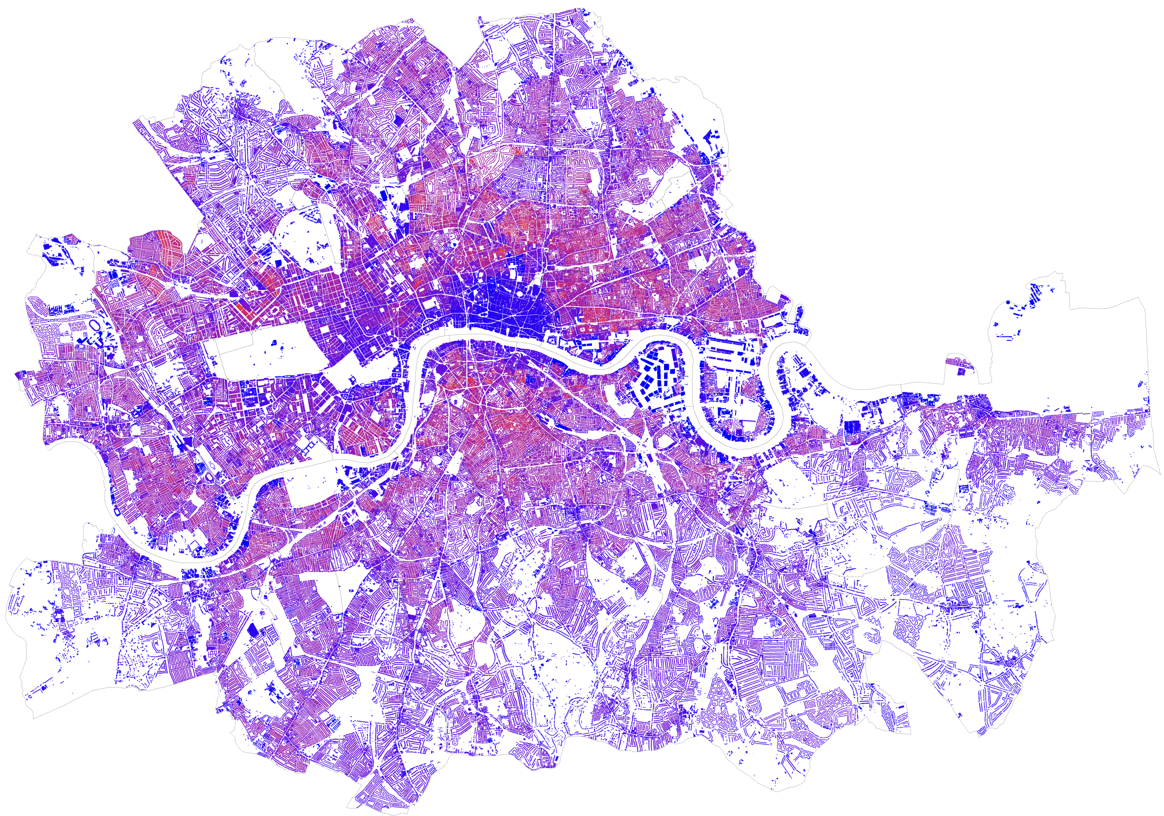
Population Density by Residential Building All steps in these procedures ignore integer constraints and hence our population estimates are not integers. Of the 9,041 output areas that make up the LCC area, 321 do not have positive population levels in 1931 according to our estimates. As the 2001 output areas were constructed in a way that they all have positive population counts in 2001, this is consistent with changes in the composition and extent of the built-up area of London between 1931 and 2001.[16]

Figure G.15 shows 1931 population density for the pre-war built-up area, measured as people per square meter of built-up area. We find an intuitive pattern of variation in population density. Most of the buildings in the most central parts of London have low population densities, because they were primarily used commercially rather than residentially in 1931. We find the highest levels of population density in the areas immediately surrounding the most central parts of London,

[16]Summing our estimated 1931 population across all Output Areas in the LCC results in a slightly lower total population than reported in the 1931 census. As discussed in Section G1, 2001 Output Areas do not exactly correspond to the LCC area. As a result, our estimation procedure allocates a very small fraction of the LCC population to built-up area that happens to be outside the boundaries of 2001 Output Areas.

with lower population density in the outer suburbs. We also observe lower population density in the areas along major roads, which typically contain primarily shops and office space.

Figure G.15: Population Density of Buildings in the London County Council (LCC) Area

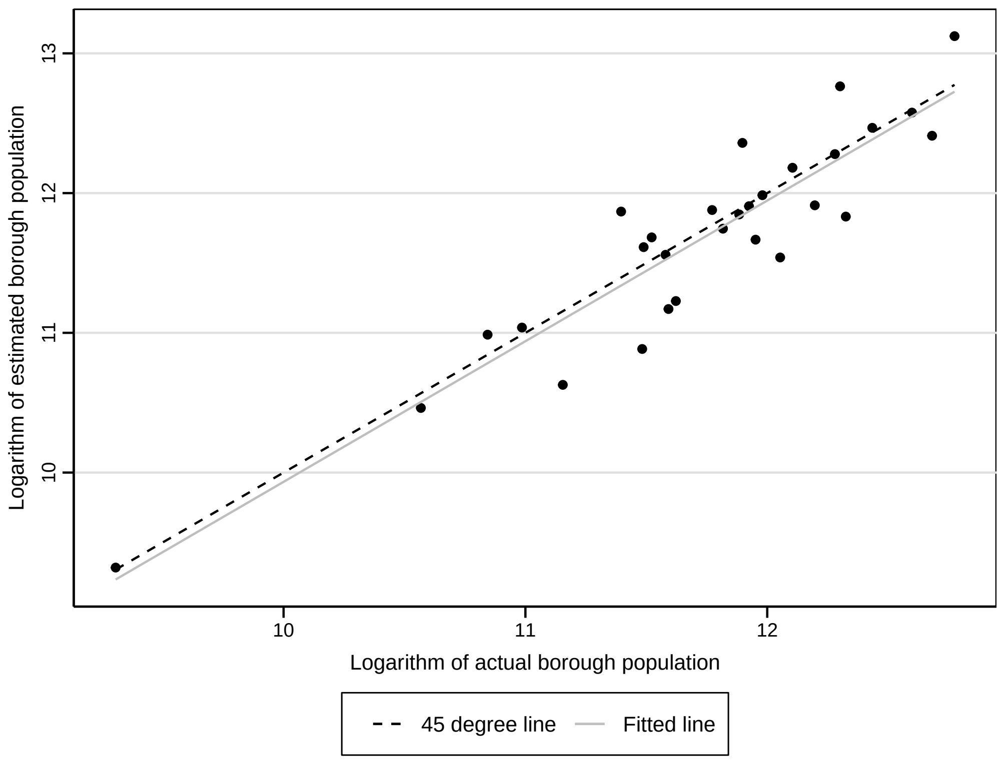


Note: Population density in 1931 for the pre-war built-up area, measured as people per square meter of built-up area, and constructed as described in the text of this subsection; blue denotes lower values; red indicates higher values.

G4.2 Robustness Checks

As a specification check on our procedure for estimating 1931 population for each 2001 Output Area using the pre-war residential built-up area, we undertake two robustness checks. First, we use our approach to allocate the total population of the LCC in 1931 to the 29 boroughs inside the LCC, for which we also observe their actual population in the 1931 Population Census. Figure G.16 compares our estimates of the borough population with the actual borough population reported in the census. We find that our estimates are highly correlated with the actual population of the boroughs with an R-squared of 0.85 and the estimates and actual values lie close to the 45 degree line. Boroughs for which our approach overestimates the borough population (points above the fitted line) are peripheral boroughs that plausibly have lower-rise buildings and generally lower population densities per square meter of built-up area.

Figure G.16: Estimating Borough Population from the LCC Population

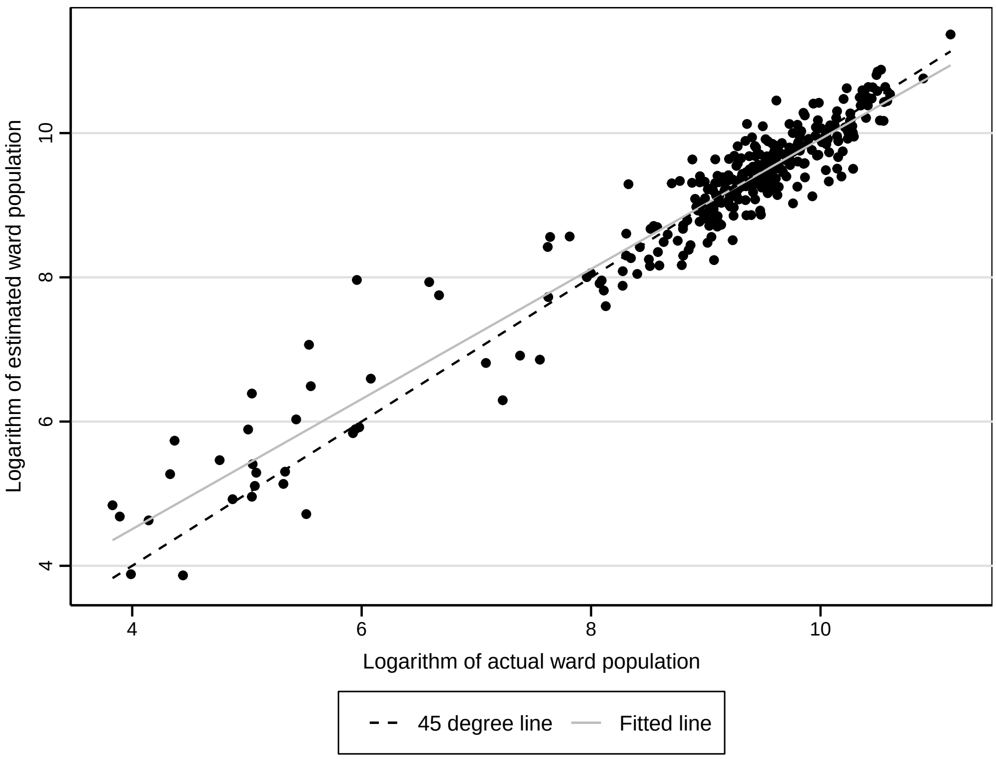


Note: The figure shows the correlation between estimated and actual 1931 borough level population totals. The estimated borough populations are derived by distributing the total population of the LCC across boroughs using the residential built-up area. The slope of the regression line is 1.006 and the R-squared of the regression is 0.85.

Second, we use the same approach to estimate the ward-level population in 1931 from the borough-level population in 1931, where again we observe the actual ward-level population in the 1931 Population Census. Figure G.17 shows the correlation between our estimates of the ward-level population and the actual ward-level population reported in the 1931 census. Implementing our approach for these smaller spatial units, we find an even higher correlation between our estimates and the actual ward population (R-squared of 0.92), and the estimates are again close to the 45 degree line.

Taken together, these two robustness checks provide strong support for our approach of estimating 1931 population for 2001 Output Areas from the ward-level totals reported in the 1931 Population Census and the pre-war residential built-up area.

Figure G.17: Estimating Ward Population from Borough Population



Note: The figure shows the correlation between estimated and actual 1931 ward level population totals. The estimates at the ward level are derived by distributing the borough population across wards using the residential built-up area. The slope of the regression line is 0.90 and the R-squared of the regression is 0.92.

G5 Pre-War Property Values Data

Our pre-war property values data are measured using rateable values for individual properties in the London County Council (LCC) area in 1936. In Subsection G5.1, we introduce the rateable values data from the 1936 valuation list. In Subsection G5.2, we discuss the digitization of this 1936 valuation list and the categorization of property types. In Subsection G5.3, we outline the geolocation of the rateable values data for individual properties.

In Subsection G5.4, we discuss the data processing to match the valuations for individual properties to building footprints from the LCC Bomb Damage Maps. In Subsection G5.5, we report descriptive statistics on our building-level data on rateable values. In Subsection G5.6, we use these building-level data to estimate average pre-war property values for each Output Area.

G5.1 Rateable Value Data for 1936

The UK has a long tradition of assessing the value of buildings and land for taxation purposes, which goes back to at least the Poor Relief Act of 1601. Since the Valuation (Metropolis) Act

1869, such valuations, which are referred to as the "rateable value" of a property, were compiled every five years. The original valuation list of the 1936 valuation for the London County Council (LCC) have survived in the London Metropolitan archives (LMA). The handwritten 1936 valuation lists run to approximately 50,000 pages. The lists are organised by London borough and are additionally split into a Part I, which contains residential and smaller commercial properties and a Part II, which contains valuations relating to industry, freight, transport and "special properties". The latter encompasses buildings used by the national and local governments, utilities, schools, hospitals, police and emergency services, and theatres and cinemas.[17]

The valuations were drawn up by individual boroughs and appeals to these valuations by property owners were heard by local assessment committees. Some of these revisions to the value of properties due to successful appeals are visible in the valuation list as corrections in red ink. The formal definition of the rateable value of a building is, "the annual rent which a tenant might reasonably be expected, taking one year with one another, to pay for a hereditament, if the tenant undertook to pay all usual tenant's rates and taxes ... after deducting the probable annual average cost of the repairs, insurance and other expenses" (see London County Council 1921).

These rateable values cover all categories of property, including public services (such as tramways, electricity works etc), government property (such as courts, parliaments etc), private property (including factories, warehouses, wharves, offices, shops, theaters, music halls, clubs, and all residential dwellings), and other property (including colleges and halls in universities, hospitals and other charity properties, public schools, and almshouses). The three main categories of exemptions are: (1) Crown property occupied by the Crown (Crown properties leased to other tenants are included); (2) Places for divine worship (church properties leased to other tenants are included); (3) Mines and quarries.

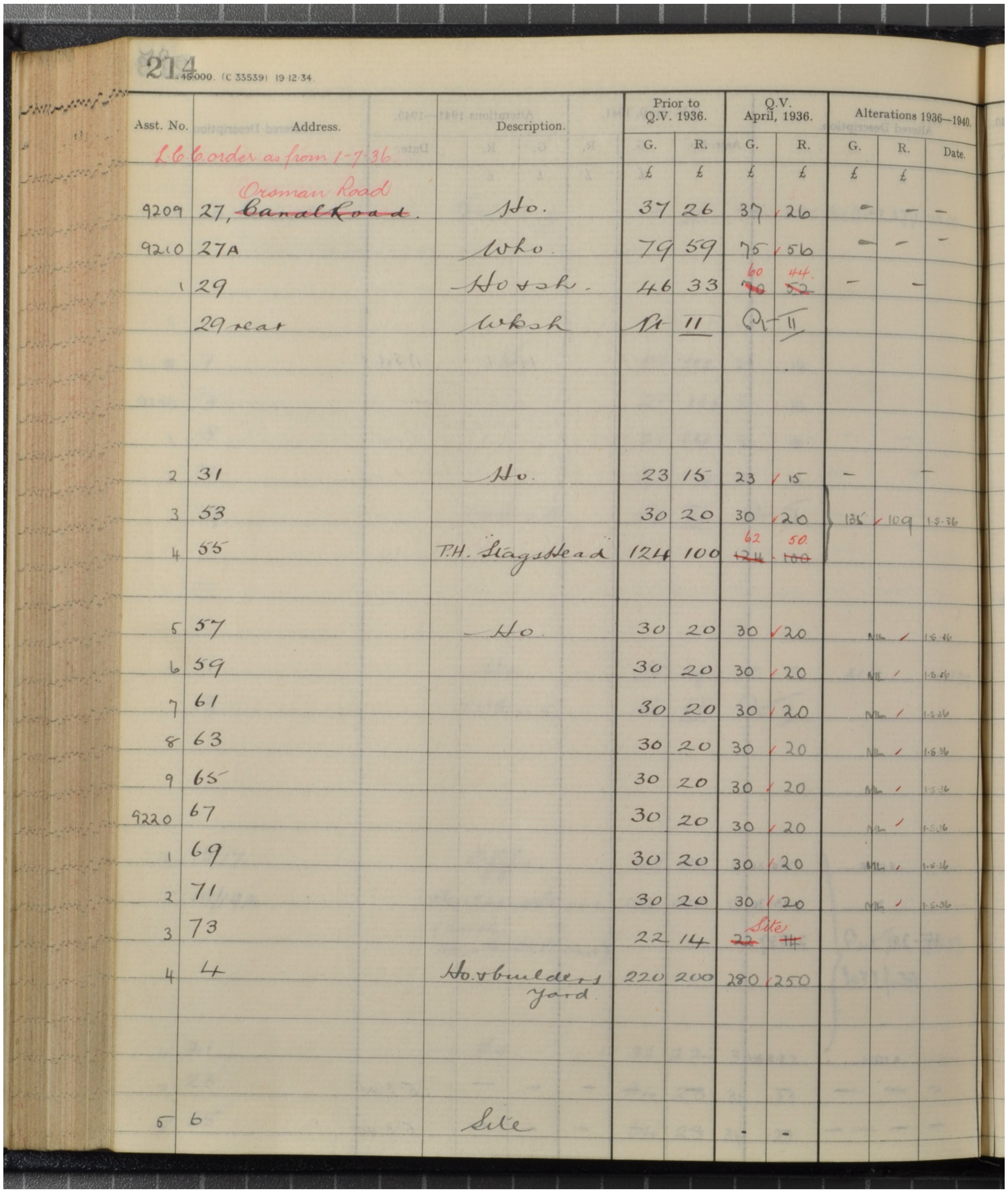
Figure G.18 shows a typical page from the valuation lists for Part I. Each valuation entry on the list reports a street and house number, a brief description of the property, which is usually abbreviated, and the rateable value. For many properties (but not all properties) the lists report both a gross value and rateable value, where the gross value does not deduct the probable annual average cost of the repairs, insurance and other expenses. For comparability and completeness, we use the gross rateable value of a property.

G5.2 Digitisation and Property Types

In a first step, we photographed the approximately 50,000 pages of valuation records and used a professional data entry company to type the information on each page into a corresponding Excel spreadsheet, with columns for the street, street number, description and gross rateable

[17] The 1936 valuation lists are held at the LMA under the catalogue reference "LCC/VA/V/L". See the LMA website for more information on the valuation material that they hold.

Figure G.18: Example Page from the 1936 Valuation List for the LCC Area



214 45000 (C 35539) 19·12·34

Asst. No.	Address.	Description.	Prior to Q.V. 1936. G.	Prior to Q.V. 1936. R.	Q.V. April, 1936. G.	Q.V. April, 1936. R.	Alterations 1936—1940. G.	Alterations 1936—1940. R.	Date.
	L.C.C. order as from 1-7-36		£	£	£	£	£	£	
9209	27, ~~Canal Road~~ Orsman Road	Ho.	37	26	37	26	–	–	–
9210	27A	Who.	79	59	75	56	–	–	–
1	29	Ho & sh.	46	33	~~70~~ 60	~~52~~ 44	–	–	
	29 rear	Wksh	Pt	11	Pt	11			
2	31	Ho.	23	15	23	15	–		–
3	53		30	20	30	20	135	109	1-5-36
4	55	T.H. Stags Head	124	100	~~124~~ 62	~~100~~ 50			
5	57	Ho.	30	20	30	20	Nil		1-5-36
6	59		30	20	30	20	Nil		1-5-36
7	61		30	20	30	20	Nil		1-5-36
8	63		30	20	30	20	Nil		1-5-36
9	65		30	20	30	20	Nil		1-5-36
9220	67		30	20	30	20	Nil		1-5-36
1	69		30	20	30	20	Nil		1-5-36
2	71		30	20	30	20	Nil		1-5-36
3	73		22	14	~~22~~ Site	~~14~~			
4	4	Ho. & builders Yard	220	200	280	250			
5	6	Site			-	-			

Notes: This page from Part I of the 1936 valuation lists shows several odd-numbered properties on Orsman Road in Shoreditch. The page shows mainly houses (Ho.), with a wharehouse (Who.), workshop (Wksh.), and the pub “Stag’s Head”. LMA reference code "LCC/VA/V/L/35/192/219"

value. To classify properties into different types such as houses, apartments or shops we ran a string detection algorithm on the descriptions entered by the valuers for each property. This assigns one or more labels to each valuation, out of 91 options such as “pub,” “workshop” or “library.” We aggregate this fine classification of properties into more aggregated categories as described below. Note that a particular building can contain several valuations such as “ground floor flat" and "first floor flat" or "workshop" and "first floor flat."

We use the descriptions of the valuers of each property to determine whether all or parts of the building were used for residential purposes. Many buildings contain a single valuation that is either residential or non-residential. If a building has two or more separate valuations, some of which are residential and some non-residential, we estimate the share of the building used for residential purposes by dividing the residential rateable value of the building by the total rateable value of the building. Sometimes both residential and non-residential uses are found within one valuation, such as a property that is described as "house and shop," which would have contained both a shop and residential areas. In such cases, if the total rateable value of the property is less than £300 in 1936 pounds, the valuation is categorised as half residential and half non-residential. Mixed valuations worth more than £300, which is substantially more than the typical rateable value of a residential house in London in 1936, we assume to be entirely non-residential to avoid classifying large factories that might have a small residential property for a warden as half residential and half non-residential.

We drop from the valuation lists a small number of special cases. First, the valuation lists contain some valuations of non-building structures such as railway tracks or gas pipework. We drop these valuations from the list and only keep buildings. Second, in a few cases, valuations refer to both buildings and non-building structures, such as pipework, and we deal with those in the data cleaning steps described below. Third, we also drop valuations of railway stations as it is difficult to determine what part of the valuation refers to the buildings versus the machinery contained within them, such as tracks. Finally, the valuation lists do not report values for churches and other religious buildings, but do include the value of associated buildings such as the houses of vicars, as discussed above.

G5.3 Geolocation

To assign the approximately 1 million entries in the 1936 valuation lists to modern output areas we need to associate each valuation with a building shown on the bomb damage maps described in Section G2.2 of this Online Appendix. Initially, we attempted to use the Google Maps API to geolocate the addresses in the digitized valuation lists. Using this approach produced relatively low quality results even after taking care of street name changes. Particularly in areas that were heavily destroyed, street layouts were sometimes changed, with old streets disappearing altogether.

To avoid these problems, we decided to manually locate the valuations on the bomb damage maps. While the bomb damage maps are very detailed, they do not show house numbers. To overcome this problem we used detailed Ordnance Survey maps at 1:1,250 scale, that show London in the early post-war period and include street names and building numbers. In cases where the built-up area is missing on the post-war map due to war-time destruction, we also consulted a Ordnance Survey map (1:1,056 scale) of London in the 1840s to 1860s to determine the likely layout of houses before the war. Georeferenced scans of both these maps are available on the webpage of the National Library of Scotland at maps.nls.uk.

To speed up the digitization, we exploited the structure of the handwritten valuation list. The order of the mainly residential valuations in Part 1 typically follows the order of properties on one side of a street. In these cases, we only recorded the coordinates of the first and last property in straight rows of terraced houses with consecutive street numbers and interpolated the locations of the addresses in between. For all other valuations, a point in the middle of the building corresponding to a valuation was selected. Industrial and commercial valuations reported in Part 2 of the valuation lists often comprise multiple buildings, such as several buildings that are part of a larger plant. In these cases, we allocated the total value reported in the valuation across all buildings that are part of this valuation in proportion to the footprint of the buildings.

G5.4 Data Processing

For the vast majority of valuations on the 1936 valuation lists, we were able to locate the building that the valuation refers to. The success rate varies across boroughs with the lowest success rate being 90% in Camberwell and Stepney. For all other boroughs the success rate exceeds 95%, with 18 boroughs having success rates higher than 98%. It is to be expected that the success rate is less than 100% as the built-up area of London changes continuously and the 1936 valuation was not taken at exactly the same time as the built-up area shown on the bomb damage maps, which we use as our baseline for the pre-war built-up area of London.

To maximise the quality of the data we undertake four data cleaning steps. First, some valuations are reported both in Part 1 and Part 2 of the original valuation lists and were as a result geo-located twice. To avoid double counting these valuations we search for all valuations with the same rateable value that are geo-located within 40m of each other. If a manual inspection of the descriptions of these two properties suggested that two valuations are very likely duplicates, one of the two valuations was dropped. Second, valuations with coordinates within 4 meters of a building on the bomb damage map are assigned to the nearest building. Third, valuations more than 4 meters away from any building on the bomb damage maps are also assigned to the nearest building on the map if their value is less £100 in 1936 pounds. Finally, valuations worth more than £100 in 1936 pounds and more than 4 meters away from a building on the bomb damage map were manually inspected and either reassigned to the correct building or dropped if no appropriate building could be identified.

After these cleaning steps a large majority of buildings shown on the bomb damage maps have valuations from the 1936 valuation list attached to them. Buildings may not have an assigned value for multiple reasons. First, the valuation of the building exists but could not be located on the map due to illegible writing. Second, the building shown on the bomb damage map was built after the 1936 valuation was conducted. Third, a small number of the volumes of the valuation lists have been lost.[18] To address this problem, the buildings without valuations were manually inspected starting with the buildings with the largest footprint. From the shape of the building, its location and sometimes descriptions on the post-war map, the buildings were assigned to one of 48 types including a category "undefined". This process continued until at least 95% of the built-up area of each borough was either matched to a valuation or categorised in this way. For the remaining buildings we assumed that they must be part of other valuations already in our data, such as houses with a "garden shed" or "garage", where the value of the shed or garage is included in the valuation of the house.[19]

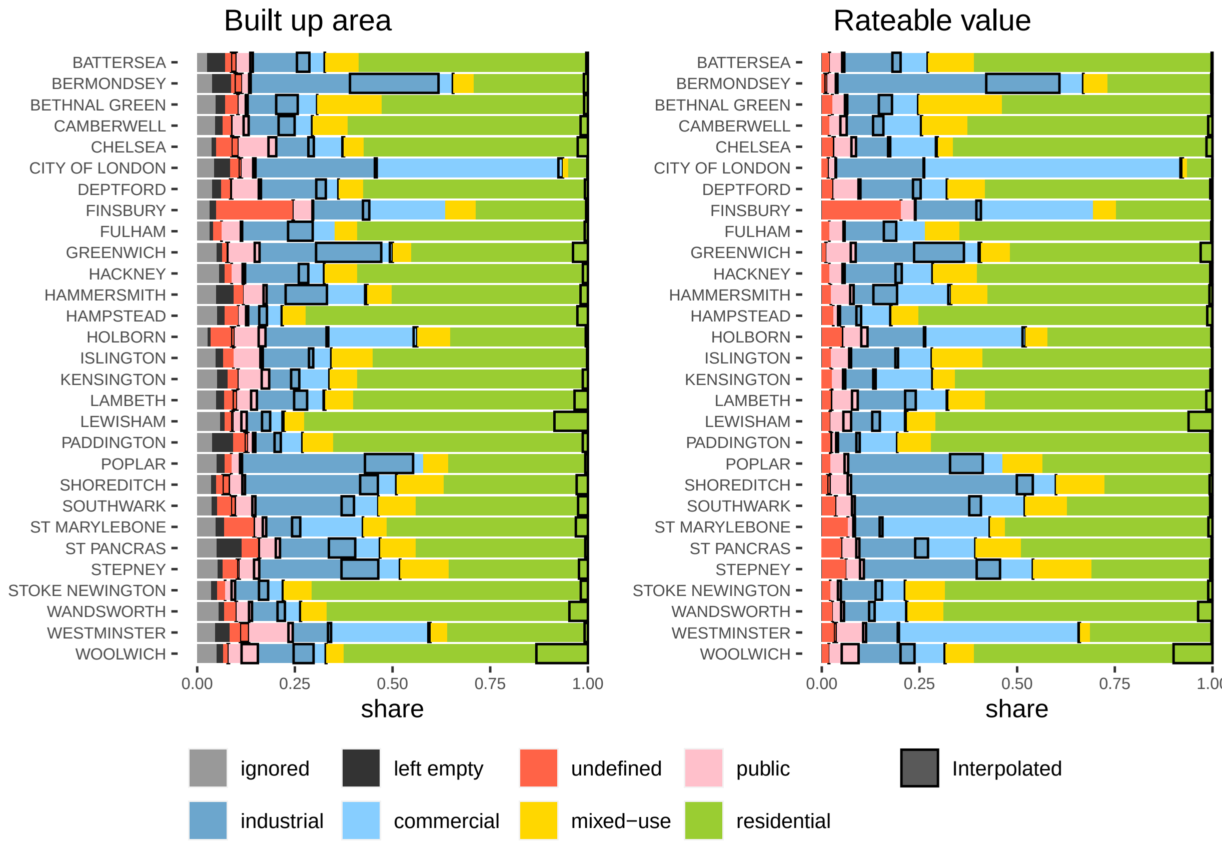
To interpolate the value of larger buildings categorised in this way, they were aggregated into five broad categories: residential (e.g., private houses or flats), commercial (e.g., shops or small workshops), industrial (e.g., factories), public (e.g., police stations, libraries) and other, where "other" contains both mixed-use buildings and buildings whose type could not be determined. For each of these five broad categories and for each of the 29 boroughs, we regress the logarithm of rateable value on the logarithm of the footprint of a building for buildings where we observe the rateable value, and use these regressions to predict the value of buildings without a rateable value. We estimate this regression separately by both building category and borough to allow for heterogeneity in the slope coefficient in this regression. The footprint of a building is a strong predictor of its rateable value within building types and boroughs, with a median R-squared in these $29 \times 5 = 145$ regressions of 0.58.

After these cleaning steps, we are left with two main datasets. The first contains one row for each unique shape-valuation pair. In this dataset a house containing three flats that have been separately valued appears as three separate rows that record the value of each flat and all three rows contain the same shape ID, i.e. the same identifier for the footprint of the building that all three flats are part off. Similarly, a factory made up of three separate buildings will appear as three separate rows with the same valuation ID and three different shape IDs. The total value of the factory is spread across the buildings proportional to their footprint as explained above. The second dataset adds up the rateable value associated with each building. We create this second dataset by collapsing the first dataset by shape ID, such that the building with three flats will appear as one row with a value equal to the sum of the three flats contained in it.[20]

[18] For Woolwich one of the Part 1 volumes covering one of the wards of Woolwich no longer exists. For Bethnal Green, Hampstead and Hammersmith no Part 2 volume with industry, freight and transport valuations is available.

[19] To reduce the influence of outliers we finally also dropped a small number of valuations that had implausibly large or small rateable values per square meter of built-up area. The values of these buildings were then also

Figure G.19: Composition of LCC Rateable Value by Borough



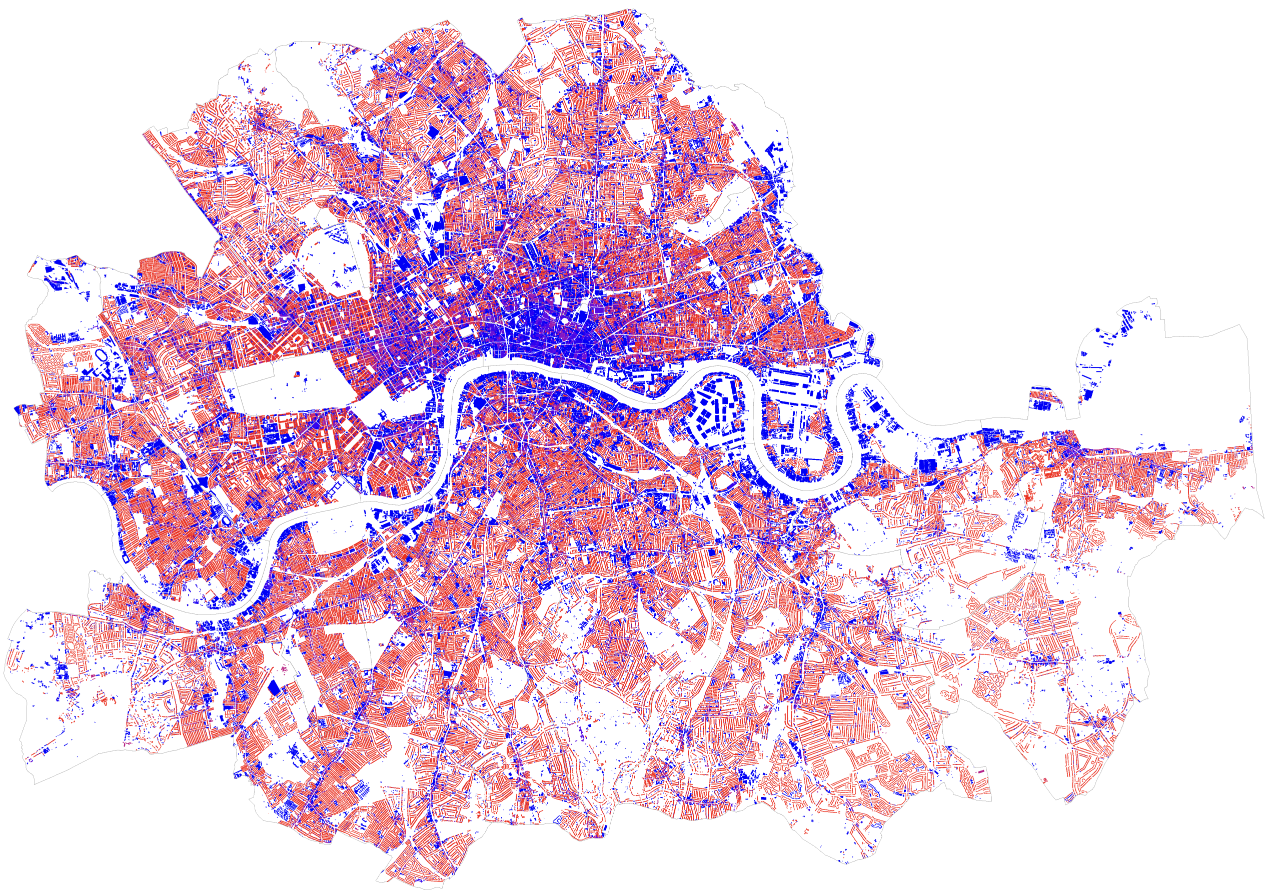
G5.5 Descriptive Statistics

Figure G.19 gives an overview of the rateable value in each of the 29 pre-war boroughs that make up the LCC area and also shows what part of this rateable value has been interpolated during the data entry process. Overall, 9.5% of the total built area is interpolated in the final dataset. After this interpolation, 92.8% of the total built area in the LCC area has been assigned a value. Another 2.4% of the total built area was deliberately left without a value due to the type of building such as religious buildings or railway stations as described above. The remaining 4.8% of the built-up area are smaller structures, the value of which is likely included in nearby buildings, such as garden sheds or garages included in the value of a house. We do not interpolate the value of these

interpolated using the approach described below.

[20]Note that it is often difficult to determine from maps what the legal boundaries of a particular building are. To speed up the digitization process we sometimes combined contiguous built-up area that had the same bomb damage grade into one polygon even if the buildings that are part of this contiguous built-up area may have had different legal titles. If a polygon contains several buildings, such as two semi-detached houses being combined into one polygon, the values of all properties with addresses inside this polygon are added up in our building level dataset.

Figure G.20: Residential and Non-Residential Buildings in Pre-War London



Notes: The map shows the built-up area of the LCC area prior to the Second World War as shown on the bomb damage maps with with blue representing non-residential buildings, purple mixed-use buildings, and red residential buildings.

structures in order to avoid overestimating the total rateable value in an area.

The distribution of the rateable value across residential and non-residential buildings visible in Figure G.19 follows an intuitive pattern. Across the entire LCC area, 50% of all rateable value is residential, but this share ranges from a low of only 7% in the City of London to shares of over 70% in predominantly residential boroughs. Figure G.20 shows the distribution of residential, non-residential and mixed-use buildings on a map. The map reveals the large concentration of non-residential buildings in the most central locations of the LCC area and in heavily industrial areas in the East of London. However, the map also shows concentrations of non-residential buildings along major roads in further outlying regions, consistent with the concentration of shops and offices on larger roads.

G5.6 Estimating Average Pre-War Property Values by Output Area

To estimate the average 1936 value of properties in each 2001 Output Area we run the following regression:

$$\ln(RV_{ij}) = \theta_j + \beta X_i + \varepsilon_{ij} \tag{G.2}$$

where the dependent variable $\ln(RV_{ij})$ is the logarithm of the rateable value of each valuation i in our shape-valuation dataset in Output Area j; θ_j is an Output Area fixed effect; and X_i is a set of characteristics of property i. Our pre-war property value in each 2001 Output Area is the exponent of the estimated Output Area fixed effect ($\exp(\theta_j)$) of this regression.

For our baseline results, we estimate this regression on properties that are either entirely residential or mixed use for comparability with our modern house price data, which is described in more detail in Section G7 of this Online Appendix and only covers residential buildings. To further increase the comparability with the modern house price data, our basic vector of control variables X_i mimics those available for the modern house price data: detached house, semi-detached house, terraced house, flat (where mixed-use valuations with both a residential and non-residential component are counted as flats). Table 1 in the main paper shows in Columns (5) and (6) specifications that use the average property value without the controls (X_i) and with these controls, respectively. We have also experimented with more detailed sets of control variables. For example, a flat might be a maisonette or a tenement, and a detached house could mean a nurses hostel or a presbytery. However including controls for every subtype results in output area fixed effects that are highly correlated with the results using our aggregated controls with a correlation coefficient of 0.95.

Figure 2 in the main text shows the distribution of our estimated property values in London in 1936 for the specification with our baseline controls. This map reveals the expensive residential areas in the center and West End of the LLC area, with other smaller clusters of higher value residential properties in many other parts of the built-up area.

G6 Post-War Census Data

Our post-war data on socioeconomic composition come from the 2001 UK Population Census, which is the first post-war census that counted a 100% sample of all questions on the census form. In robustness exercises, we also use data from the 2011 and 1981 UK Population Census. Below we first describe the data from the 2001 census and then those from the 2011 and 1981 censuses.

In contrast to many other countries, the UK does not allow access to individual level census data for research purposes and instead publishes census data aggregated to different spatial units. The smallest spatial units in which data is published are Output Areas (which were called

Enumeration Districts in older censuses), as discussed in more detail in Section G1 of this Online Appendix. Unless otherwise specified, we work with the Output Area level data.

G6.1 2001 UK Census

Data from the 2001 UK Census are available electronically from the "Nomis - Official Census and Labour Market Statistics" webpage at https://www.nomisweb.co.uk/.

Socioeconomic Status The 2001 census data report the usual resident population aged 16 to 74 by their socioeconomic classification. Individuals are classified into nine socioeconomic groups according to the National Statistics Socioeconomic Classification (NS-SeC), which is an occupationally-based classification that provides coverage of the whole adult population (https://www.nomisweb.co.uk/datasets/ks014a). Table G.8 shows how we aggregate the nine groups to reflect the three socioeconomic categories of low, middle and high-income, as observed in our pre-war data.

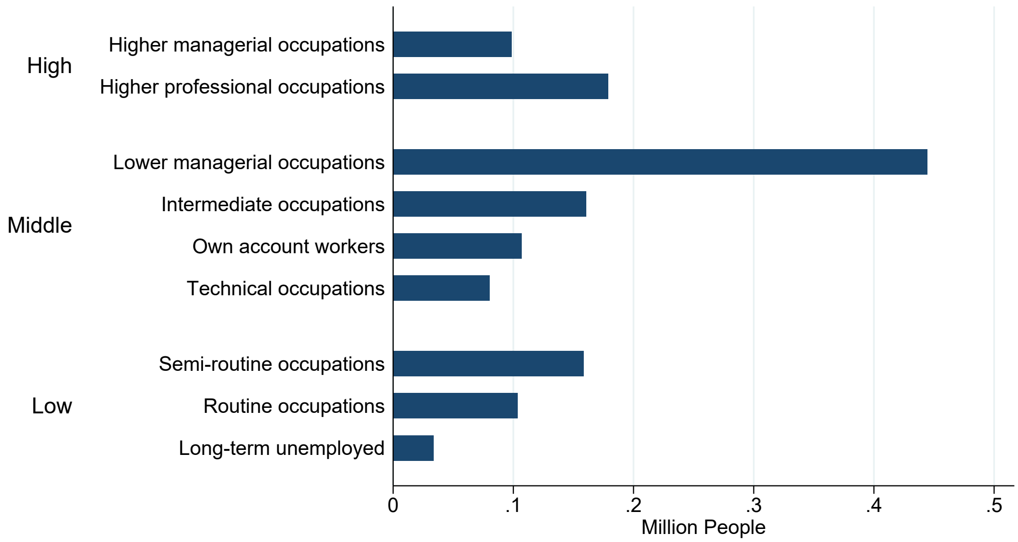
Table G.8: Post-War Socioeconomic Status Classification

Our Classification	National Statistics Socioeconomic Classification
High	Higher managerial, administrative and professional occupations
	Higher professional occupations
Middle	Lower managerial, administrative and professional occupations
	Intermediate occupations
	Own account workers
	Lower supervisory and technical occupations
Low	Semi-routine occupations
	Routine occupations
	Never worked and long-term unemployed

Note: National Statistics Socioeconomic Classification (NS-SeC) by occupation in the 2001 and 2011 population censuses. For further details, see: https://www.ons.gov.uk/methodology/classificationsandstandards/otherclassifications/thenationalstatisticssocioeconomicclassificationnssecrebasedonsoc2010.

In Figure G.21, we show the proportion of the population in each of the nine NS-SeC categories and our three aggregations of these categories (low, middle and high) in 2001. We find a relative stable share of the population across our three socioeconomic categories between the pre-war and post-war periods. In the 2001 census, the low socioeconomic status category (long-term unemployed and routine and semi-routine occupations) accounts for around 22 percent of the population; the middle socioeconomic category (lower-managerial, intermediate occupations, own account workers, technical occupations) makes up about 58 percent of the population; and the high socioeconomic category (higher managerial, higher professional) includes around 20 percent of the population, as shown in Table G.7 in Section G3 of this Online Appendix. In comparison, these low, middle and high categories account for 24, 47 and 28 percent of the population in our pre-war data on socioeconomic status, respectively.

Figure G.21: LCC Population by Socioeconomic Status in 2001



Notes: Population in millions in the London County Council (LCC) area in each of the nine NS-SeC categories in 2001 and our aggregation of these groups into our low, middle and high socioeconomic groups.

Employment The 2001 Census publishes employment data by NS-SeC group both for the resident and workplace population on the Output Area level.

Council Housing Tenants The 2001 Census publishes the number of households in each Output Area renting their accommodation from the local council.

Disclosure Protection Measures To prevent the inadvertent disclosure of information about identifiable individuals, a number of measures are applied to the 2001 UK Census data. These disclosure protection measures include:

- Factors involved in the design of output tables.
- Random record swapping in the output database.
- Application of thresholds to determine which areas can have results produced for them.
- Application of a small cell adjustment procedure to final tables to modify the values contained in small count cells.

- Conditions of use applied to all output products.

The two methods of particular importance to the quality of our data are random record swapping and the small cell adjustment procedure. The former consists of 'swapping' a sample of records with similar records in other geographical areas, the latter of 'adjusting the values of small cells up or down according to rules that say a proportion of the cells with that small value will be adjusted up, while the rest of the cells with that value will be adjusted down' (Office for National Statistics 2011a). The official document states that 'information on what constitutes a small cell count could not be provided as this may have compromised confidentiality protection.' However, documentation on the statistical disclosure regarding the flow data downloaded from the WICID database which are based on Census data states that 'small values are understood to be in the range 0-3. Cells with an initial value of 1 have been rounded to either 0 or 3, with 0 being the more likely result. Cells with an initial value of 2 have also been rounded to either 0 or 3, but with 3 being the more likely result. Cells with initial values of either 0 or 3 have retained these values although in each case it is impossible to distinguish between rounded values and genuine 0s or 3s.' (UK Data Service 2001). As our Census data do not contain any values equal to either one or two it is our understanding that the procedure outlined on the WICID webpage is the one that was applied.

G6.2 2011 UK Census

As a robustness exercise, we collect the socioeconomic status, employment and council housing data also from the 2011 Census. The Output Areas used to publish the 2011 Census are marginally different to those of the 2001 Census. The 2011 Census also improves on the statistical disclosure mechanisms in the 2001 Census by replacing the measures discussed above with 'targeted record swapping.' This measure entails calculating a household risk score for every household based on a small number of characteristics. A sample of households for swapping is then selected, with chances of selection being higher for households with higher risk scores. Once selected, the household is matched to a 'similar' household based on some basic characteristics and records are swapped. For a more detailed explanation, see Office for National Statistics (2011a).

G6.3 1981 UK Census

As discussed in Section 4.4 of the main paper, Margaret Thatcher's 1980 Housing Act gave council housing tenants the "right to buy" their home at significant discounts. The 1988 Housing Act also created the possibility for councils to transfer ownership of their housing stock to housing associations. The number of people living in council housing in the 2001 and 2011 censuses only captures people living in units that are still owned by local councils in the respective census year. Residents of council properties who bought their units under the "right to buy" scheme or whose units were transferred to housing associations are no longer counted as council housing residents. To explore the robustness of our results to these changes, we also collect data on the number of households living in council housing from the 1981 Census from the CASWEB webpage at https://casweb.ukdataservice.ac.uk/.

As discussed in Section 3 of the main paper, the 2001 population census was the first post-war census for which all census questions were counted for the full population, rather than for a 10% sample. Our main outcome variables on the socioeconomic composition of enumeration districts were only counted for a 10% sample in the 1981 population census, making these variables unreliable for small spatial units. In contrast, information on whether respondents live in council housing is available for the 100% sample of the 1981 population census and published for the 7,321 enumeration districts from that census year that fall within the LCC area.

G7 Post-War Property Values Data

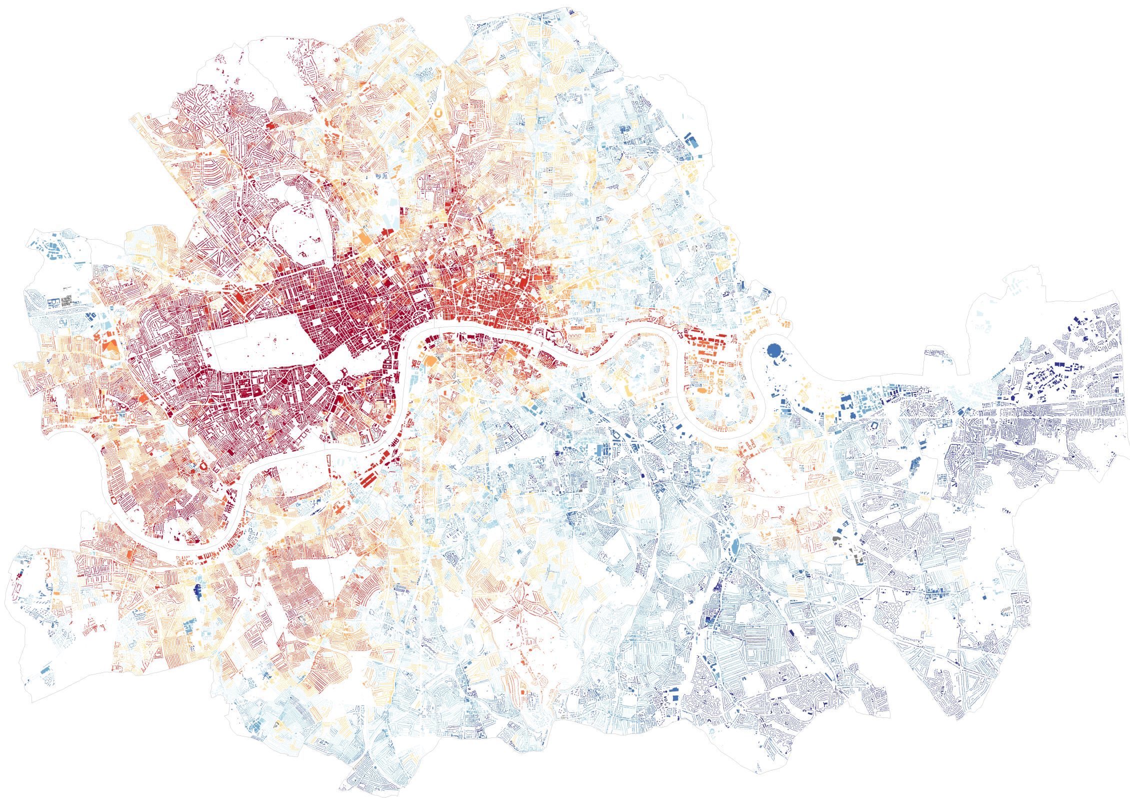
Our post-war property values data are based on property transactions data from the UK Land Registry. We begin by discussing the Land Registry data. We next summarize the matching of property transactions to 2001 Output Areas. Finally, we discuss the hedonic regressions that we use to measure average property prices for each 2001 Output Area.

Land Registry Data The Land Registry data include prices paid, postcodes and a range of characteristics for each property transaction. From 1995-2020, there are 1,186,317 transactions registered within the LCC area. Unlike the rateable value data in Section G5 of this Online Appendix, this Land Registry data only includes residential properties. It also excludes sales that were not at full market value (e.g., sales under subsidized government programs, gifts or compulsory purchases).

The Land Registry data includes information on a number of property characteristics, but does not report property size. Therefore we augment it with the dataset from Chi et al. (2021), which matches Land Registry transactions to Domestic Energy Performance Certificates (EPC) for the period from 2011-2019. Within this period, 74 percent of Land Registry transactions in the LCC area have been address matched, adding information on the floor space of the transacted property in square metres and the number of rooms. For a subsample of 226,000 transactions, we thus obtain information on property size.

Matching to 2001 Output Areas We match each property transaction to its 2001 Output Area using the centroid of the property's postcode. These postcodes cover a small geographical area and sometimes correspond to individual buildings. In the United Kingdom as a whole, there are

Figure G.22: Post-War Property Values in London



Note: Post-war property values for each 2001 Output Area; we estimate a hedonic regression of log property prices on observed property characteristics, Output Area fixed effects, and year fixed effects, using all property transactions from 1995 to 2019; the displayed post-war property values for each Output Area correspond to the estimated fixed effects from this hedonic regression.

1.8 million postcodes for around 30 million addresses. On average, there are 133 transactions per Output Area in the LCC area, and only 3 percent of these Output Areas have less than 10 transactions. For the subsample of transactions on which we have property size information, there are 25 transactions per Output Area, and 22 percent of these Output Areas have less than 10 transactions.

Hedonic Regressions We measure post-war property values using hedonic regressions that control for observed property characteristics. We estimate the following regression for the log transaction price ($\log \mathbb{P}_{ht}$) of property h at time t on observed property characteristics (X_{ht}), a fixed effect for the Output Area i in which the property is located (η_i), and year fixed effects (d_t):

$$\log \mathbb{P}_{ht} = X_{ht}\beta + \eta_i + d_t + u_{ht}, \tag{G.3}$$

where u_{ht} is the regression error.

We estimate this regression using all property transactions from 1995 to 2020. Observed property characteristics (X_{ht}) include the type of property (detached, semi-detached, terraced or flat), whether the property was newly built, and whether it was freehold or leasehold. Our baseline measure of the post-war property value for each 2001 Output Area is the exponent of the estimated fixed effect from this regression ($\exp\left(\widehat{\eta}_i\right)$). We also report robustness tests in which we estimate this hedonic regression for sub-periods.

As further robustness check, we estimate an analogous regression for the subset of the transactions that include floor space information, in which we use the log of the price paid per square meter instead of the log price paid as the dependent variable. Although the price paid per square meter data arguably provides a more accurate measure of the value of a unit of floor space, we find that the property price and the property price per square meter have a correlation coefficient of 0.92 after we condition on the observed property characteristics (X_{ht}). This finding suggests that most of the variation in floor space within an Output Area is accounted for by our observed property characteristics (including the distinction between house and flat). This pattern of results also validates our use of the estimated fixed effect from the regression with property price data and observed property characteristics as our baseline measure of post-war property values.

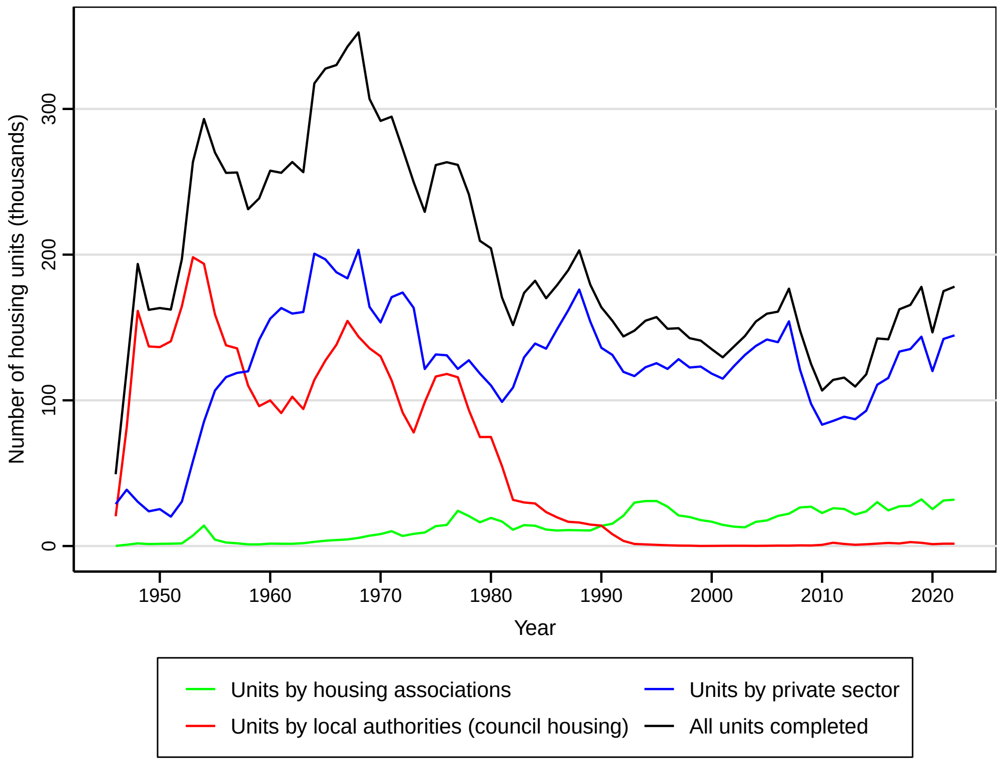
In Figure G.22, we display our baseline measure of post-war property values across 2001 Output Areas in the LCC area. We find a broadly similar gradient in property values as in our pre-war data, with higher values in the West End than in the East End, and substantial variation even with narrow geographical areas. Nevertheless, we find evidence of substantial changes in relative property values within narrow geographical areas, which we exploit in both our reduced-form and structural estimation in the paper.

G8 Post-War Housing Construction Time Series

Figures G.23 and G.24 show annual data from 1946 to 2022 on the number of housing units completed in England and Wales and the LCC, respectively. The graphs show the total number of housing units completed in each year and also a breakdown of the total number of housing units into three different types: (1) units constructed by local authorities (i.e., council housing) (2) units constructed by housing associations and (3) units constructed by the private sector.

The aggregate data for England and Wales is available from the UK governments online resource "Live Tables on Housing Supply", which are available at https://www.gov.uk/government/statistical-data-sets/live-tables-on-house-building. The data for the LCC area has been collected from multiple sources. Data for the period 1946 to 1965 are taken from the "Housing Service Handbook" published by the Greater London Council in 1966. Data for the period 1966 to 1980 are taken from "Local Housing Statistics – England and Wales" published by the UK Government Statistical Service in 1986. Data from 1980 to 2022 are available electronically in the Life Tables on Housing Supply referred to above. The data from 1966 onwards is reported for the modern boroughs of London as the LCC area ceases to be the administrative area of London in 1965. We

Figure G.23: Post-War Housing Construction in England and Wales

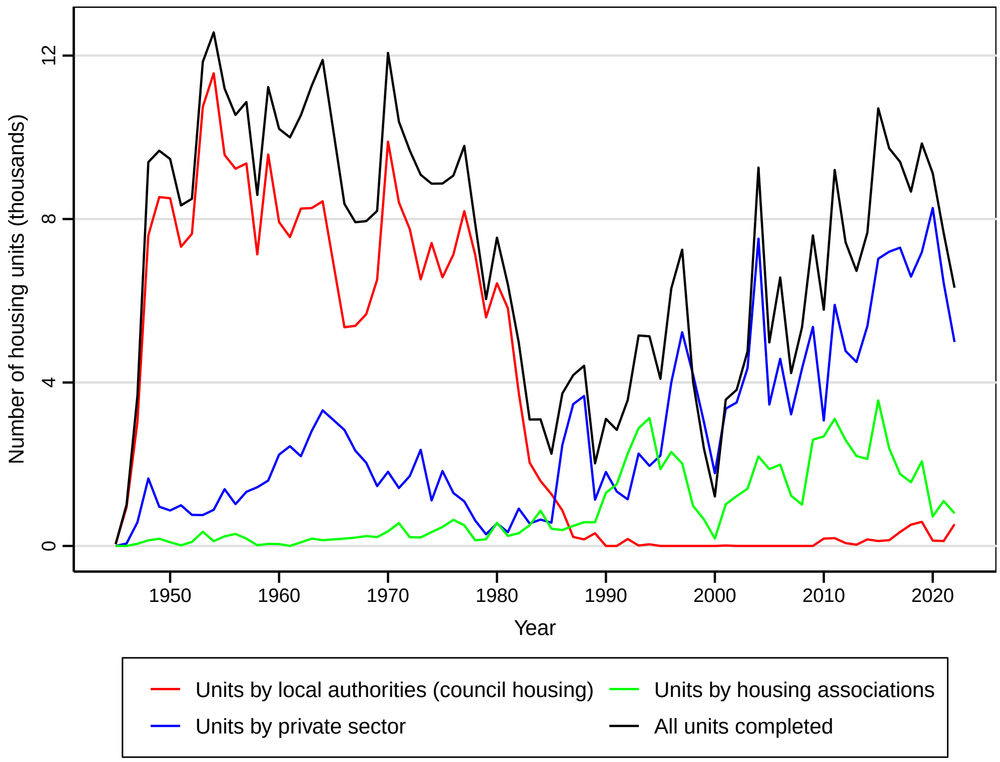


Note: The figure shows a line graph for the number of housing units completed in each year between 1946 and 2022 in England and Wales and its break down into units constructed by local authorities, housing associations and the private sector. The number of housing units are reported in thousands.

aggregate the data for the 13 modern boroughs that make up the LCC area to calculate the LCC totals for the post 1965 period.

The figures show two striking patterns. First, in the years immediately after the Second World War a substantial share of housing construction is in the form of council housing. For England and Wales 63% of all construction between 1946 and 1960 is council housing. For the LCC area this share is even higher with council housing making up 87% of all construction between 1946 and 1960. Even for a much longer post-war period from 1946 to 1980 the share of council housing in total construction remains substantial, accounting for 48% and 82% of housing construction in England and Wales and the LCC area, respectively. Second, there is a sharp drop in both council housing construction and also overall construction around 1980 after Margaret Thatcher comes to power in 1979. Construction only slowly recovers after the downturn with the private sector now accounting for the majority of new units completed.

Figure G.24: Post-War Housing Construction in the LCC Area



Note: The figure shows a line graph for the number of housing units completed in each year between 1946 and 2022 in the LCC area and its break down into units constructed by local authorities, housing associations and the private sector. The number of housing units are reported in thousands.

G9 Post-War Built-Up Area and Building Heights

Our data on the post-war built-up area and building heights are taken from the Ordnance Survey's "MasterMap Topography Layers" and are available electronically to academic users through Edina Digimap at: https://digimap.edina.ac.uk/. We downloaded this data in November 2014 and the data shows the footprint and height of every building in London at the time. The data reports two measures of the height of buildings, which are the overall height and the eave height. We work with the eave heights as our default measure of building height.

By overlaying the shapefile of the modern built-up area with our data of the pre-war built-up area from the bomb damage maps discussed in Section G2 of this appendix, we determine which modern buildings have the same footprint as the buildings shown on the bomb damage maps. If the historical and modern footprint of a building are identical on visual inspection, we code the variable "exists now" as one and as zero otherwise. It there are differences in the modern and historical built-up area, they tend to be substantial and it is therefore easy to spot such changes visually. Our variable "exists now" therefore captures changes in the footprint of buildings and will not capture buildings re-built within the same footprint after suffering wartime damage. In Figure G.8 of Section G2 of this appendix we show the "exists now" variable is intuitively related to the amount of wartime damage that an output area experienced.

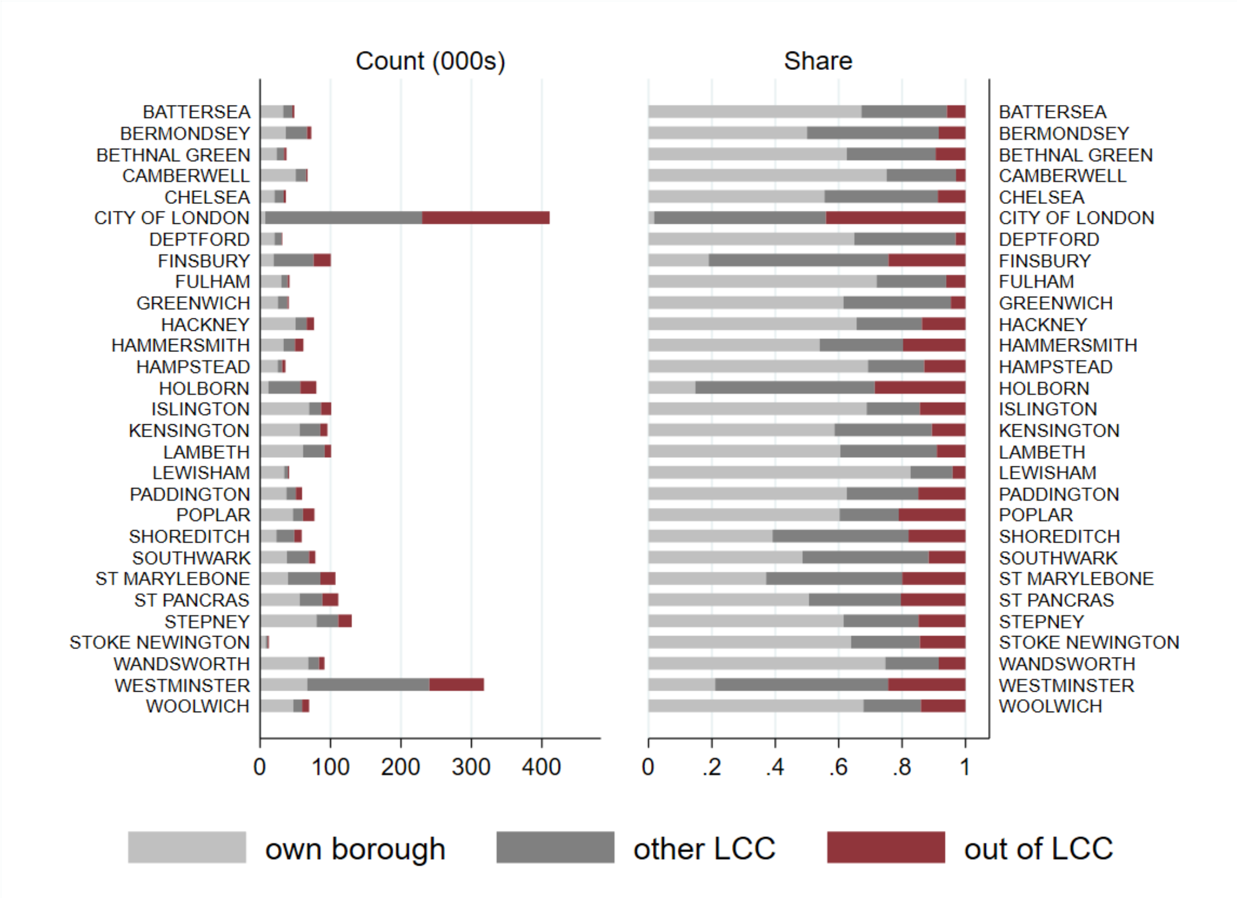
G10 Commuting Data

Pre-war Commuting Data The 1921 population census of England and Wales is the first census that reports bilateral flows of commuters between their residence and workplace boroughs. The 1921 census did so by asking respondents to report the address of their regular place of work. The same question was not repeated in the last pre-war census in 1931, likely in part due to pressures to reduce the costs of the census. The 1921 data was first published for London and the five Home Counties (Essex, Hertfordshire, Kent, Middlesex and Surrey) in Census of England & Wales 1921 (1923) "Workplaces in London and the Five Home Counties, Tables Part III (Supplementary)" and then for all boroughs in England and Wales in Census of England & Wales 1921 (1925) "Workplaces." We used the publication for London and the Home Counties as our main data source and used the later publication for all of England and Wales to obtain information on inflows of workers to London and the Home Counties from other parts of England and Wales.

The residence of a worker is the borough in which the person was located on census evening whether as a permanent resident or as a temporary visitor. The workplace of a worker is reported in four categories: (a) workers who work in the borough in which they were located on census evening; (b) workers who have no fixed workplace; (c) workers whose workplace is not known; (d) workers who work in another borough than the borough in which they were on census evening. Groups (b) and (c) are typically very small, and we assume that these workers work in their borough of residence. For workers with a workplace outside their residence borough, the census reports flows to each destination borough. Bilateral flows of less than 20 people are not reported for confidentiality reasons and are omitted. Summing these reported bilateral flows, the resulting sums of workplace employment and residence employment are close to the totals for workplace employment and residence employment (including flows of less than 20 people) that are separately reported in the population census.

We use data on bilateral commuting flows between the 29 boroughs of the London County Council (LCC) area, which is a subset of the larger number of boroughs considered in Heblich, Redding and Sturm (2020). For each of these 29 boroughs, we compute total gross in-commuting flows from residences outside of the LCC area, and total gross out-commuting flows to workplaces outside of the LCC area. We find that the gross out-commuting flows to workplaces outside the LCC area are small in magnitude for each borough, reflecting the fact that the areas outside the LCC boundaries were primarily either residential or agricultural in 1921. Therefore, for residents of the LCC as whole, 52 percent work in the same borough where they live, 45 percent work

Figure G.25: Total In-Commuting Flows in 1921 By Borough of Workplace and the Residential Origin of these Flows



Note: Left panel shows total gross in-commuting flows in 1921 for each LCC borough (in thousands) by the residential origin of those flows (within the same borough, from other LCC boroughs, and from outside of the LCC); right panel shows the share of these gross in-commuting flows for each borough by residential origin.

in another LCC borough, and only 3 percent work outside of the LCC. In contrast, gross in-commuting flows from residences outside the LCC boundaries are larger. For workers in the LCC area as a whole, 43 percent live in the same borough where they work, 37 percent live in other LCC boroughs, and 20 percent live outside the LCC.

In Figure G.25, we display total in-commuting flows in 1921 by borough of workplace and residential origin. For most of the 29 boroughs, in-commuting flows from residences outside the LCC boundaries are also relatively small. The two main exceptions are the City of London (the historical Roman city and commercial center) and Westminster (the center of Royal and parliamentary government). In Section D3 of this Online Appendix, we include an explicit correction for in-commuting from residences outside the LCC boundaries.

Post-war Commuting Data The pre-war commuting data described above only reports the total number of commuters between boroughs but does not break these flows down by socioeconomic group. The first UK census that reports commuting flows broken down by socioeconomic group is the 2011 Census of England and Wales. This data is available electronically for academic users from the WICID database at https://wicid.ukdataservice.ac.uk/ (UK Data Service 2011). The smallest spatial units for which this data is published are Middle Super Output Areas (MSOAs), which are aggregations of the Output Areas used for our main reduced-form results. There are 356 MSOAs in the LCC area in the 2011 census. The census reports the commuting flows between MSOAs for each of the nine National Statistics Socioeconomic Classification (NS-SeC) groups that we use to construct our post-war data on socioeconomic status. We aggregate commuting flows for these nine groups to the commuting flows of low, middle and high-income workers using the correspondence shown in Table G.8 of Section G6.1 of this Online Appendix. The commuting flow data is subject to the same confidentially adjustments as the rest of the 2011 Census data, as discussed in Section G6.1 of this Online Appendix.

G11 Commuting Times

Pre-war Commuting Times To calculate pre-war travel times by public transport we use the transport networks developed by Heblich et al. (2020). That paper constructs detailed transport networks for London for census decades from 1801 to 1921. This transport network includes shapefiles for underground and overground railway lines, which are based on the shapefiles for the entire UK produced by the Cambridge Group for the History of Population and Social Structure, which digitized the work of Cobb (2003). Heblich et al. (2020) also constructed shapefiles of the omnibus and tram network in London for census decades from 1801 to 1921.

We use the railway, underground and omnibus network for 1921 from Heblich et al. (2020) to compute the shortest pre-war public transport travel time between each of the 9,041 Output Areas of the 2001 Census that are within the LCC boundaries. For connections where walking is faster than public transport, walking times are used as the fastest transport connection. In these calculations we follow Heblich et al. (2020) in allowing workers to walk in a straight line from the centroid of an output area to any railway or underground station or bus stop within 10 km, since the nearest station or stop may not be on the fastest bilateral connection. We also assume that changing modes of transport, such as changing from a bus to an underground line, incurs a wait time of 3 minutes.

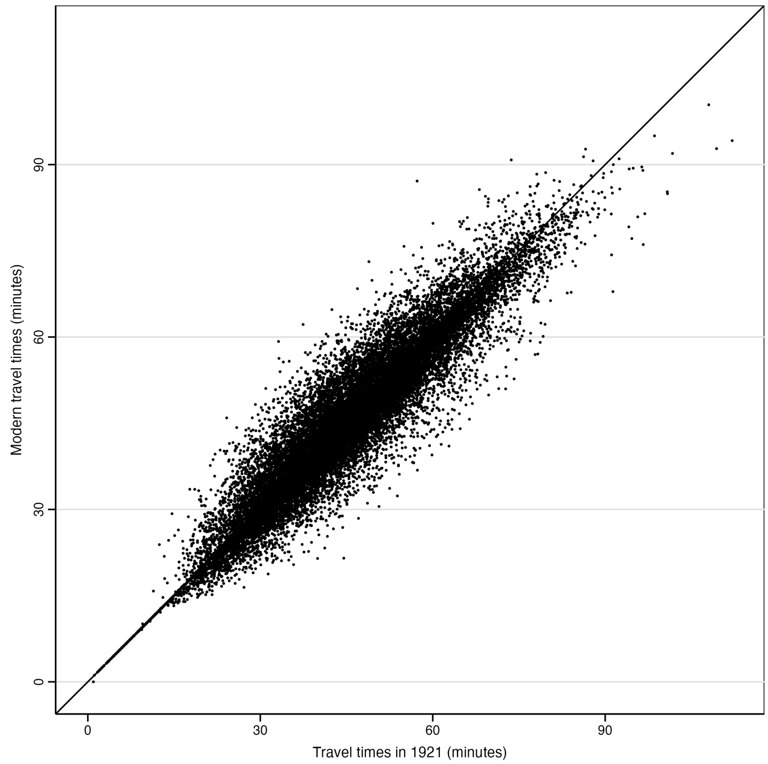
Post-war Commuting Times To compute post-war public transport travel times between the 2001 Output Areas we have updated the transport networks used in Heblich et al. (2020) as described below. We then use this updated network combined with data about the modern speed of travel to compute modern public transport travel times between all 2001 Output Areas.

Both the underground and rail network and also the bus network have changed since 1921. To capture these changes, we first use OpenStreetMap to add new railway and underground lines that have been constructed in the post-war period. On the underground system, the main additions were small extensions of the Northern line in the 1920s, the Central line extension past Liverpool Station to Stratford (opened in 1947), the Victoria line (opened in 1968) and the Jubilee line extension to Stratford (opened in 1999). From 1987 the Dockland Light Rail (DLR) system has operated in East London. The DLR system expanded several times over the next decades with the last smaller change to the network in 2011. We do not include the Elizabeth line, which was only opened in 2022.

There have also been a number of closures of rail lines and stations that were present in 1921. We use the information in Cobb (2003) to determine which stations and line segments from the 1921 network have been closed, and remove them from our post-war network. Finally, the bus network today is substantially different from the bus network in 1921. We use OpenStreetMap to create a shapefile of all London bus lines and bus stops in operation in 2023 and assume that this network has not changed materially over the last few years.

We have collected information on the modern speed of travel on the underground, railways and buses in the LCC area. For each underground line, we determined the two stations furthest apart but still inside the LCC area and determined the 2023 travel time and distance between these two stations using the Transport for London (TfL) journey planner available at https://api.tfl.gov.uk/journey. The unweighted average speed of underground lines from this data is 17.8 miles per hour and we use this average value for all underground lines. We use the same approach to determine the average speed of buses and railways in the LCC area. The unweighted average speed of bus lines in the LCC area using this method is 7.0 miles per hour, while the unweighted speed of rail lines is 20.5 miles per hour. Overall, these speeds are only marginally different from the speeds in 1921 from Heblich et al. (2020)

Figure G.26: Comparison of 1921 and Modern Commuting Times for a Random Sample of Bilateral Connections Flows



Note: The figure compares the travel times in 1921 with the modern travel times for a random sample of 25,000 bilateral connections in the LCC area. Travel times are computed in minutes using a combination of public transport and walking as described in more detail in the main text.

Using the updated network and travel speeds, together with the assumptions about walking to stations and waiting times to change modes of transport discussed above, we compute modern travel times for the LCC area. Figure G.26 shows the correlation between the 1921 and modern travel times for a random sample of 25,000 bilateral connections out of the roughly 81 million bilateral travel times between our 9,041 Output Areas. The 1921 and modern travel times are highly correlated with a correlation coefficient of 0.93. The median change in travel times between the pre-war travel times and the modern travel times across all bilateral connections in the LCC is a marginal increase in travel times of 0.1 percent. There is some variation around this median change, with a decrease in travel times by 11 percent at the 10th percentile, and an increase in travel times of 14 percent at the 90th percentile.

While commuting to work in private cars was virtually absent in London in 1921, there is some commuting to work by car today. Office for National Statistics (2011b) “Special Workplace Statistics: WU03UK” reports for each of the 13 modern boroughs in the LCC area the mode of transport of people with a workplace in these boroughs. The employment-weighted average percentage of commutes by car for people working in these 13 boroughs is only 15.8 percent. This is a much lower share of commuting by car than in many other cities and we therefore abstract from commuting by car.[21]

[21]The same publication reports the number of people cycling to work and this mode choice is in small single-digit percentages for the 13 modern boroughs in the LCC area, which is consistent with the vast majority of workers commuting by public transport or walking, as captured by our travel time calculations.

References

Ahlfeldt, G., S. Redding, D. Sturm and N. Wolf (2015) “The Economics of Density: Evidence from the Berlin Wall,” *Econometrica*, 83(6), 2127-2189.

Allen, Treb, Costas Arkolakis and Xiangliang Li (2024) “On the Equilibrium Properties of Spatial Models,” *American Economic Insights*, forthcoming.

Antràs, Pol, and Hans-Joachim Voth (2003) “Factor Prices and Productivity Growth During the British Industrial Revolution,” *Explorations in Economic History*, 40, 52-77.

Bertrand, M., E. Duflo, and S. Mullainathan (2004) “How Much Should We Trust Differences-in-Differences Estimators?” Quarterly Journal of Economics, 119, 249-275.

Bohr, Clement, Martí Mestieri and Frederic Robert-Nicoud (2023) “Heterothetic Cobb-Douglas: Theory and Applications,” UPF mimeo.

Booth, C. (1889) *Life and Labour of the People in London*, London: Macmillan

Campbell, C. (2012) *Target London*, London: Little-Brown.

Chi, Bin, A. Dennett, T. Oléron-Evans and Robin Morphet (2021) “A New Attribute-Linked Residential Property Price Dataset for England and Wales, 2011-2019,” *UCL Open Environment*, 2(07), https://dx.doi.org/10.14324/111.444/ucloe.000019.

Clarke, R. D. (1946) “An Application of the Poisson Distribution,” *Journal of the Institute of Actuaries*, 72, 481.

Cobb, M. H., “The Railways of Great Britain: a Historical Atlas”, (Shepperton : Ian Allan Publishing, 2003).

Conley, T. G. (1999) “GMM Estimation with Cross Sectional Dependence,” *Journal of Econometrics*, 92, 1-45.

Combes, Pierre-Philippe, Gilles Duranton and Laurent Gobillon (2019) “The Costs of Agglomeration: House and Land Prices in French Cities,” *Review of Economic Studies*, 86, 1556-1589.

Dericks, G. H. and H. R. Koster (2021) “The Billion Pound Drop: The Blitz and Agglomeration Economics in London,” *Journal of Economic Geography*, 21(6), 869-897.

Dingel, J. I. and F. Tintelnot (2023) “Spatial Economics for Granular Settings,” *NBER Working Paper*, 27287.

Epple, D., B. Gordon and H. Sieg (2010) “A New Approach to Estimating the Production Function for Housing,” *American Economic Review*, 100(3), 905-924.

Fetzer, T. (2023) “Regulatory Barriers to Climate Action: Evidence from Conservation Areas in England,” *CAGE Working Paper*, 654.

Gazeley, I., Newell, A., Hawkins, M., Walker, J., Scott, P. (2016) “Living Standards of Working Households in Britain, 1904-1954,” UK Data Service. SN: 7916, http://dx.doi.org/10.5255/UKDA-SN-7916-1.

Heblich, S., S. Redding and D. Sturm (2020) “The Making of the Modern Metropolis: Evidence from London," *Quarterly Journal of Economics*, 135(4), 2059-2133.

Holland, J. (2012) *The Battle of Britain: Five Months That Changed History*, May-October 1940, London: St. Martin’s Griffin.

Jackson, J. (2004) *The Fall of France: The Nazi Invasion of 1940*, Oxford: Oxford University Press.

Johnson, D. (1981) *V-1 V-2: Hitler’s Vengeance on London*, New York: Stein and Day.

Katz, Lawrence F. and Kevin M. Murphy (1992) "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 107(1), 35-78.

Kreindler, Gabriel E. and Yuhei Miyauchi (2023) "Measuring Commuting and Economic Activity Inside Cities with Cell Phone Records," *Review of Economics and Statistics*, 105(4), 899–909.

London County Council (1921) *London Statistics*, London: London County Council.

Matthews, Robert C. O., Charles H. Feinstein, and John C. Odling-Smee (1982) *British Economic Growth: 1856-1973*, Oxford: Oxford University Press.

Mckinstry, L. (2014) *Operation Sea Lion*, London: The Overlook Press.

Llewellyn Smith, Sir Hubert (1930) *The New Survey of London Life and Labour*, London: P. S. King & Son.

Office for National Statistics (2011a) *Statistical Disclosure Control for 2011 Census*, London: Office for National Statistics.

Office for National Statistics (2011b) *Special Workplace Statistics: WU03UK*, London: Office for National Statistics.

O'Brien, Terence H. (1955) *Civil Defence*, London: Her Majesty's Stationery Office.

Orford, S., D. Dorling, R. Mitchell, M. Shawn and G. D. Smith (2002) "Life and Death of the People of London: a Historical GIS of Charles Booth's Inquiry," *Health and Place*, 8, 25-35.

Parliamentary Papers (1887) "Tabulations of the Statements Made by Men Living in Certain Selected Districts of London in 1887," London: House of Commons.

Ray, J. (2004) *The Night Blitz: 1940-1*, London: Booksales.

Rosenthal, S. S. and W. C. Strange (2004) "Evidence on the Nature and Sources of Agglomeration Economics," in *Handbook of Regional and Urban Economics*, (eds) J. V. Henderson and J. F. Thisse, Vol 4. Amsterdam: Elsevier North Holland.

Santos Silva, J. M. C. and Tenreyro, Silvana (2006) "The Log of Gravity," *Review of Economics and Statistics*, 88(4), 641-658.

Saiz, Albert (2010) "The Geographic Determinants of Housing Supply," *Quarterly Journal of Economics*, 125(3), 1253-96.

Seltzer, A. and J. Wadsworth (2023) "The Impact of Public Transportation and Commuting on Urban Labor Markets: Evidence from the New Survey of London Life and Labor, 1929-32," *Explorations in Economic History*, forthcoming.

Severen, Christopher (2023) "Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification," *Review of Economics and Statistics*, 105(5), 1073-1091.

Shaw, Liam P. and Luke F. Shaw (2019) "The Flying Bomb and the Actuary,"' *Significance Magazine*, October, 13-17.

Tsivanidis, Nick (2023) "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio," *American Economic Review*, forthcoming.

UK Data Service (2001) *2001 Census: Special Workplace Statistics*, https://wicid.ukdataservice.ac.uk/cider/info.php?datatype=25.

UK Data Service (2011) *2011 Census: Special Workplace Statistics.*

Ward, Lawrence (2016) *The London County Council Bomb Damage Maps, 1939-1945*, London: Thames and Hudson.

White, Jerry (2014) *Zeppelin Nights: London in the First World War*, London: Random House.

Wooldridge, Jeffrey M. (2014) "Quasi-Maximum Likelihood Estimation and Testing for Nonlinear Models with Endogenous Explanatory Variables," *Journal of Econometrics*, 182, 226-234.