Tax Incentives and the Supply of Low-Income Housing

Evan Soltas*

October 13, 2023

(Click for Most Recent Version)

Abstract

Subsidies to developers are a core instrument of housing policy. How do they affect housing markets, and who benefits? I assess their impacts and incidence with a dynamic model and new data on developers competing for Low-Income Housing Tax Credits. I estimate the model using three sources of variation: quasi-random assignment of subsidies, shocks to subsidy generosity, and nonlinear incentives to reduce rents. I find that, due to displacement of unsubsidized housing, subsidies add few net units to the housing stock and instead reallocate units progressively. Households benefit from developer competition for subsidies, but competition also results in high entry costs, and developers still capture nearly half of the welfare gains. In counterfactuals, a stylized voucher program can generate the same household benefits at less fiscal cost.

*MIT Department of Economics and Federal Reserve Bank of Boston. Email: esoltas@mit.edu. This project would not have been possible without the generous assistance of state housing finance agency staff. I am greatly indebted to Jim Poterba, David Autor, and Jon Gruber for their advice and support, and also to Paul Willen for facilitating data access. For helpful comments and conversations, I thank Rebecca Diamond, Denise DiPasquale, Ingrid Gould Ellen, Colin Higgins, Mike Hollar, Allan Hsiao, Dev Patel, Charlie Rafkin, Anna Russo, Tobias Salz, Ben Sprung-Keyser, Bill Wheaton, and participants in the MIT Public Finance Lunch. Umang Bansal, Cristine Bren, and Jayson Roxas provided excellent research assistance. I gratefully acknowledge support from the Jerry A. Hausman Dissertation Fellowship, the Lincoln Institute of Land Policy’s C. Lowell Harriss Dissertation Fellowship, the Lynde and Harry Bradley Foundation, the George and Obie Shultz Fund, and the National Science Foundation Graduate Research Fellowship under Grant No. 1122374. The views expressed in this paper are solely those of the author and not necessarily those of the Federal Reserve Bank of Boston nor the Federal Reserve System.
1 Introduction

Housing policy—and the housing markets it shapes—plays an important role in the lives of lower-income Americans. At the center of U.S. federal housing policy since the 1970s are two ways to subsidize rents: “tenant-based” and “project-based” assistance. Unlike public housing or “tenant-based” vouchers, project-based programs work through developers and the supply side of housing markets, offering incentives to build, renovate, and operate housing for lower-income tenants at regulated rents. In 2022, about seven million U.S. households received some form of housing assistance, more than half through “project-based” programs.

The largest project-based subsidy, the Low-Income Housing Tax Credit (LIHTC), has funded one in five of all new U.S. multifamily units since 1987 through grants. About two percent of all U.S. households lived in LIHTC units in 2022, more than received rent vouchers in the same year and more than lived in public housing at its historical peak. Structured as a corporate income tax expenditure, the LIHTC reduces annual federal receipts by about $10 billion. Its rise as a share of the housing stock is set to continue, as its share of new units exceeds its share of existing units, driving a momentous but little-discussed change in the direction of U.S. housing policy.

This shift toward project-based subsidies has unfolded without any real progress in a longstanding debate about its merits relative to tenant-based programs. Project-based assistance offers the potential for spatially-targeted investment, and it is more easily combined with supportive services, helping populations with housing needs that private rental markets can struggle to meet. Yet economists are generally skeptical of project-based programs, often due to concerns about their incidence. Glaeser and Gyourko (2008), for instance, argue the LIHTC “essentially functions as a transfer program” to developers, and Quigley (2011) writes project-based subsidies “must be justified on some other basis” than redistribution to lower-income households. These concerns closely resemble ones that prompted the introduction of vouchers fifty years ago, in great part because they have faced little empirical scrutiny since then.1

This paper revisits the impacts and incidence of project-based assistance using newly-collected data on the applications of developers competing for the LIHTC. I use the responses of developers to three sources of variation—quasi-random assignment of subsidies, changes in subsidy generosity, and variation in incentives to offer lower rents—to estimate a dynamic model of developer behavior in the markets for subsidized and unsubsidized housing. These analyses show that few applicants are marginal to development, that many are marginal to applying, and that applicants’ proposed rents respond to incentives. The model rationalizes their behavior through high outside options, high values of winning, and high costs to enter the competition. Using the estimated model, I

---

1The latest review of research on U.S. housing policy (Olsen and Zabel, 2015) writes that there is “no [recent] high-quality evidence on cost-effectiveness” and describes this topic as the “highest priority for housing policy research.”
find the LIHTC does little to expand the housing stock on net and mostly displaces unsubsidized housing, but in doing so, it reallocates units to lower-income households. Overall, my results give a mixed review of project-based subsidies in terms of their incidence. Households benefit from developer competition for subsidies, but entry costs are high, and so developers still capture nearly half of the welfare gains. In counterfactuals that compare the LIHTC to a stylized voucher program, I find the same household benefits could be achieved at less fiscal cost. Despite these weaknesses of project-based subsidies, their coexistence with tenant-based subsidies in U.S. housing policy appears broadly reasonable in light of their context-specific advantages and disadvantages.

Section 3 presents the model of developer behavior. In the model, each developer is associated with a land parcel and may enter it into a grant competition. The developer’s problem is intrinsically dynamic. As an alternative to entry, they can develop the parcel for a private use or wait, possibly to apply or to develop privately in the future—and they may also reapply in future competitions if they lose. Developer behavior is shaped by three primitive objects: their value of winning the subsidy, their outside option, and their cost of applying for the subsidy. Households rent the entire housing stock, with welfare impacts occurring through both rent savings in below-market units and the general-equilibrium effects on market-rent units. This aspect of the model echoes the two-sector incidence analysis of Harberger (1962): Project-based subsidies create a market for subsidized housing, in which resources are allocated by grants, that is linked to the unsubsidized market through developer and household substitution. Developers take win probabilities as given, and equilibrium follows from profit maximization, market clearing, and rational expectations.

New data, introduced in Section 2, enable my analysis. Through public-information requests, I collected administrative records on 453 rounds of LIHTC competitions from 40 states, covering 22,241 applications from 2005 through 2019 with requested subsidies of $200 billion in total. I link these records to parcel-level tax assessments and neighborhood-level outcomes. Finally, I code states’ grant rules from regulatory documents, which I use to compute ex-ante probabilities with which developers could expect to win by simulating the mechanism. These rich data enable careful study of the key actor in project-based subsidies: the developer.

Initial tabulations of the data immediately raise questions about the subsidy’s impacts and incidence. Average rents in LIHTC units are only about 12 percent below my estimates of market rents for new units in the same neighborhoods, though the rent savings vary greatly. Remarkably, this appears to be the first national estimate of the LIHTC rent discount. At the same time, I also find that LIHTC tenants are much poorer than the likely tenants of counterfactual developments. To understand how the LIHTC affects developers, I turn to the three quasi-experimental analyses.

---

2My estimates appear consistent with expert knowledge. Relying on their policy experience, Khadduri and Wilkins (2008) write that LIHTC rents are “indistinguishable from market rents or only slightly below market.” An analysis of LIHTC properties in Tallahassee, Florida finds similar rent savings (Burge, 2011).
I show how each source of variation is linked to model primitives, and I then structurally estimate the model so as to jointly match the quasi-experimental moments.3

The first empirical strategy (Section 4) uses quasi-random assignment of subsidies to estimate the causal effects of winning on development activity. Developers apply under uncertainty about rivals, making subsidy assignment as good as random conditional on an estimated win probability. This strategy is drawn from analyses of auctions (Hortaçsu and McAdams, 2010) and is related to a broader literature on treatment-effect estimation in mechanisms (e.g., Abdulkadiroğlu et al., 2017). I find small causal effects of winning on development: 75 percent of applicant parcels would be developed within ten years if they were to lose. However, the LIHTC does pull development forward in time, and it replaces some market-rate units with (slightly) subsidized ones. These results are confirmed by neighborhood-level event studies that compare areas where winners and losers applied with similar ex-ante win probabilities. From the model’s point of view, such behavior suggests most applicants have profitable outside options.

The second empirical strategy (Section 5) estimates application supply responses to changes in subsidy generosity. Through an event study and regression discontinuity design (RDD), I exploit annual policy variation by Census tract in the LIHTC’s value. This variation is induced by a cutoff-based rule whose inputs are measured with substantial sampling error. Both empirical approaches find applications are responsive to subsidy generosity: A ten-percent reduction in the net-of-tax price of low-income housing development leads to an increase in developer applications of at least three percent. The model-based implication is that entry costs are high, as a substantial mass of developers applies only when the government assumes almost the entire construction cost of low-income housing, even when these units command near-market rents.

The third empirical strategy (Section 6) infers the value to developers of winning the LIHTC from their rent-setting behavior. In their application, a developer may “bid” a lower rent than the maximum allowed under federal regulations, so as to raise their win probability. Most state rules create strong and nonlinear incentives to reduce rents. Developers appear highly responsive to these incentives, bidding rents down and bunching sharply at kink points. Such behavior implies substantial developer profits from winning, as they would otherwise be reluctant to trade rental income for a higher win probability on the margin. To quantify profit incidence, I simulate unilateral deviations to different rents and obtain developer valuations through bid inversion. I find developers are indifferent on average between a 0.9-percentage-point increase in win probability and a $1,000 increase in present-value rental income per unit. These findings imply a developer incidence share of about 45 percent, matching the structural estimates that incorporate dynamic considerations.

3Holmes and Sieg (2015) describe this approach as a way to unite the virtues of the quasi-experimental and structural approaches: The former credibly estimates behavioral responses, while the latter maps this description of behavior into the primitives of an equilibrium model.
I use these three sets of results to estimate the model in Section 7, adopting methods from dynamic discrete choice (Rust, 2000; Sweeting, 2013) and indirect inference (Gourieroux et al., 1993). Each quasi-experimental analysis is motivated by a conceptual link to one of the three primitives: outside options for the causal effects of subsidy awards, entry costs for responses to subsidy generosity, and win values for the trade-off between rental income and win probability. In estimation, the key econometric challenge is sample selection: I observe applicants, not potential applicants, and moreover the counterfactual development decisions of losing applicants only. The quasi-experimental analyses overcome both selection issues. Policy variation in subsidy generosity is an instrument for applicant self-selection, and quasi-random subsidy assignment addresses the selection of winners on unobservables. Both the problem and the solution are analogous to Walters (2018) and Van Dijk (2019), who also study self-selection into mechanisms.

Section 8 reports the structural results. Estimates of model primitives reconcile the three empirical analyses with a simple explanation: Many applicants are marginal to application but not to development, with large and heterogeneous entry costs deterring non-applicants. Such estimates imply the LIHTC is ineffective in expanding the overall housing stock, although it appears to reallocate it progressively. For every ten LIHTC units, I find eight units displace private housing that would have otherwise been built, and two units are net additions to the housing stock. Applying these displacement estimates to the LIHTC in aggregate, I conclude the LIHTC has expanded the U.S. housing stock by about 500,000 units, or by 0.4 percent. Due to displacement, the fiscal cost of the LIHTC is about $1 million per net new unit on average.

I also estimate that households capture about 31 percent of the welfare gains from the LIHTC, in a clear rejection of the null hypothesis of full incidence on developers. These gains arise mostly from below-market rents rather than general-equilibrium effects (23 and 8 percent of incidence): Below-market rents are a transfer to LIHTC tenants, and by reducing market rents, the LIHTC aids other households. Developer competition thus succeeds at least partly in its aim of driving incidence to households. Yet the household benefits of developer competition are diluted in two ways: Developers do capture significant subsidy (44 percent of incidence), and much of the LIHTC’s value is competed away by the entry costs that developers pay to apply (25 percent of incidence).

Is this a lot or a little incidence on households? To provide a benchmark, I implement counterfactuals that compare the LIHTC to a stylized voucher program. I find vouchers could provide the same welfare benefit to households as the LIHTC at a 25-percent fiscal savings. Conversely, a balanced-budget reform that shifts from project- to tenant-based assistance would expand housing supply and raise household welfare. These findings reflect two considerations. First, these subsi-
dies have opposite-signed pecuniary externalities on the unsubsidized market. To offset effects on unsubsidized households, vouchers would require more spending to achieve the same household benefit, all else equal. The second difference between the two policies is that vouchers reduce developer incidence and eliminate entry costs. On net, the balance of these forces favors vouchers, but the disadvantage of project-based assistance is surprisingly modest.

This paper contributes to several literatures in public finance and urban economics. The developer's choice to take up the subsidy imposes a participation constraint—a supply of subsidized housing—on the social problem of optimal housing policy (Soltas, forthcoming). Understanding this constraint is a natural starting point for public finance to inform housing policy, given its shift from direct provision via public housing to market support via subsidies (Collinson et al., 2015). This paper pursues this agenda for the largest U.S. project-based housing policy, bringing together a new model, new data, and new sources of quasi-experimental variation.

The incidence and effectiveness of project-based assistance are classic issues in housing policy. Early research was critical of public housing and project-based programs and proved influential in the rise of rent vouchers. Despite the LIHTC's ascendance, these issues have largely not been revisited. Most recent research on project-based subsidies, and the LIHTC in particular, studies its amenity effects and the "crowding-out" of unsubsidized housing (Sinai and Waldfogel, 2005; Baum-Snow and Marion, 2009; Eriksen and Rosenthal, 2010; Freedman and Owens, 2011; Davis et al., 2019; Diamond and McQuade, 2019). My analysis connects displacement to its welfare implications, illuminates why this literature finds significant displacement of unsubsidized housing, and suggests that housing is being reallocated despite overall displacement. Economists often argue for vouchers over project-based assistance (Aaron, 1972; Rosenthal, 2014), although views are diverse (Favilukis et al., 2023).

Urban economists have recently adopted empirical techniques from industrial organization and mechanism design (Murphy, 2018; Diamond et al., 2019; Waldinger, 2021; Calder-Wang, 2022; Almagro and Domínguez-Iíno, 2022; Hsiao, 2022). This paper is also related to work that uses housing-policy variation to study housing supply (Gyourko, 2009). In particular, my approach to combining a dynamic model and quasi-experimental evidence may be well-suited for other housing-supply topics, as policy variation abounds but dynamic considerations loom large. Three recent and closely related papers find developer responses to zoning reform in Brazil (Anagol et al., 2021), to French housing investment tax credits (Levy, 2021), and to French housing quality standards (Levy, 2023). By modeling the project pipeline from application to completion, the framework may also be useful to study grants in other contexts (e.g., Jacob and Lefgren, 2011).

---

*See Olsen (2003), Weicher (2012), and Olsen and Zabel (2015) for discussions of this research.

†Other research has considered the effects of subsidized housing on resident outcomes (Currie and Yelowitz, 2000; Chyn, 2018). Ellen et al. (2016), Derby (2021), and Sportiche (2022) specifically study LIHTC properties.
2 Setting and Data

2.1 The Low-Income Housing Tax Credit

Established in the Tax Reform Act of 1986 and now defined in Section 42 of the Internal Revenue Code, the LIHTC is an investment tax credit issued to developers against federal corporate income tax liabilities.\(^8\) The credits reduce tax liabilities over ten years by a specified share of a project’s “qualified basis.” The basis typically includes construction costs associated with low-income units, as well as a developer fee, but it excludes land and predevelopment expenses.

**Context.** Figure 1 depicts the evolution of three leading housing programs as shares of the U.S. housing stock: public housing, rent vouchers, and the LIHTC. The era of public housing prevailed from the Great Depression until 1973, when it was ended by a construction moratorium and a watershed report, the National Housing Policy Review (1973), which found that public housing was not cost-effective. A rapid transition to tenant-based subsidies followed, with sporadic use of project-based assistance (Orlebeke, 2000). By the 1990s, debates over the proper form of housing assistance were viewed as definitively resolved in favor of vouchers (Winnick, 1995). The rise of the LIHTC has quietly overturned this consensus. Since 2000, the LIHTC’s growth has coincided with the stagnation of vouchers, reorienting policy toward project-based subsidies.

**Applications.** Each year, state housing finance agencies may issue credits up to a per-capita maximum. In 2022, the maximum is $2.60 in ten-year credits per state resident, with an additional allowance for small states, amounting to an annual nationwide tax expenditure of about $10 billion (Joint Committee on Taxation, 2023). Agencies award credits through competitions in which they receive proposals and select some for funding.

Each agency awards funding according to its Qualified Allocation Plan (QAP), a public document of selection criteria. In 2019, all but two states used numerical rubrics that prevent discretion at the level of individual applications. An application \(i\)’s tax-credit assignment depends upon its QAP score \(q_i\) and set-asides \(z_i\), along with those of other applications \((Q_{-i}, Z_{-i})\). The primary purpose of set-asides is to balance the distribution of subsidies over geographies and demographic constituencies (e.g., seniors). The assignment \(W_i \in \{0, 1\}\) is thus characterized by

\[
W_i = W(q_i, z_i; Q_{-i}, Z_{-i}),
\]

where I refer to \(W(\cdot)\) as the “grant rule.” I programmed the grant rules for each round in my data

\(^8\)Throughout this paper, I focus on the “9%” (also known as “70%”) LIHTC, which is awarded competitively. Another program, the “4%” (a.k.a. “30%”) LIHTC, funds developments at a lower credit rate. In general, this program funds rehabilitation projects, rather than new construction, usually with tax-exempt bonds and without a competitive process. Collinson et al. (2015) provide a broader review of the LIHTC alongside other low-income housing programs.
Figure 1: Three Waves of Postwar U.S. Housing Policy

Percentage of Households

- Public Housing
- Section 8 Voucher
- Low-Income Housing Tax Credit

Notes: This figure displays the assisted shares of households for three housing policies: public housing (built under Section 9 of the U.S. Housing Act of 1937), tenant-based rental assistance (Section 8 vouchers and certificates), and project-based developments financed by Low-Income Housing Tax Credits (9% and 4% LIHTC). Annual assisted unit counts are drawn from Olsen (2003) and Vale and Freemark (2012). I extend the data to the present using the HUD Picture of Subsidized Households and LIHTC Databases. I also adjust for incomplete reporting in the final three years of the LIHTC time series. Household counts are from Census Table HH-1.

following my reading of the QAP.

Scoring rules \( q_i = q(x_i) \) determine scores from application characteristics \( x_i \), which I take to contain \( z_i \) and other variables. These rules are complex and differentiated across states (Shelburne, 2021), as federal regulations allow them to be “appropriate to local conditions.” Common criteria include site selection, building amenities, and developer characteristics. Developers are thought to “‘chase points’ when they make decisions” (Ellen and Horn, 2018). A key federal stipulation on QAPs is that they must incorporate a preference for “projects serving the lowest income tenants.” As U.S. housing policy links income and rent levels, states meet this requirement with rules favoring applications that set rents below federally-determined levels. Typically, the scoring rules are additively separable in characteristics \( x_i \). This feature makes it possible to compute counterfactual scores \( q_i' = q(x_i') \) for counterfactual applications \( x_i' \) without observing all dimensions of \( x_i \).

Most states hold annual or semiannual application rounds. These are open to a variety of

\[ \text{The federal maximum on monthly rent works out to 1.25 percent of the local Area Median Income (AMI), although property-level “income averaging” introduced at the end of my sample allows some units to charge higher rents.} \]
entrants: for-profit developers, non-profit developers, and public housing agencies. Application appears costly. Developers must show their “readiness to proceed” with construction if funded. Satisfying this criterion usually entails “site control”—that is, land ownership, or more commonly, a purchase option contract contingent on winning. Applications also often include a construction plan, zoning approvals, pro-forma income statements, contracts for construction financing, and letters of support from local politicians.\(^{10}\)

**Awards.** Upon winning, developers transfer the credit to taxable equity limited-partner investors who finance construction (Desai et al., 2010). Units are income- and rent-restricted for at least 30 years. If an applicant loses, they may reapply in subsequent cycles. There is also a landscape of other subsidies they may pursue, including a less-generous credit that is not awarded competitively in most states. Losing applicants are also free to not develop or to develop for alternative uses.

Winning applicants receive nonrefundable tax credits with a face value of 70 percent of their proposal’s basis in present value, applied to tax liabilities over ten years. In some areas, winners qualify for a 30-percent “boost” to their basis. Thus, when boosted, each basis dollar translates to $0.91 in credits. Areas can qualify for the boost in two ways: as a Qualified Census Tract (QCT) or as a Difficult Development Area (DDA). Federal regulations assign the boost based on Census data, and the regulations are complicated. A tract can be designated a QCT according to its poverty rate or its income distribution relative to the median income of the metropolitan area. Areas are designated DDAs based on a ratio of rents to median incomes. Since 2016, this determination occurs at a combination of zipcode, county, and metropolitan-area levels.

### 2.2 Data

The next three subsections introduce the data. For further information, see Appendix B.

**Applications.** I built a new database of LIHTC applications covering 40 U.S. states from 2005 to 2019. The data are from the administrative records of state housing finance agencies, which I compiled, digitized, and standardized.\(^{11}\) The database contains records of 22,241 LIHTC applications, both winning and losing, with a total funding request of approximately $200 billion in 2022 constant dollars, adjusted for inflation using the Consumer Price Index.

For each of the 453 application rounds in the data (see Appendix Figure A1), I have every application considered. Due to agencies’ varying record-keeping practices, other rounds cannot be fully reconstructed and are thus excluded from the data. For each application, I have the proposed name and address of the development, the primary demographic group and income levels to be

\(^{10}\)The detail in applications fundamentally shapes this paper’s approach. On the one hand, enough is known about losers to provide a counterfactual to winners. On the other hand, self-selection into application is a major concern.

\(^{11}\)To minimize the burden of my requests on agencies, I requested records produced in their review processes. Most agencies were highly cooperative. I won or settled appeals against several agencies after they denied my initial requests.
served, the unit count (below-market and market-rate), the value of tax credits requested, the identity of the applicant (their name, contact details, and nonprofit status), and whether the funds are for new construction or rehabilitation. The data further include scores and all set-asides, enabling me to simulate tax-credit assignments. I identify reapplications by linking across rounds.

**Parcel Outcomes.** By combining several data sources, I observe whether and what, if anything, was developed at the level of the application parcel. The sources are CoreLogic and the National Housing Preservation Database (NHPD). I manually link the applications to both datasets.

The CoreLogic data collate tax assessments from localities and therefore include variables that are consistently available in such records. In particular, the data provide the assessed values of land and improvements, year built (the end year of construction), land use (e.g., multifamily residential), floor space (in square feet), and ownership information.

The NHPD is the most extensive database of U.S. subsidized housing, covering nearly all federal programs. I use the NHPD to measure if a property is subsidized, and by which program if not the LIHTC. If an application is never funded in my data (including reapplications), is not matched to an NHPD record, but appears in CoreLogic, I assume it is not subsidized.

**Neighborhood Data.** I use data by Census tract and block group as outcomes and as covariates. The outcome data come from the U.S. Postal Service (USPS) and a mail-marketing firm (Data Axle, formerly InfoGroup). I align both to consistent Census 2000 geographies. I use covariates from the 2000 Census, not from later years, to avoid confounding from the developments themselves.

The USPS data are counts of deliverable residential addresses by tract and quarter since 2006. The Data Axle files contain annual address-level microdata, also since 2006. I collapse the files to the block group and tract, so as to complement the USPS counts with details on demography and land use (age, estimated income decile, rental or owner-occupant, single- or multi-family).

### 2.3 Win Probabilities

A key input in my analysis is an application’s ex-ante win probability, taking rival applications as a random variable whose distribution but not realization is known when a developer applies. The win probabilities have two main purposes. First, I use the the probabilities to balance winners and losers on unobservables in Section 4. Second, I use them as developers’ beliefs in Section 8.

I estimate the win probabilities of applicants through a simulation procedure that I take from research on auctions (Hortaçsu and McAdams, 2010) and market design (Abdulkadiroğlu et al., 2017). The market-design literature uses explicit randomization embedded in mechanisms, such as tie-breaking by lottery. The auctions literature obtains win probabilities under informational assumptions (e.g., independent private values). LIHTC competitions only very rarely involve random assignment, so I assume application choices are made independently within round.
The estimated win probability is a valid balancing score in the sense of Rosenbaum and Rubin (1983) under two assumptions. First, the grant rule $W(\cdot)$ must be correctly specified. Second, developers must know no more than their own application and the distribution of potential rivals. Under these assumptions, the propensity-score theorem then gives that any residual variation in tax-credit assignment is independent of potential outcomes.

Variation in assignment conditional on win probability, however, is not guaranteed in non-stochastic mechanisms. A necessary condition for the existence of such variation is that the grant or scoring rule is multidimensional, so that applicants with the same propensity score can differ in their characteristics and thus assignments. Set-asides will create such variation in my context.

To obtain win probabilities, I first estimate the distribution $\hat{\Psi}_{it} = \hat{\Psi}(Q_{-it}, Z_{-it})$ of potential rivals by resampling from applications within the same round. That is, for a bootstrap replication $b = 1, \ldots, B$, I draw $N - 1$ applications uniformly with replacement from each applicant’s distribution of actual rivals. This yields simulated rivals $(Q_{b_{it}}, Z_{b_{it}})$ for an application $i$ in replication $b$.\footnote{I implement this procedure in a way that is more computationally efficient, so as to avoid $BN$ draws of the sample. Before each bootstrap replication, I split realized applications into $K$ “folds.” I retain the realized applications in one fold and resample among the $K - 1$ others. In calculating $\hat{\Psi}_{it}$, I only those replications in which application $i$ is in the retained fold. This procedure only requires $BK$ draws of the sample.}

Next, I run the mechanism on the applicant and each simulated draw of rivals, which assumes developers know their own score and set-asides.\footnote{An interesting institutional feature of the LIHTC is that applications often contain “self-scores” that agencies review, so as to limit the need for ex-post appeals. In Appendix B, I use this data to validate my assumption that developers are highly informed about their own scores.} This yields $B$ simulated assignments $\hat{W}_{it} = W(q_{it}, z_{it}; Q_{b_{it}}, Z_{b_{it}})$, and I calculate the estimated win probabilities by $\hat{p}_{it} = \frac{1}{B} \sum_{b=1}^{B} \hat{W}_{it}$.

### 2.4 Summary Statistics

**Application Characteristics.** Table 1 summarizes the data. Columns 1 and 2 show means of characteristics for winners and losers respectively. A typical application is for a development of roughly 60 units, is entirely income-restricted, and charged a monthly rent of approximately $900 in 2019. About 45 percent of developments are intended for elderly or other non-family groups (typically: disabled, homeless, or supportive housing). The competitions are dominated by “repeat players” (i.e., developers specialized in low-income housing), and 28 percent of applications involve nonprofit organizations. Relative to the national average, neighborhoods with applications are poor and densely-populated but do not have elevated rental vacancy rates.

Differences between winners and losers on observed characteristics are statistically significant but of small magnitude. Without mechanism-based win probabilities, a researcher might rely on simple comparisons of outcomes between winners and losers. However, Table 1 also suggests that such a design would be dangerous. Despite observable similarity, winners and losers differ sharply
Table 1: Covariate Balance Between Winning and Losing Applications

<table>
<thead>
<tr>
<th></th>
<th>Means</th>
<th>Differences (SE)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Winners (1)</td>
<td>Losers (2)</td>
<td>Uncontrolled (3)</td>
<td>Controlled (4)</td>
</tr>
<tr>
<td>Unit Count</td>
<td>61.62</td>
<td>64.94</td>
<td>-3.32*** (0.55)</td>
<td>0.20 (0.66)</td>
</tr>
<tr>
<td>% Low-Income Units</td>
<td>97.5</td>
<td>98.1</td>
<td>-0.6*** (0.1)</td>
<td>-0.0 (0.1)</td>
</tr>
<tr>
<td>Monthly Rent</td>
<td>$885.77</td>
<td>$894.37</td>
<td>-8.60** (4.16)</td>
<td>4.44 (5.40)</td>
</tr>
<tr>
<td>% New Construction</td>
<td>61.9</td>
<td>64.6</td>
<td>-2.8*** (0.8)</td>
<td>-2.6*** (0.9)</td>
</tr>
<tr>
<td>Family</td>
<td>54.1</td>
<td>56.7</td>
<td>-2.6*** (0.9)</td>
<td>1.1 (1.0)</td>
</tr>
<tr>
<td>Elderly</td>
<td>31.4</td>
<td>33.5</td>
<td>-2.1*** (0.8)</td>
<td>-1.3 (1.0)</td>
</tr>
<tr>
<td>Other</td>
<td>14.5</td>
<td>9.8</td>
<td>4.7*** (0.6)</td>
<td>0.2 (0.7)</td>
</tr>
<tr>
<td>Win Probability</td>
<td>69.9</td>
<td>23.6</td>
<td>46.3*** (0.5)</td>
<td>-</td>
</tr>
<tr>
<td>PDV Tax Credits Per Unit</td>
<td>$144,900</td>
<td>$155,182</td>
<td>-10,282*** (1,056)</td>
<td>-2,093 (1,314)</td>
</tr>
</tbody>
</table>

Panel A: Project Characteristics

Panel B: Applicant Characteristics

Panel C: Location Characteristics

Notes: This table reports means and differences in means between winning and losing applications. Columns 1 and 2 respectively report means of characteristics for winning and losing applications. Column 3 reports the unconditional win-minus-lose difference in means, and Column 4 reports the difference controlling nonparametrically for the win propensity score. Standard errors are clustered by tract. Observations indicate counts of winning and losing applications, but some covariates are unavailable for some applications. * = p < 0.10, ** = p < 0.05, *** = p < 0.01.
in their ex-ante win probabilities: on average, by 46.3 percentage points. By contrast, if one naively predicted win probabilities from a regression on all other application characteristics in this table, the analogous difference would be 7.0 percentage points. The gap reflects that the information contained in scores is not closely related to other observed characteristics.

Table 1 also shows we can easily reject balance on observables between winners and losers, even conditional on the win probabilities. This likely reflects minor misspecifications of QAP rules, not agencies’ secret discretion.\textsuperscript{14} I provide two checks regarding these imbalances, since I use winners-versus-losers comparisons in Section 4. First, in the tract event study, I control for variables where there is clear imbalance (Appendix Figure A2). This approach, however, does not address residual imbalance on unobservables. I therefore also instrument for the actual assignment using my version of the grant rule to simulate the assignment, while controlling for the win probability (Appendix Figure A3). Both tests suggest the remaining imbalance is innocuous.

Figure 2 shows the calibration and distribution of the simulated win probabilities. In the left panel, I present a binned scatterplot of the empirical win probability against the simulated values, along with a 45-degree line for reference. The simulation would be perfectly calibrated if the scatter points fell exactly along the reference line. The win probabilities are well calibrated but not perfectly so. As above, the issue appears to be difficulties in coding the QAP rules.

\textsuperscript{14}In a few cases, a QAP provides an opaque description of a provision in the grant rule $W(\cdot)$, and I have simplified a few very complex provisions.
**Rent Discounts.** Figure 3 describes the LIHTC in terms of the rent discounts it provides to tenants. Panel A show that on average, the monthly rents of subsidized units are close to the rent that I estimate could be charged by the median new unit in the same Census block group and with the same number of bedrooms. All values are as of the year 2019.

Panel B shows a binned scatterplot of regulated LIHTC rents versus new-unit market rents. Rent regulations bind strongly in high-rent neighborhoods, generating large discounts. However, few LIHTC developments are built in such areas. I estimate that, on average, a new market-rate unit in a neighborhood with a LIHTC application would rent for $990 per month, which is about 40 percent below the national-average rent for new units. Accounting for place—and, in particular, that low-income housing is built mostly in places with low willingness to pay for new housing—is crucial to estimate rent savings. Indeed, I find regulated rents do not bind for about half of applications, so considerable subsidy goes to developments charging local market rents.

Assuming that subsidized units charge the minimum of the market rent and their regulated rent, the average monthly rent discount is 12 percent. In present-value terms, the LIHTC thus buys little in rent savings. At a 5-percent annual discount rate, $1 in the present value of tax credits reduces the present value of rents by 27 cents.

**Notes:** This figure compares rents in LIHTC units to the median rents of new units in the same Census block group. See the text for explanations of the panels.

---

15“New” means built between 2010 and 2019. New-unit rents are observed in only 17 percent of Census tracts. I therefore impute new-unit rents using a hedonic regression that incorporates local rents on older units and a sample-selection correction that accounts for selective development with respect to potential rents. Other processing steps project these rents to the block-group level and align regulated and market rents on bedrooms. See Appendix B.
Figure 4: Housing for Whom? LIHTC Tenants and Counterfactual Tenants

<table>
<thead>
<tr>
<th>Household Income (Thousands of Dollars)</th>
<th>Percentage of Households</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0 to $10</td>
<td>LIHTC</td>
</tr>
<tr>
<td>$10 to $20</td>
<td></td>
</tr>
<tr>
<td>$20 to $35</td>
<td></td>
</tr>
<tr>
<td>$35 to $50</td>
<td></td>
</tr>
<tr>
<td>$50 to $75</td>
<td></td>
</tr>
<tr>
<td>$75 to $100</td>
<td></td>
</tr>
<tr>
<td>$100+</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This figure contrasts the household-income distribution in LIHTC units with the predicted distribution in the counterfactual market-rate development. See the text and Appendix B for details on this comparison.

Several caveats are useful in interpreting these estimates. First, benchmarking low-income housing to the median new rental unit nearby assumes these units are comparable on other dimensions, which ignores potential differences in quality as well as any non-rent services provided by low-income units. Second, interpreting rent discounts as a measure of tenant welfare requires that tenants view the benefit as equivalent to cash. Third, my estimates are of the rent discounts when the building is new. As the building depreciates, its counterfactual market rents fall, but regulated rents do not, reducing the rent savings. Finally, some subsidized units surely appear the new-unit comparison group, attenuating the estimated rent savings.

Tenant Composition. I also examine whether LIHTC tenants differ in income from the likely tenants of counterfactual market-rate developments. I do so using property-level HUD data on the household incomes of LIHTC tenants. To overcome privacy-related censoring and limited distributional information in the released data, I fit property-level normal distributions by method of moments (see Appendix B). For each property, I impute the non-LIHTC counterfactual income distribution using the joint distribution of rents and household incomes in the surrounding tract. In particular, this approach assumes the counterfactual development rents at what I estimate above.

Figure 4 shows that LIHTC tenants have lower incomes than counterfactual non-LIHTC tenants.
The fact in the raw data that drives this result is that the median LIHTC household earns much less than the median household in the same tract (see Appendix Figure A23). In Section 4, I present some direct evidence consistent with reallocation of new housing: When a tract wins the LIHTC, it sees increases in the share of households with Section 8 vouchers and the share in the bottom decile of the national income distribution.

These results raise the question of whether one should, as I do, measure the LIHTC’s household welfare effect by its impact on housing costs. If housing markets facing low-income households are competitive, this approach is indeed sensible: Households then face no barriers other than price, so the reallocation of new housing across households is not a source of welfare gains. However, the assumptions underpinning this view are hardly beyond question. For instance, Bergman et al. (forthcoming) and Christensen and Timmins (2023) find barriers for voucher recipients and minorities in accessing high-quality neighborhoods.

3 A Dynamic Model of Housing Markets

This section introduces a dynamic equilibrium model of the markets for subsidized and unsubsidized housing. Its primary aim is to jointly explain the application and building behavior of developers. The model’s household side is kept simple, with the minimum needed for equilibrium effects of subsidies in the unsubsidized market. I then provide welfare and incidence measures. Appendix C discusses the model’s nonparametric identification.

3.1 Setup

**Choices.** A developer \( i \in I \) has the exclusive right to develop a land parcel. They make two profit-maximizing choices in each time period \( t \). The first, \( A_{it} \in \{0, 1\} \), is whether to submit their parcel to a grant competition and, if they win, must build low-income housing. If they lose the competition, the developer may reapply in the future. The second choice, made subsequently, is whether to build if they do not win or if they do not apply for the grant: \( B_{it} \in \{0, 1\} \). Developers can defer the building decision indefinitely. Whether subsidized or not, development is an absorbing state. Figure 5 depicts the structure of choices in the model.

**State Variables.** The developer’s vector of state variables is denoted by \( s_{it} = \{d_{it}, h_{it}, x_{it}, \eta_{i}, r_{it}^{m}\} \). The variable \( d_{it} \) indicates whether the parcel is developed by \( t \). The “history” variable \( h_{it} \) indicates whether the developer has previously applied for the grant by \( t \). The developer also has characteristics \( x_{it} \) and scalar unobservable characteristic \( \eta_{i} \), which has a distribution \( G \) and is fully persistent

---

16Indeed, the literature on household valuations of in-kind transfers usually interprets this distortion in housing consumption as a source of welfare loss. Estimates reviewed in Olsen (2003), for instance, suggest tenants value housing benefits at about 80 cents on the dollar.
through time. The market rent for housing is \( r^m_t \).

Each period, the developer also draws temporary unobservable shocks \( \varepsilon_{it} = (\varepsilon^A_{it}, \varepsilon^B_{it}) \) from a distribution \( F \). The shocks are assumed to be independently and identically distributed (i.i.d.) and additively separable from payoffs. Developers’ draws are private information, so that each faces uncertainty as to which rivals will apply. They therefore perceive a probability \( p_{it} \) of winning the grant if they apply. Developers are atomistic, taking their individual win probabilities and rents as given but determining both collectively, following Hopenhayn (1992).

3.2 Application and Building Decisions

I express the developer’s problem through two Bellman equations. The first, which I call the application value function, concerns the developer’s application choice. The second, which I call the building value function, concerns the developer’s building choice. Developers move between these value functions by deciding not to apply and by deciding not to build.

**Application.** The developer’s application value function in the state \( s_{it} \) is given by:

\[
V^A(s_{it}, \varepsilon_{it}) = \max_a \left\{ \Pi^A(a, s_{it}) + \varepsilon^A_{it}(a) + \beta \mathbb{E}[V^A(s_{it+1}, \varepsilon_{it+1}) \mid a, s_{it}] \right\},
\]

where their application choice is \( a \) and \( \Pi^A(a, s_{it}) \) denotes their expected flow payoff upon taking the action \( a \) from their state. The expected continuation value \( \mathbb{E}[V^A(s_{it+1}, \varepsilon_{it+1}) \mid a, s_{it}] \) is conditional on their choice and is discounted by a time-preference parameter \( \beta \). Application choices are therefore \( A(s_{it}, \varepsilon_{it}) = \arg \max_a \left\{ \Pi^A(a, s_{it}) + \varepsilon^A_{it}(a) + \beta \mathbb{E}[V^A(s_{it+1}, \varepsilon_{it+1}) \mid a, s_{it}] \right\} \).
The developer’s expected payoffs are

\[ \Pi^A(a, s_{it}) = \begin{cases} 
  p_{it} \pi_1(s_{it}) - \kappa(s_{it}) & \text{if } a = 1 \\
  \mathbb{E}[V^B(s_{it}, \epsilon_{it}) | s_{it}] & \text{if } a = 0.
\end{cases} \quad (3) \]

In choosing to apply, the developer compares the expected value of applying to a reservation value, their building value function. If they apply, they receive the win payoff \( \pi_1(s_{it}) \), scaled by the win probability, less their entry cost \( \kappa(s_{it}) \). They pay entry costs whether they win or lose. Upon winning, the developer receives no further payoffs. If they apply, the state is updated such that \( h_{it} = 1 \); the update is to \( h_{it+1} = d_{it+1} = 1 \) if they win.

If the developer does not apply, their payoff is the building value function \( \mathbb{E}[V^B(s_{it}, \epsilon_{it}) | s_{it}] \), taking the ex-ante expectation over \( \epsilon_{it}^B \). By not applying, the continuation value \( \mathbb{E}[V^A(s_{it+1}, \epsilon_{it+1}) | a = 0, s_{it}] \) is also set to zero for all states \( s_{it} \). A decision not to apply thus moves the developer over to the building value function, but it does not rule out applications in future periods. Not applying preserves the developer’s history and development state variables: \( h_{it+1} = h_{it} \) and \( d_{it+1} = d_{it} \).

**Building.** The building value function is

\[ V^B(s_{it}, \epsilon_{it}) = \max_b \left\{ \Pi^B(b, s_{it}) + \epsilon_{it}^B(b) \right\}, \quad (4) \]

where \( b \) is the building choice. These are given by \( B(s_{it}, \epsilon_{it}) = \arg \max_b \left\{ \Pi^B(b, s_{it}) + \epsilon_{it}^B(b) \right\} \). The expected payoffs are

\[ \Pi^B(b, s_{it}) = \begin{cases} 
  \pi_0(s_{it}) & \text{if } b = 1 \\
  \beta \mathbb{E}[V^A(s_{it+1}, \epsilon_{it+1}) | s_{it}] & \text{if } b = 0.
\end{cases} \quad (5) \]

The developer’s building choice thus involves a comparison of their outside option and their application value function. In particular, if the developer chooses to build, their payoff is \( \pi_0(s_{it}) \). They receive no further payoffs, and their development state variable is updated to \( d_{it+1} = 1 \). If they do not build, their payoff is the (ex-ante) application value function \( \beta \mathbb{E}[V^A(s_{it+1}, \epsilon_{it+1}) | s_{it}] \). They also move back to the application value function and preserve history and development states.

### 3.3 Grant Mechanism

Applications are scored according to \( q_{it} = q(x_{it}, \eta_i) \). The dependence on the unobservable induces both selection in awards and self-selection at application, as potential applicants anticipate their win probabilities. If a developer applies, their grant assignment is given by Equation 1. To compute their win probability, a developer evaluates the conditional expectation of their assignment over the
distribution of potential rival applicants $\Psi_{it}$:

$$p_{it} = \int W(q_{it}, x_{it}; Q_{-it}, X_{-it}) \, d\Psi_{it}(Q_{-it}, X_{-it}).$$ (6)

The potential-rival distribution is given by the product of the distribution $\varphi$ of potential-applicant characteristics and the conditional probability of application for such a potential applicant: $\Psi_{it} = \prod_{j \in I^i} \Pr(A_{j-it} \mid s_{j-it}) \varphi(s_{j-it})$, where $j$ indexes potential applicants.

That the shocks $\varepsilon_{it}$ are private information has two implications in the model. First, developers cannot determine their rivals’ application decisions before making their own, so $\Pr(A_{j-it} \mid s_{j-it})$ takes intermediate values in $\Psi_{it}$. Second, potential outcomes are independent of the grant assignment, conditional on the win probability and on deciding to apply: $B_{it} \perp W_{it} \mid A_{it} = 1, p_{it}$.

### 3.4 Housing Demand

A representative household rents the entire housing stock each period. They view subsidized and unsubsidized housing as perfect substitutes, and they have constant elasticity of substitution (CES) preferences over housing $H_t$ and a non-housing numeraire good $C_t$:

$$\max_{H_t, C_t} u(H_t, C_t) = \left[ H_t^\rho + (\phi C_t)^{\rho-1} \right]^{\frac{1}{\rho}} \quad \text{s.t.} \quad r_t H_t + C_t = Y_t, \tag{7}$$

where $\rho$ is the elasticity of substitution and $\phi$ is a preference weight on the numeraire. The household has an exogenous nominal income of $Y_t$. These preferences give rise to a price index

$$P(r_t) = \left[ r_t^{1-\rho} + \phi^{\rho-1} \right]^{1/(1-\rho)},$$

where $r_t = [\lambda_t (1-\delta) + (1-\lambda)] r^m_t$ is the average rent. There is an equilibrium share $\lambda_t$ of subsidized units, renting at a discount of $\delta$, set by the government, relative to the market rent.

I measure household welfare by the present value of its indirect utility, $V^H = \sum_t \beta^t (Y_t/P(r_t))$. For the counterfactuals, I assume that changes in rents enter developer primitives as present values: $\Delta \pi_0(s_{it}) = \sum_t \beta^t \Delta r^m_t H(s_{it})$ and $\Delta \pi_1(s_{it}) = (1-\delta) \sum_t \beta^t \Delta r^m_t H(s_{it})$, where $H(s_{it})$ is the developer’s potential number of units. This adjustment would result from treating the outside option as a difference of construction cost and the present value of rent: $\pi_0(s_{it}) = -C(s_{it}) + \sum_t \beta^t r^m_t H(s_{it})$.

### 3.5 Equilibrium and Welfare

**Equilibrium.** Given primitives $\{\pi_0, \pi_1, \kappa, F, G, \varphi\}$, parameters $\{\beta, \phi, \rho, Y_t\}$, and an initial state, an equilibrium is defined by a sequence of endogenous developer quantities $\{A_{it}, B_{it}\}$, household quantities $\{H_t, C_t\}$, and probabilities and prices $\{p_{it}, r^m_t\}$ such that
1. **Developers and households optimize**: For all \(i, t\), application and development choices \(A_{it}\) and \(B_{it}\) maximize Equations 2 and 4, and housing consumption \(H_t\) maximizes Equation 7.

2. **Housing market clears**: Market rents \(r_i^m = r^m\) balance supply and demand in expectation.

3. **Rational expectations**: Win probabilities \(p_{it}\) satisfy Equation 6.

**Welfare.** Social welfare is the sum of developer and household welfare:

\[
W = E[V^A(s_{it}, \varepsilon_{it})] + V^H.
\]

Let \(\Delta\)’s indicate differences with respect to some policy change, and let the present value of entry costs be \(K = E[\sum_{i,t} \beta^{-1} A_{it} \kappa(s_{it})]\). In the incidence analysis, I will refer to developer and household shares, defined respectively as \(\Delta E[V^A(s_{it}, \varepsilon_{it})]/(\Delta W + \Delta K)\) and \(\Delta V^H/(\Delta W + \Delta K)\). I also refer to the share lost to entry costs, \(\Delta K/(\Delta W + \Delta K)\). I include entry costs in the incidence-share denominator to account for the use of real resources.

## 4 Causal Effects of Tax Credit Awards

This section estimates the causal effects of winning the LIHTC on development outcomes. It is the first of three empirical analyses that I aim to match in estimating the model in Section 3.

### 4.1 Setup

The causal effects of awards on development are informative about outside options. If winning greatly raises the probability of development, then the counterfactual land use of applicants is usually to leave it undeveloped. In the model, the probability of building without the LIHTC is

\[
\log \frac{\Pr(B_{it} = 1 \mid s_{it})}{1 - \Pr(B_{it} = 1 \mid s_{it})} = \frac{1}{\sigma_b} \left[ \pi_0(s_{it}) - \beta E[V^A(s_{it+1}, \varepsilon_{it+1}) \mid s_{it}] \right],
\]

where I suppress the probability’s condition of having applied and lost, \(A_{it} = 1\) and \(W_{it} = 0.17\) All else equal, a developer with a strong outside option (large \(\pi_0(s_{it})\)) is likelier to develop if they lose. However, the build probability does not directly reveal outside options, in that it also depends on the value of waiting to apply or build later (\(\beta E[V^A(s_{it+1}, \varepsilon_{it+1}) \mid s_{it}]\)). The identification would be one-to-one in a static model, where there is no value of waiting.

By conditioning on the unobservable part \((\eta_t)\) of the state \(s_{it}\), Equation 8 also warns about naive “winners-versus-losers” estimators of grant impacts. For instance, if applications viewed by the

---

17This equation also assumes that \(\varepsilon_{it}^R\) is distributed Type I extreme value with dispersion \(\sigma_b\), as in Section 8.
government as strong also tend to have unobservably strong outside options, then naive estimates will overstate true causal effects and understate winners’ outside options.

Motivated by this concern, I estimate winners-versus-losers comparisons controlling for the win probability. Depending on context explained below, I estimate various event-study specifications, but all comparisons fundamentally take form

\[
Y_i = \beta(\hat{p}_i)W_i + f(\hat{p}_i) + u_i, \tag{9}
\]

where \(Y_i\) is a development outcome for application \(i\), \(W_i\) is an indicator for \(i\)’s tax-credit assignment and \(\hat{p}_i\) is its win probability. To keep \(f(\cdot)\) flexible, I use a cubic basis spline with four knots spaced evenly through the distribution of win probabilities. The \(\beta(\hat{p}_i)\) notation allows for heterogeneous effects by win probability.

Let \(\hat{\beta}\) denote the estimator in Equation 9 imposing constant win effects. Under conditions discussed in Section 2.3, the win probability is a sufficient statistic for the government’s information set with respect to application \(i\). It thus balances winners and losers on \(\eta_i\), with \(E[u_iW_i] = 0\). By consequence, \(\hat{\beta}\) is a valid estimator of a weighted-average effect of tax-credit awards on winners, \(E[y_i(W_i = 1) | W_i = 1] - E[y_i(W_i = 0) | W_i = 1]\), where \(y_i(W_i)\) is the potential development outcome of \(i\) under a tax-credit assignment \(W_i\).

### 4.2 Parcel Effects

Figure 6 shows the results of a parcel-level event study of development outcomes.\(^{18}\) The left half of Panel A shows effects of winning and losing on the annual probability of development, relative to earlier and later applicants. The right half of Panel A shows win-versus-lose estimates consistent with Equation 9. In both, the outcome is an indicator for whether the parcel is recorded as having construction or rehabilitation in that year.

Taken together, the two halves of Panel A tell a clear story. When developers win, construction follows almost immediately. When developers lose, construction typically still happens, albeit at a lag of several years and not substituting one-for-one with the LIHTC. I estimate that, at a horizon of ten years after the LIHTC round, about 75 percent of winning parcels would have been counterfactually developed had they lost. I compute this “displacement rate” by summing the lose and win coefficients from Panel A and dividing the former by the latter.\(^{19}\)

---

\(^{18}\)Some applications concern multiple parcels. When construction years differ across parcels, I code the outcome for the largest parcel by floor space. Weighting results by floor space yields essentially identical results (see Appendix Figure A4). Appendix Figure A5 shows some limited evidence for heterogeneous effects by win probability, with greater displacement among applications that are likelier to win.

\(^{19}\)Two calculations illuminate the extent to which development is pulled forward in time. First, I adjust the displacement rate for discounting at five percent per year. I obtain a rate of 71 percent, as compared to 75 percent without discounting. Second, I calculate the number of additional “building–years” within the 10-year horizon (linear
Figure 6: Parcel Effects of Tax Credit Awards

Panel A: Total Development

Panel B: Development by Type

Notes: Panel A plots win and lose effects (on left) and the win–lose difference (on right), of tax credit awards on the property construction is recorded as completed in a given year. Panel B decomposes these development impacts into LIHTC, other-subsidy, and unsubsidized development. In both, the specification is an event-study analog of Equation 9 on applicant parcels: \( B_{it} = \alpha_i + \alpha_t + \sum_k [\beta_k \text{Win}_{it} + \gamma_k \text{Lose}_{it} + f_k (\hat{\beta}_{it})] + \epsilon_{it} \), for \( k \) years from the competition. Standard errors in Panel A are clustered by the consolidated parcel.

discounting). The result is 2.1 building–years, implying a displacement rate in these terms of 67 percent. Incorporating time preference thus modestly reduces estimates of displacement.
What gets built? I answer this in Panel B. Not surprisingly, winners build LIHTC units. There is also a small effect on unsubsidized development on parcels I associate with winning applications. When developers lose, some reapply, succeed in their reapplication, and therefore end up with LIHTC developments. Around half of development among losers is due to reapplication. However, many losers also find other ways to develop the property. About 21 percent of developments among losers secure a project-based subsidy other than the LIHTC, while the remaining 29 percent appears not to be subsidized. Overall, reapplication, substitution across subsidies, and substitution to unsubsidized development are all relevant contributors to the high displacement rate.

The rightmost figure of Panel B provides the win–lose difference by type of development. Winning the LIHTC pulls forward development in time and changes its nature from one subsidy to another or from unsubsidized to subsidized. It largely does not induce development where it otherwise would not happen within a few years.

4.3 Neighborhood Effects

Figure 7 shows the results of event-study comparisons of neighborhoods with winning and losing applications to similar non-applicant tracts. All specifications control for win probabilities in applicant tracts and for pre-award neighborhood observables.²⁰

Panel A estimates quarterly impacts on a tract’s total occupied housing stock. It confirms the parcel-level findings. Following a LIHTC win, there is a tract-level development “boom,” relative to ex-ante similar non-applicant tracts, which occurs sharply with the timing of the LIHTC round. The increase in its occupied housing stock corresponds to 150 additional households, which is about three times the average number of proposed units.²¹

On its own, the comparison of winners to ex-ante similar non-applicants would seem quite compelling as evidence that awards have large causal effects on development. Yet the comparison group of tracts with losing applications suggests it is not so: The losers also “boom.” Indeed, at standard levels of statistical significance, it is hard to reject zero net impact of winning the LIHTC on a tract’s occupied housing stock (see Appendix Figure A6). The null is sufficiently precise as to rule out effects larger than 30 households, or half of the average size of a single application.

Panel B estimates annual impacts on block groups’ household counts, splitting up each tract’s block groups into those with LIHTC applications and those without. Nearly all of the tract-level development “boom” occurs within the block group of the application, whereas little occurs in

²⁰The neighborhood controls are those listed in Panel C of Table 1. I take the logs of population density and household income, and I also include the tract’s cumulative population growth rate from 1990 through 2000. In a few tracts with multiple applications in the same year, I compute the probability any application wins in the tract, assuming independence across applications.

²¹Some of the excess increase in the housing stock appears to reflect follow-on applications. It may also be attributable in part to “crowded-in” developments that are unaffiliated with the applicant.
Figure 7: Neighborhood Effects of Tax Credit Awards

Panel A: Tract (USPS)

Change in Occupied Housing Stock (p.p.)

Panel B: Block Group (Data Axle)

Notes: Panel A plots win and lose effects on the occupied housing stock (change in percentage points). Panel B plots win and lose effects on the change (in percentage points) in the household count in the same Census block group as the LIHTC application, or in the same tract but other block groups. In both, the specification is an event-study analog of Equation 9 that includes non-applicant tracts: \( \Delta \log E[Y_{i,t+k}] = \alpha_k \text{Win}_{it} + \beta_k \text{Lose}_{it} + X_{it} \gamma_k \), for \( k \) quarters or years from the award. The specification is estimated via Poisson regression, and it includes controls \( X_{it} \) for the win probability, pre-award tract characteristics, and the baseline level of the outcome. Winning and losing are both defined as a neighborhood’s first event in sample. In both panels, the color bands depict pointwise 95-percent confidence intervals, and standard errors are clustered at the tract level.
adjacent block groups. These findings help to establish that we are observing the same economic phenomenon, of parcel-level displacement, in the parcel- and tract-level results.

The impacts of winning the LIHTC on local demographics and land use suggest that counterfactual developments are observably similar to subsidized development but that their residents are poorer. Winning and losing neighborhoods both see growth in the stocks of multi-family and rental housing (Appendix Figure A7). Appendix Figure A9 finds clear differences in voucher use between winners and losers, and Appendix Figure A8 shows suggestive evidence of changes in local income distributions. Appendix B also contains robustness checks.

5 Application Responses to Subsidy Generosity

This section estimates application responses to variation in the LIHTC’s value. I leverage the “basis boost” introduced in Section 2, which raises the rate at which construction expenses reduce tax liabilities. I take two empirical approaches to this variation: an event study around entry and exit from the boost, and an RDD at the boost threshold. In structural estimation (Section 8), I draw on results from both approaches as targeted moments.

From the model’s perspective, developer responses to subsidy generosity are informative about the entry cost $\kappa(s_{it})$. To see this, consider the developer’s application probability:

$$
\log \frac{\Pr(A_{it} \mid s_{it})}{1 - \Pr(A_{it} \mid s_{it})} = \frac{1}{\sigma_a} \left[ p_{it} \pi_1(s_{it}) - \kappa(s_{it}) + (1 - p_{it}) \beta E[V^A(s_{it+1}, \epsilon_{it+1}) \mid s_{it}] - E[V^B(s_{it}, \epsilon_{it}) \mid s_{it}] \right],
$$

where I impose the assumption (as I will in Section 7) that $\epsilon_{it}^A$ is distributed Type I extreme value with dispersion $\sigma_a$. This equation implies that, when $\pi_1$ rises, developers become more likely to apply. As before, however, one cannot directly “read off” $\pi_1$ or $\kappa$ from these behavioral responses. The policy variation also affects the value functions $V^A$ and $V^B$, motivating the structural approach, but would directly identify $\kappa(s_{it})$ in a static setting.

5.1 Event Study

**Approach.** Areas are assigned to the boost according to the tract-level poverty rate and median income (see Section 2). Changes in boost thus reflect, in principle, a combination of sampling variation and genuine changes in local economic conditions. As changes in conditions may independently affect the desirability of local LIHTC investment, the credibility of the event-study approach depends upon whether sampling variation or true economic variation predominate.

---

22It is helpful here to appreciate the fine geographic scale of a block group. In the 2000 U.S. Census, there were 211,267 block groups, or 3.2 block groups per tract, each containing about 500 households on average.

23It thus complements the win–lose analysis in Section 4, which I show is informative about outside options $\pi_0$. 

---
Simulations suggest the identifying variation in the boost is, fortunately, almost entirely due to sampling. To arrive at this conclusion, I reassign the boost using the actual assignment rule but drawing two sets of new values for its assignment variables. I draw the values using Census estimates of tract-level standard errors (see Appendix B for details). In 2019 data, for instance, around 3.9 percent of tracts are differently assigned between two simulation runs. By comparison, 8.0 percent of tracts changed their actual designation from 2018 to 2019, as I show in the left panel of Figure 8. Sampling variation is so large because tract-level estimates from national surveys are noisy. Furthermore, my event-study approach exploits variation in the timing of switches, compounding the importance of sampling variation.24

I estimate impacts of the boost by the following generalized event-study specification:

\[ y_{it} = \alpha_i + \alpha_t + \sum_s \beta_s \Delta \text{Boost}_{i,t-s} + X_{it} \gamma_t + \epsilon_{it}, \]  

where tracts are indexed by \( i \) and time \( t \) is in years. The terms \( \alpha_i \) and \( \alpha_t \) are thus tract and year fixed effects, and \( X_{it} \) is a vector of controls with potentially time-varying coefficients \( \gamma_t \). Standard errors are clustered by tract. The specification is estimated via a Poisson regression.

I note several implementation details. The event variable \( \Delta \text{Boost}_{i,t} \) can take three values: 0 (no year-to-year change in boost), 1 (gains boost), or −1 (loses boost). This specification thus assumes symmetric effects of entry and exit from boost, although I relax this assumption by allowing for differential effects of entry and exit. I include leads of the event of up to eight years ahead and for lags up to six years after. I bin any events occurring beyond these endpoints and thereby impose constant effects in these ranges (Schmidheiny and Siegloch, 2020). I exclude always-boosted tracts from the sample, so that never-treated tracts form the only pure control group in Equation 10. Due to a timing convention that LIHTC funds notionally for year \( t \) are sometimes committed in the fourth quarter of \( t - 1 \) (and are therefore recorded in my data as applications in \( t - 1 \)), the base event-time period in this event study is two years prior to the year a tract gains or loses its boost.

**Results.** The right panel of Figure 8 shows estimates of Equation 10, splitting effects into entry- and exit-specific estimates. The count of applications from a tract changes by approximately 30 percent in the years immediately after a change in boost designation, with symmetric effects of entry and exit. Pre-period coefficients appear steady before the change in boost. These results imply the application elasticity with respect to the net-of-tax price is roughly 0.25.

Appendix Figure A13 presents a key test of whether contemporaneous local shocks confound the event study. In this analysis, I predict application volume from the running variables among

---

24In related work, Freedman and Owens (2011) use panel variation in the boosted share of a county’s Census tracts as an instrument in a two-way fixed-effects model to study the effects of LIHTC on crime. In other policy contexts, Feiveson (2015), Suárez-Serrato and Wingender (2016), and Chodorow-Reich et al. (2019) all exploit measurement error and nonlinear transformations of data embedded in policy rules as sources of variation.
the non-boosted tracts and then use these predicted volumes as the event-study outcome. I find that year-to-year variation in the running variables is essentially unrelated to application volumes in this sub-sample, consistent with noise. The expected change in applications around boost entry and exit is thus essentially zero. Appendix B presents additional robustness checks.

5.2 Fuzzy Regression Discontinuity Design

**Approach.** The RDD directly exploits the cutoff rules in the boost designation. Some tracts are boosted despite being on the “unboosted” side of the cutoff and vice versa, due to other rules, making the discontinuity slightly fuzzy (see Appendix B). Relative to the event study, the main strengths of this strategy is that I can investigate changes in application characteristics, as well as its perhaps-finier control for local economic conditions.

I define the running variable as a tract’s distance from the cutoff with respect to its rank in the running-variable distribution for its metropolitan area. Thus, a running variable value of $-0.1$ implies a tract is 10 percentiles away in its metro-specific distribution from being designed for the boost. This approach consolidates cutoffs in the assignment rule. The median metro area contains 64 tracts, so comparisons within a bandwidth of 0.1 of the running variable represent comparisons of roughly the six tracts nearest to either side of the cutoff in a typical metro area.

Both Baum-Snow and Marion (2009) and Davis et al. (2019) use the same discontinuity to
identify effects of LIHTC projects on local areas. I make several adjustments to the definition of the running variable that strengthen the sharpness of the first stage (see Appendix B). These adjustments incorporate steps in the boost assignment rule omitted in prior work.

I model the effect of being boosted as

\[ Y_{it} = \beta_{\text{Boost}it} + f_2(c_{it}, \delta) + u_{it}, \]

where \( f_2(c_{it}, \delta) \) is a locally-linear specification in the running variable \( c_{it} \) estimated with a bandwidth \( \delta \). The coefficient \( \beta \) captures the effect on LIHTC outcomes of a tract just barely qualifying for the boost as a Qualified Census Tract (QCT). The “first stage” in the fuzzy RD is

\[ \text{Boost}_{it} = \gamma_{\text{QCT}it} + f_1(c_{it}, \delta) + v_{it}, \]

where \( f_1(c_{it}, \delta) \) is also locally linear in the running variable and \( \text{QCT}_{it} \) indicates QCT status.\(^{25}\)

**Results.** Figure 9 presents plots of the treatment and main outcome variable, the annual count of applications per 10,000 households in the tract, around the cutoff. Table 2 reports accompanying estimates for a broader array of application-supply outcomes. Overall, I find an increase in application volumes of 90 percent at the threshold, exceeding the event-study impacts.

The left panel of Figure 9 shows treatment assignment around the threshold. The probability a tract is a QCT jumps from approximately 20 percent to 80 percent at the threshold, as shown by the hollow blue dots. Some tracts are boosted for other reasons, and so the solid blue dots show that approximately 30 percent of tracts are boosted just below the QCT threshold. About 80 percent of tracts are boosted just above the QCT threshold.

The right panel of Figure 9 visualizes the application response to the boost. There are an additional 0.16 applications per 10,000 households in barely-boosted tracts, relative to a base level of 0.34 applications in barely-unboosted tracts, yielding an IV-estimated increase of 90 percent. The general upward slope in the outcome variable in Figure 9 reflects that tracts are poorer as the running variable rises. Table 2 shows similar increases in other measures of application supply.

Appendix A includes supporting analyses. Appendix Figure A18 plots the discontinuity for all outcomes in Table 2. Appendix Figure A19 shows the running variables in the assignment rule evolve smoothly through the cutoff, along with four other tract characteristics. Tracts just above and below the cutoff appear similar on observable characteristics.\(^{26}\) I also find no evidence of heterogeneity by six application characteristics (Appendix Figure A20) and modest, if imprecise,

\(^{25}\)Abdulkadiroğlu et al. (2017) provide conditions under which this design estimates a convex-weighted average treatment effect over the multiple cutoffs.

\(^{26}\)Due to transformations involved in the running variable, there is a discrete change in the mass of tracts at the cutoff, but institutional considerations rule out potential concerns of precise control.
Figure 9: Application Volume Around the Qualified Census Tract Threshold

Notes: In the left panel, this figure shows that the probabilities that a tract is designated a Qualified Census Tract (QCT) and is boosted both rise discontinuously in its distance to the QCT threshold. In the right panel, this figure shows that the count of LIHTC applications per 10,000 households living in the Census tract (in the year of application) also jumps at the QCT threshold. In both panels, the running variable is a distance defined with respect to tract rank within its metropolitan area. Both panels split the data into 15 equal-interval bins on either side of the QCT threshold.

Evidence of heterogeneity across tracts (Appendix Table A1). Taken together, these heterogeneity analyses suggest increases in the tax-credit rate would not attract more, but not “better,” applications.

Implications. Both sets of results in this section suggest entry costs are high on average. Many applicants apply only when the net-of-tax price of low-income housing is almost zero, as it is in boosted tracts. There are two possible explanations for that result. One is that low-income housing is not valuable to developers, but this is belied by the “bidding” analysis in Section 6. The remaining possibility is that, although winning is valuable, entry costs negate most of the win value.

6 Bidding for Subsidies

Developers face a trade-off between their win probability and the rent they may charge if they win. This section characterizes developer behavior in the face of this trade-off and estimates preferences over these objects. Using the preferences, I obtain a semi-structural estimate of developer incidence. This provides a useful check on the structural estimation in Section 8, where I will target this section’s reduced-form results directly, rather than its incidence estimate.
Table 2: Application Responses to Grant Value: QCT Threshold Estimates

<table>
<thead>
<tr>
<th></th>
<th>Is Boosted (1)</th>
<th>Applications (2)</th>
<th>Wins (3)</th>
<th>Proposed Units (4)</th>
<th>Funded Units (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: RD Estimates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QCT Threshold</td>
<td>0.541***</td>
<td>0.155***</td>
<td>0.069**</td>
<td>8.37**</td>
<td>3.46**</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.052)</td>
<td>(0.027)</td>
<td>(3.26)</td>
<td>(1.68)</td>
</tr>
<tr>
<td><strong>Panel B: Fuzzy RD Estimates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is Boosted</td>
<td>0.308***</td>
<td>0.123***</td>
<td>16.34***</td>
<td>5.50*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
<td>(0.046)</td>
<td>(6.06)</td>
<td>(2.82)</td>
<td></td>
</tr>
<tr>
<td>Bandwidth</td>
<td>0.030</td>
<td>0.044</td>
<td>0.056</td>
<td>0.044</td>
<td>0.054</td>
</tr>
<tr>
<td>Untreated Mean at Threshold</td>
<td>0.340</td>
<td>0.341</td>
<td>0.134</td>
<td>17.47</td>
<td>6.72</td>
</tr>
<tr>
<td>Estimate / Mean</td>
<td>0.904</td>
<td>0.914</td>
<td>0.935</td>
<td>0.819</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table reports estimates of LIHTC application responses to the Qualified Census Tract (QCT) designation, which raises grant value discontinuously at the QCT threshold. Outcomes in Columns 2–4 are in terms of levels per 10,000 Census tract households per year. I define the running variable as a tract’s distance in ranks from the QCT designation within the distribution of Census tracts in its metropolitan area. The untreated mean at the threshold is estimated by local linear regression to the left of the QCT threshold using the corresponding bandwidth in Panel B. The bottom row divides the fuzzy-RD estimate by the untreated mean at the threshold. * = p < 0.10, ** = p < 0.05, *** = p < 0.01.

### 6.1 Trade-Off Between Win Probability and Rental Income

Incentives in QAPs for lower rents, as noted in Section 2, allow developers to adjust their win probability by setting rents higher or lower. Three considerations make rent a useful dimension of applications to study. First, rent is particularly flexible as a choice variable, while other dimensions may be inflexible or require adjustments on other dimensions. Second, rent is a key component of application scores. In the median state in 2019, I calculate that rent preferences amounted to one quarter of the maximum point score. Third, the cost to developers of rent restrictions is measured in dollars, facilitating conclusions about their valuations of winning the LIHTC.\(^\text{27}\)

To analyze bidding decisions, I compute the consequences if each developer were to deviate unilaterally by setting rents modestly higher or lower than those they actually propose. I implement this approach by re-scoring 6,793 applications from 13 states (30 percent of my sample). For each, I searched for “up” and “down” deviations that were compliant with the QAP, not strictly dominated by another choice, and which solely involved a change in rents. I also determined the highest rent that would still receive full points in the rent category. With the new scores in hand, I simulate new

---

\(^{27}\)A key objection to rent as a trade-off variable is that, when rent regulations do not bind, rent concessions are free to developers on the margin. Appendix Table A2 uses market rents estimated in Section 2 to adjust for bindingness. With such adjustments, I estimate higher marginal rates of substitution and lower developer incidence.
Figure 10: Trade-Off Between Win Probability and Rental Income

Notes: The left panel is a binned scatterplot of two dimensions of alternative applications relative to those developers actually submitted. The horizontal axis is the present-value rent difference in thousands of dollars per unit. The vertical axis is the change in win probability in percentage points. The right panel is a histogram of the difference between developers’ proposed average unit rent and the score-maximizing rent level.

win probabilities for applications as in Section 2. For additional details, see Appendix B.

The left panel of Figure 10 shows the rent-versus-win-probability frontier facing developers, visualized as a binned scatterplot of both up and down deviations from actual applications. On the horizontal axis, I plot the change in applications’ present values of rental income per unit. On the vertical axis, I plot the change in applications’ win probabilities. The (negative of the) slope of the line through these points can be interpreted as an average marginal rate of transformation (MRT) between rents and probabilities. For a $1,000 reduction in the present value of rent per unit over a project’s first thirty years of occupancy, developers can raise their probability of winning the LIHTC by about 0.9 percentage points on average.28

The right panel of Figure 10 shows developers respond to incentives. It presents a histogram of applications’ proposed rents, computed as an application-level average, relative to the highest rent level that would qualify for full points. Providing rent reductions beyond this level has no direct QAP-score benefit, and thus QAPs generate strongly “kinked” incentives with respect to rent around this level. About 30 percent of applications bunch precisely at the score-maximizing kink point.29 Appendix Figure A21 uses variation in the location of the kink across states and over time

\footnote{Federal regulations do not require commitments to rent restrictions beyond 30 years, though some states encourage such extended commitments. I do not analyze these, as the present value of further years of rent regulation is highly sensitive to assumptions on developer discount rates.}

\footnote{Many developers restrict rent even beyond what the LIHTC directly encourages. My inspection of applications}
to nonparametrically estimate the bunching mass. Looking beyond the kink, Appendix Figure A22 shows that developers with stronger local incentives are more likely to set lower rents.

Figure 10 immediately suggests developer incidence of the LIHTC. Developers routinely bid rents down to the minimum so as to maximize their win probabilities. If equilibrium profits from the LIHTC were minimal, developers would be unwilling to trade lower rents for higher win probabilities at the margin. Intuitively, if the benefit of a higher win probability exceeds the cost of less rent if won, the prize must be worth winning.

6.2 Estimating Developer Preferences and Incidence

**Approach.** Consider a developer facing a menu of possible rents and win probabilities for their application \( \{r_{it}, p_{it}(r_{it})\} \), holding fixed all other dimensions. Denote the rent difference between two alternatives by \( \Delta r_{it} \) and the win-probability difference by \( \Delta p_{it} \). In general, rent levels could be related to other developer costs, which I represent as \( \Delta \xi_{it} = \Delta \pi_1(s_{it}) - \Delta r_{it} \).

From Equation 2, the difference between these alternatives in the application value function is

\[
\Delta V^A(s_{it}) \approx \Delta p_{it} \left[ \pi_1(s_{it}) - \beta E[V^A(s_{it+1}, e_{it+1}) | s_{it}] \right] + p_{it} \left[ \Delta r_{it} + \Delta \xi_{it} \right],
\]

where \( \pi_1(s_{it}) \) and \( p_{it} \) are respectively the win value and the win probability associated with an arbitrary base level for the rent.

This equation formalizes the developer’s trade-off between win probability and rental income. In particular, indifference between two alternatives such that \( \Delta p_{it} \neq 0 \) reveals the win value:

\[
\pi_1(s_{it}) = \beta E[V^A(s_{it+1}, e_{it+1}) | s_{it}] - \frac{p_{it}(\Delta r_{it} + \Delta \xi_{it})}{\Delta p_{it}}.
\]

The indifference condition in Equation 12 is a bid inversion, akin to those in research on auctions. Ignoring the value-function term, for example, a developer indifferent between a $1,000 per-unit rise to the present value of rent and a one-percentage-point increase in win probability, starting at a baseline probability of 50 percent, must value winning at $50,000 per unit in present value.\(^{30}\)

Here I consider a simplified version of Equation 11, and in particular, the following auxiliary model of developer choice:

\[
\Delta u_{ij} = \beta_1 \Delta p_i(r_{ij}) + \beta_2 \Delta \log r_{ij} + \Delta \xi_{ij} + \Delta e_{ij},
\]

suggests that developers may do so to qualify for other point categories in some QAPs (e.g., SRO housing), to comply with external rent restrictions (e.g., in public housing rehabilitation), or to pursue additional sources of funding.

\(^{30}\)It is innocuous to ignore the value-function term here. While the static value of winning is \( \pi_1(s_{it}) \), the dynamic equivalent is \( \pi_1(s_{it}) - \beta E[V^A(s_{it+1}, e_{it+1}) | s_{it}] \). Thus, the estimates in Table 3 apply to this dynamic object.
Table 3: Estimating Valuations from Bidding Behavior

<table>
<thead>
<tr>
<th></th>
<th>(1) OLS</th>
<th>(2) IV</th>
<th>(3) Cond. Logit.</th>
<th>(4) + Ctrl. Funct.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Win Probability</td>
<td>0.416***</td>
<td>1.708***</td>
<td>1.278***</td>
<td>6.423***</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.042)</td>
<td>(0.071)</td>
<td>(0.195)</td>
</tr>
<tr>
<td>Log Average Rent</td>
<td>0.620***</td>
<td>2.900***</td>
<td>1.826***</td>
<td>10.249***</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.098)</td>
<td>(0.147)</td>
<td>(0.395)</td>
</tr>
<tr>
<td>Applications</td>
<td>6,785</td>
<td>6,785</td>
<td>6,779</td>
<td>6,779</td>
</tr>
<tr>
<td>Marg. Rate of Substitution</td>
<td>0.923</td>
<td>1.051</td>
<td>0.884</td>
<td>0.987</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.021)</td>
<td>(0.039)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>Mean Win Value Per Unit</td>
<td>$54,957</td>
<td>$48,292</td>
<td>$57,385</td>
<td>$51,385</td>
</tr>
<tr>
<td></td>
<td>(2,441)</td>
<td>(1,022)</td>
<td>(2,608)</td>
<td>(1,040)</td>
</tr>
<tr>
<td>Developer Incidence Share</td>
<td>0.456</td>
<td>0.401</td>
<td>0.476</td>
<td>0.427</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.010)</td>
<td>(0.023)</td>
<td>(0.010)</td>
</tr>
</tbody>
</table>

Notes: This table reports estimates of coefficients in Equation 13. In Column 4, I take a control-function approach to instrument for the win probability in the conditional logit. Marginal rates of substitution are calculated as a ratio of coefficients; see the text for detail on the other calculations. Standard errors are clustered by application. * = p < 0.10, ** = p < 0.05, *** = p < 0.01.

where \(j\) denotes an alternative (that is, the actual choice, the “up” deviation, or the “down” deviation) and \(\Delta e_{ij}\) is an error term that is conditional mean zero. I use the latent utility index \(\Delta u\) to clearly distinguish from Equation 11. I estimate this equation as a conditional logit and as a fixed-effects linear probability model (LPM).

The parameters to be estimated are \(\beta_1\), the developer’s preference for a higher win probability, and \(\beta_2\), their preference for higher rents. With a structural error \(\Delta \xi_{ij}\), these parameters are not identified. That is, applications likely to win may be unattractive to developers on dimensions other than rents (e.g., high construction costs). To obtain identification, I therefore assume that within the comparisons I have constructed, differences in unobservables are uncorrelated with differences in rents and win probabilities: \(E[\Delta \xi_{ij} \Delta p_{ij}] = E[\Delta \xi_{ij} \Delta r_{ij}] = 0\). The justification for this assumption is that I control the comparison choices and consider a tightly restricted subset of alternatives.\(^31\)

**Results.** Table 3 presents estimates of Equation 13. Both the LPM (Column 1) and the conditional logit (Column 3) find an average marginal rate of substitution between win probability and rents

\(^31\)This identification assumption follows Pakes (2010), who uses a similar comparison-based sample to difference out a structural error term. It rules out the possibility of changes in other costs that flow from serving a different tenant population. For instance, extremely-low-income tenants can require supportive services in addition to housing; this would lead to some overstatement of developer incidence.
of approximately 0.9. On average, developers would be indifferent between a 0.9-p.p. rise in their win probability and a $1,000 rise in present-value rent. Encouragingly, this marginal rate of substitution matches the marginal rate of transformation in Figure 10. The final two rows of Table 3 use the approach in Equation 12 to estimate the mean per-unit value of winning the grant and for the incidence of the grant on developer profits. I find developers value winning at approximately $50,000 per unit in present value. By comparison, the average grant amount per unit is $120,000, implying that developers capture around 45 percent of the grant.

I explore a key concern with the use of simulated win probabilities to estimate preferences in Columns 2 and 4 of Table 3. If the simulated probabilities provide noisy estimates of developers’ beliefs, as seems likely, then choices will be more responsive to win probabilities than Columns 1 and 3 suggest. This bias will cause me to understate the marginal rate of substitution and, in turn, to overstate developer incidence. As a solution, I instrument for the win probability using the application’s rank in the round-specific distribution of QAP scores. Consistent with an attenuation bias, this approach yields slightly higher estimates of marginal rates of substitution and slightly lower developer valuations and incidence shares.

7 Structural Estimation

This section estimates the model introduced in Section 3. First, I introduce the parameterization of model primitives, along with other significant choices. I then explain the estimation procedure, which combines parametric policy iteration (PPI, see Rust, 2000; Sweeting, 2013) with simulated minimum distance (SMD, see Gourieroux et al., 1993).

7.1 Setup

Parametric Assumptions. I parameterize the win value $\pi_1(s_{it})$, the outside option $\pi_0(s_{it})$, the entry cost $\kappa(s_{it})$, the distribution $G$ of persistent unobservables $\eta_i$, and the distribution $F$ of temporary unobservables $\mathbf{e}_{it}$. I collect the parameters in the vector $\theta$. The matrix $\mathbf{x}_i$ contains observable application characteristics. The parameterizations are as follows:

- **Outside Option, Win Value, and Entry Cost:** The primitives $\pi_0(s_{it})$, $\pi_1(s_{it})$, and $\kappa(s_{it})$ are all assumed to be linear in the history $h_{it}$, the unobservable $\eta_i$, and the observable characteristics of applications and tracts. The characteristics are the number of units, the tax credits requested per unit, the tract poverty rate, and the tract (log) population density.

---

32 Across columns, the coefficients fluctuate in level due to differences in econometric approach. Incidence and the other objects of interest are determined by the ratio of coefficients, which is stable across approaches.
• **Application Characteristics:** I assume that potential applicants’ number of units and credits per unit are distributed independent log-normally. To avoid drawing from the distributions of scores and set-asides—an extremely high-dimensional vector—I directly simulate applications’ chains of win probabilities for their initial application and any reapplications. I use a bivariate beta distribution and assume the win probabilities follow a Markov process.

• **Unobservables:** I assume the persistent unobservable heterogeneity $\eta_i$ is distributed normally, setting the mean to zero and variance to unity, as it is rescaled in the primitive objects. For the temporary unobservables, developers draw i.i.d. Type I extreme value shocks $\varepsilon_{it} = (\varepsilon^A_{it}, \varepsilon^B_{it})$ each period. The shocks are additively separable from mean payoffs. The dispersion parameters are respectively $\sigma_a$ and $\sigma_b$.

The unit of simulation is a tract–year. Each tract draws a new potential applicant each year, which then makes a sequence of application and building decisions until reaching a terminal state. In estimating the model, I define parameters in units of dollars of the potential application’s qualified basis, normalizing for differences in scale. For instance, a value $\kappa(s_{it}) = 0.1$ would mean that the entry cost is a tenth of the basis, or about 14 percent of tax-credit value.

The model’s household side is calibrated. In particular, I take the elasticity of substitution $\rho = 0.691$ between housing and the non-housing good from Albouy et al. (2016), and I set $\phi$ to achieve a housing consumption share of 0.15.\(^{33}\) From Section 2, I set the rent discount $\delta = 0.12$. For further details regarding the structural setup and counterfactuals, see Appendix B.

### 7.2 Part 1: Parametric Policy Iteration

The first part of the estimation procedure (PPI) generates choice probabilities for simulated developers from the data and structural parameters. My exposition of PPI follows Sweeting (2013).

Let $\hat{P}^A(a, s_{it})$ be an initial estimate of the developer’s application probability from the state $s_{it}$, $\Pr(A_{it} = a \mid s_{it})$. Similarly, let $\hat{P}^B(b, s_{it})$ be an initial estimates of the building probability.\(^{34}\) Using Equation 2 and these initial choice probabilities, I compute the unconditional expectation of the developer’s application value function as

$$
\tilde{V}^A(s_{it}) = \sum_a \hat{P}^A(a, s_{it}) \left[ \hat{\Pi}^A(a, s_{it}) + \beta E[\tilde{V}^A(s_{it+1}) \mid s_{it}, A_{it} = a] \right],
$$

where $\hat{\Pi}^A(a, s_{it}) = \Pi^A(a, s_{it}) + E[\varepsilon^A_{it}(a) \mid s_{it}, A_{it} = a]$. Under the distributional assumptions on errors, $E[\varepsilon^A_{it}(a) \mid s_{it}, A_{it} = a] = \sigma_a(\gamma - \log \Pr(A_{it} = a \mid s_{it}))$, where $\gamma$ is Euler’s constant.

---

\(^{33}\)Estimated from U.S. Bureau of Economic Analysis data on the housing share of personal consumption expenditures.

\(^{34}\)I initialize the choice probabilities using the predicted values of a flexible logit model on tract-level observables.
The first step of PPI, termed “policy valuation” in Rust (2000), computes the expected value function \( \tilde{V}^A(s_{it}) \) under a given set of choice probabilities through a regression-based approximation. To set up the regression, let the vector \( \tilde{\Pi}_P \) collect the expected payoffs at each state: \( \tilde{\Pi}_P = E_P[\tilde{\Pi}^A(a, s_{it})] = \sum_a Pr(A_{it} = a | s_{it})\tilde{\Pi}^A(a, s_{it}) \), integrating over application choices. I then assume the application value function can be approximated by basis functions \( \phi_k(\cdot) \) of the state:

\[
\tilde{V}^A(s_{it}) \approx \sum_{k=1}^K \lambda_k \phi_k(s_{it}) = \Phi_{it} \lambda,
\]

where I let \( \Phi_{it} = \Phi(s_{it}) \), and where \( \lambda_k \) is a coefficient on the \( k \)th basis function. I use quadratic polynomials of the state variables as basis functions. Using this approximation and Equation 14, the application value function can be represented by

\[
\Phi_{it} \lambda = \tilde{\Pi}_P + \beta E_P[\Phi_{it+1}] \lambda,
\]

with \( E_P[\Phi_{it+1}] \) collecting the \( K \) values of \( E_P[\phi_k(s_{it+1}) | s_{it}] \).

Let the matrix \( \Phi \) stack the basis-function vector \( \Phi_{it} \) across individuals and periods. If there were exactly \( K \) unique states, equal to the number of basis functions, one could obtain coefficients \( \lambda \) on the basis functions by \( (\Phi - \beta E_P[\Phi])^{-1} \Pi_P \). In the typical (overidentified) case, the number of unique states exceeds the number of basis vectors, and estimates of \( \lambda \) can be obtained by

\[
\hat{\lambda} = ((\Phi - \beta E_P[\Phi])'(\Phi - \beta E_P[\Phi]))^{-1}(\Phi - \beta E_P[\Phi])' \Pi_P
\]

which are the coefficients from an OLS regression of \( \tilde{\Pi}_P \) on the matrix \( \Phi - \beta E_P[\Phi] \).

To complete the policy-valuation step, I use the estimates \( \hat{\lambda} \) to approximate the continuation value functions by \( \tilde{F}V(a, s_{it}) \approx E[V^A(s_{it+1}) | s_{it}, A_{it} = a] \). In particular,

\[
\tilde{F}V(a, s_{it}) = E[\Phi_{it+1} | s_{it}, A_{it} = a] \hat{\lambda}.
\]

Using these approximate continuation values, I construct the choice-specific values of applying and not applying by

\[
V^A(a = 1, s_{it}) = \Pi^A(a, s_{it}) + \beta(1 - p_{it})\tilde{F}V(1, s_{it})
\]
\[
V^A(a = 0, s_{it}) = \tilde{F}B(b, s_{it+1})[\pi_0(s_{it+1}) + E[E_{it}^B(1) | b = 1]] + [1 - \tilde{F}B(b, s_{it+1})] \tilde{F}V(b = 0, s_{it}),
\]

where, with some misuse of notation, \( \tilde{F}V(b = 0, s_{it}) = E[\Phi_{it+1} | s_{it}, B_{it} = 0] \hat{\lambda} \). By Equations 4 and
5, I can also form choice-specific values for the building decision:

\[ V^B(b = 1, s_{it}) = \pi_0, \quad V^B(b = 0, s_{it}) = \hat{F}(b = 0, s_{it}). \]

In the second step of PPI, termed “policy improvement,” I use the choice-specific values to update the choice probabilities. In particular,

\[
\Pr(A_{it} | s_{it}) = \Lambda\left(\sigma^{-1}_a [V^A(1, s_{it}) - V^A(0, s_{it})]\right) \quad \text{and} \quad \Pr(B_{it} | s_{it}) = \Lambda\left(\sigma^{-1}_b [V^B(1, s_{it}) - V^B(0, s_{it})]\right)
\]

where the logistic function is \( \Lambda(\cdot) = \exp(\cdot)/(1 + \exp(\cdot)) \).

The iteration can now proceed. I use the probabilities from the latest policy-improvement step to recomputed the policy-valuation step, iterating until the choice probabilities converge.

### 7.3 Part 2: Simulated Minimum Distance

The second part of the estimation procedure (SMD) finds the structural parameters that minimize a distance between the actual data and simulated data from the first part. Under SMD, I am free to target a general set of moments in estimation. I use this freedom to target the quasi-experimental moments estimated above, along with key descriptive patterns in the data.

Formally, the SMD step finds parameter estimates \( \hat{\theta} \) that jointly minimize a distance between empirical moments \( \hat{\beta} \) and their simulation analogs \( \tilde{\beta}(\theta) \). In particular, the estimates \( \hat{\theta} \) solve

\[
\arg \min_{\theta} [\hat{\beta} - \tilde{\beta}(\theta)]^T \Sigma^{-1} [\hat{\beta} - \tilde{\beta}(\theta)],
\]

where \( \Sigma^{-1} \) is a weight matrix. \(^{35}\) I target three groups of empirical moments \( \hat{\beta} \): (1) causal effects of winning a grant using quasi-random assignment, from Section 4; (2) application supply responses to changes in grant value, from Section 5; and (3) developer bidding behavior in response to incentives to reduce rents, from Section 6. These are joined by moments to match descriptive patterns.

**Winners Versus Losers.** In the simulated data, I can compare winners and losers, controlling for win probability (Equation 9). The model presumes that winners must build, so I estimate an equivalent specification among losers only: \( B_{it} = X_{it} \gamma_0 + \gamma_1 p_{it} + u_{it} \), where \( B_{it} \) is an indicator for the development outcome and \( X_{it} \) contains application characteristics. The coefficients \( \gamma_0 \) and \( \gamma_1 \) are targeted moments.

**Application Supply.** First, I match application responses to the boost. The simulated response is calculated as an average derivative across all potential applicants, and I target the event-study

---

\(^{35}\)I specify \( \Sigma \) as block-diagonal. Each block corresponds to the covariance matrix of moments that are readily estimated simultaneously, such as a vector of coefficients from the same regression.
estimate for the boost’s effect on the annual probability (in percentage points) of any application from the tract. Second, I match the boost’s (small) effects on the composition of the applicant pool from the RD estimates, targeting effects on win probabilities, credits requested per unit, and proposed unit counts (Appendix Figure A19). Third, I match the coefficients of a cross-sectional regression of whether a tract has any applications in a given year on tract characteristics.

**Bidding.** I target the fixed-effects LPM coefficients from Column 1 of Table 3. I enrich the regression specification in the table by interacting the win-probability and rent variables with the two tract characteristics, poverty and population density.

**Distributional Moments.** I target the means and variances of application characteristics: the number of units and the credit amount requested per unit. I also target the mean and variance of the distribution of win probabilities. Among reapplicants, I match the coefficients of a regression of their current-round win probabilities on their prior-round probabilities.

### 7.4 Coefficient Estimates and Model Fit

I briefly discuss the parameter estimates in Appendix Table A3, reserving further discussion for the distributions of model primitives in Section 8, which are more readily interpreted. Parameter estimates imply that potential applicants with high unit counts or in dense areas have stronger outside options on average. Applications with larger credit requests per unit generally have weaker outside options. There is state dependence: Applying improves applicants’ outside options, and the cost of reapplying is lower than the cost of the initial application.

The tax credit makes up a larger share of win value in low-income areas and for projects with high credits per unit. There are substantial returns to scale in entry costs, measuring size both by unit count and in credits requested per unit. Entry costs are lower in poor and low-density areas. The persistent unobservable induces a strong positive correlation in outside options and win values, implying there is advantageous selection on outside options into the LIHTC.

Appendix Table A4 reports information on the model’s fit on targeted moments. Model fit appears strong. This is informative in itself: The model does not seem to struggle to explain developer behavior across three distinct empirical analyses. The largest gaps in fit occur in the cross-sectional regression of application supply on tract characteristics.

### 8 Results and Counterfactuals

This section reports model results. After reviewing estimates of model primitives, I turn to the paper’s two main questions: the LIHTC’s housing-market impacts and its incidence. I also assess its cost-effectiveness in comparison to a stylized voucher program.
### Table 4: Estimates of Model Primitives and Potential-Applicant Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Applicants</th>
<th></th>
<th></th>
<th>Non-Applicants</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>P25</td>
<td>P75</td>
<td>Median</td>
<td>P25</td>
<td>P75</td>
</tr>
<tr>
<td>Ex-Ante Value</td>
<td>1.41</td>
<td>1.13</td>
<td>1.76</td>
<td>0.56</td>
<td>0.25</td>
<td>0.94</td>
</tr>
<tr>
<td>Outside Option</td>
<td>1.25</td>
<td>0.88</td>
<td>1.65</td>
<td>-0.03</td>
<td>-0.71</td>
<td>0.63</td>
</tr>
<tr>
<td>Win Value</td>
<td>1.45</td>
<td>1.01</td>
<td>1.89</td>
<td>0.47</td>
<td>-0.11</td>
<td>1.06</td>
</tr>
<tr>
<td>Entry Cost</td>
<td>0.21</td>
<td>0.08</td>
<td>0.31</td>
<td>0.55</td>
<td>0.40</td>
<td>0.65</td>
</tr>
</tbody>
</table>

**Panel A: Model Primitives as a Share of Basis**

**Panel B: Model Primitives in Thousands of Dollars Per Unit**

**Panel C: Potential-Applicant Characteristics**

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Win Probability</td>
<td>0.48</td>
<td>0.05</td>
<td>0.71</td>
<td>0.29</td>
<td>0.10</td>
<td>0.55</td>
</tr>
<tr>
<td>Unit Count</td>
<td>51</td>
<td>34</td>
<td>89</td>
<td>43</td>
<td>31</td>
<td>59</td>
</tr>
<tr>
<td>Tax Credits Per Unit</td>
<td>218</td>
<td>137</td>
<td>301</td>
<td>184</td>
<td>133</td>
<td>259</td>
</tr>
</tbody>
</table>

**Notes:** Panels A and B report estimates of the model primitives, the ex-ante value of the application value function $E[V^A(s_{it}, \epsilon_{it}) | s_{it}]$, the outside option $\pi_0(s_{it})$, the win value $\pi_1(s_{it})$, and the entry cost $\kappa(s_{it})$. Panel C reports estimates of potential-applicant characteristics. Tax credits per unit are in thousands of dollars. Appendix Table A6 provides a similar analysis for winning and losing applicants.

### 8.1 Model Primitives and Potential-Applicant Characteristics

Panels A and B of Table 4 reports estimates of the key model primitives. These are the (ex-ante) application value function $E[V^A(s_{it}, \epsilon_{it}) | s_{it}]$, the outside option $\pi_0(s_{it})$, the win value $\pi_1(s_{it})$, and the entry cost $\kappa(s_{it})$. To summarize the heterogeneity, I report medians and interquartile ranges. In Panel A, I report these values as shares of the potential application’s eligible basis. Panel B rescales the values into dollar terms.

Estimates of $\pi_1$ show winning the LIHTC is valuable to applicants. Taking the reciprocal, the credits are worth about the full project value for the median applicant ($1/(0.7 \cdot 1.45) = 1.01$). Beyond the grant, the project value reflects construction costs, future rents, and other subsidies, so it could be greater or less than the grant amount in principle.

For the median applicant, winning is about 16 percent ($1.45/1.25 - 1 = 0.16$) better than the outside option. Thus, the LIHTC is narrowly “buying out” such an applicant’s next-best-alternative land use. Moreover, outside options are positive (i.e., statically preferred to not building that
period) for nearly all applicants. In dollar terms, outside options appear plausible by comparison to capitalized rents—that is, the former is less than the latter under reasonable discount rates. This discussion illuminates how the model reconciles the application responses and the award effects: by making applicants marginal to applying but not marginal to development. Applicants vary greatly in win values and outside options, which is ultimately the source of developer incidence.

I now turn to the win values and outside options of non-applicants. Win values generally exceed outside options, and outside options are negative for approximately half of non-applicants, implying much land with little productive use. Model results suggest the key driver of selection into application is variation in entry costs: The median applicant has small entry costs, whereas the median non-applicant would burn the subsidy’s entire value in applying. Such an applicant would not find it profitable to apply even if they would win with certainty. This conclusion is natural: Had selection into application been driven by variation in outside options, for instance, losers would be unlikely to develop privately. Overall, estimated entry costs appear large, at 10 percent of win value for the median applicant.36

I also examine the observable characteristics of potential applicants. Applicants are sensibly selected on their win probabilities relative to non-applicants. Relative to applicants, non-applicants would propose slightly smaller projects (in unit count) if they were to apply, and they would be eligible for less in tax credits per unit.

### 8.2 Incidence and Impacts

Table 5 shows the model-based estimates of incidence and impacts. Column 1 reports overall averages. Columns 2 to 5 split the simulation according to whether a potential applicant’s tract is above or below the median on two characteristics, the poverty rate or the population density.

Overall, I find that households receive about 30 percent of the welfare gains from the LIHTC. Instead, the LIHTC is heavily incident on developer profits and entry costs. The entry-cost share has risen relative to the comparisons of primitives above, as costs are paid by both winners and losers, whereas other impacts result only for winners. The gains to households are split between rent savings and general-equilibrium effects but mostly arise from the former. LIHTC residents are poorer than the typical resident of the same tract, so the general-equilibrium component is likely less targeted than the rent-savings component (see Appendix Figure A23).

The model also finds that subsidized units substantially displace other construction on the same parcel and not one-for-one net additions to the local housing stock. For every ten subsidized units, there are about two net units added to the housing stock, an estimate that is consistent with

36These estimated entry costs are difficult to benchmark against real-world costs and are ultimately a structural object needed to explain entry behavior. Although I take entry costs as arising per application, there is likely a large fixed component at the developer corporate level.
Table 5: Model-Based Incidence and Impacts

<table>
<thead>
<tr>
<th></th>
<th>(1) Poverty Rate</th>
<th>(2) Poverty Rate</th>
<th>(3) Poverty Rate</th>
<th>(4) Population Density</th>
<th>(5) Population Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household Share</td>
<td>0.314</td>
<td>0.265</td>
<td>0.361</td>
<td>0.391</td>
<td>0.396</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.026)</td>
<td>(0.029)</td>
<td>(0.047)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>Rent Savings</td>
<td>0.233</td>
<td>0.180</td>
<td>0.284</td>
<td>0.265</td>
<td>0.177</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.016)</td>
<td>(0.022)</td>
<td>(0.022)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Gen. Eqm. Effects</td>
<td>0.081</td>
<td>0.085</td>
<td>0.076</td>
<td>0.125</td>
<td>0.219</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.023)</td>
<td>(0.021)</td>
<td>(0.045)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Developer Share</td>
<td>0.437</td>
<td>0.480</td>
<td>0.401</td>
<td>0.397</td>
<td>0.382</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.032)</td>
<td>(0.034)</td>
<td>(0.058)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>Entry Cost Share</td>
<td>0.250</td>
<td>0.255</td>
<td>0.239</td>
<td>0.212</td>
<td>0.222</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.030)</td>
<td>(0.022)</td>
<td>(0.045)</td>
<td>(0.020)</td>
</tr>
</tbody>
</table>

Panel B: Impacts

<table>
<thead>
<tr>
<th></th>
<th>(4) Displacement Rate</th>
<th>(5) Cost Per Net Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Displacement Rate</td>
<td>0.777</td>
<td>0.758</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Cost Per Net Unit</td>
<td>960</td>
<td>870</td>
</tr>
<tr>
<td></td>
<td>(100)</td>
<td>(98)</td>
</tr>
</tbody>
</table>

Notes: This table reports model-based estimates of the incidence and impacts of the LIHTC. The displacement rate is one minus the net change in the housing stock per LIHTC unit. For tract characteristics, I split at the median among simulated applicants. Costs per unit are in thousands of dollars. Bootstrap standard errors are reported in parentheses.

Baum-Snow and Marion (2009). Unlike the results in Section 4, however, these estimates account for reapplication, general-equilibrium effects, and state dependence in behavior.

There are notable spatial differences in the displacement rate and thus in costs per net unit. Displacement is weaker in high-density than in low-density areas, consistent with variation in local housing-supply elasticities. Given such displacement, the LIHTC’s cost per net new unit varies sharply—and essentially inversely with its “accounting” cost per subsidized unit. Spatial heterogeneity in incidence is also considerable. In low-poverty areas, the LIHTC provides sizable transfers to households, consistent with spatial variation in rent savings. In other areas, the LIHTC is more of a pure subsidy for development with less redistributive benefit.

Appendix Table A5 analyzes the sensitivity of the structural results. Following Andrews et al. (2017), I report the changes in estimates of model primitives, incidence, and impacts that result from small perturbations to the empirical moments. There are three important lessons. First, the direction of sensitivity of parameters to moments is generally intuitive. For instance, a higher
application-supply elasticity would push down estimated entry costs. Second, specific equations do not identify specific structural parameters in isolation, so estimating the model jointly over the moments does matter. Third, incidence estimates are especially sensitive to reapapplication behavior, a result that suggests the value of dynamic models in this context.

8.3 Vouchers as Policy Alternative

The government could replace the LIHTC with an expansion of vouchers or other tenant-based subsidies. I evaluate this policy alternative in two model-based counterfactuals. The first replaces the LIHTC with a stylized voucher program that achieves the same aggregate benefit to households as the LIHTC, while allowing the fiscal cost to change. The second holds the fiscal cost fixed and allows the transfer to households to change.

Figure 11 reports results for the first counterfactual. Holding household utility fixed, vouchers reduce fiscal costs by 25 percent, or $56,000 per LIHTC unit in present value. This difference seems modest in light of economists’ objections to project-based assistance. I next interpret the model forces that determine this cost difference, before discussing considerations beyond the model.

Vouchers and the LIHTC impose opposite-signed pecuniary externalities on market rents, so all else equal, the government must make a larger outlay under vouchers to achieve the same overall welfare gain for households. I find that, to offset such impacts on unassisted households, vouchers must make a transfer to assisted tenants that is more than twice the per-unit present value of the rent.
discounts received by LIHTC tenants. Weighing in the other direction, vouchers avoid the LIHTC’s entry costs, and they incur less incidence on developer profits. On balance, these latter forces leave vouchers with a modest fiscal advantage over the LIHTC. Due to the canceling-out of these two forces, my estimates of the compensated difference in fiscal costs are quite close to accounting-style comparisons of housing policies: Relative to public housing and various project-based subsidies over the decades, they also generally find moderate advantages for vouchers (Weicher, 2012).

In the fixed-budget counterfactual, the switch to vouchers expands the housing stock and raises household welfare. Holding the budget fixed, vouchers’ effect on the aggregate housing stock is about 40 percent larger than that of the LIHTC. The total welfare benefit to households would be about 60 percent larger, rising from from $87,000 per LIHTC unit to $138,000. In both counterfactuals, I find little difference in where vouchers and the LIHTC add units on net in terms of tract poverty rates, but the LIHTC seems to be somewhat more pro-density than vouchers.

The model omits a variety of other interesting differences between tenant- and project-based assistance. First, as noted above, project-based programs facilitate the provision of other supports, particularly healthcare, alongside housing. This aspect is challenging to value and possibly applies to the half of LIHTC properties that provide elderly or supportive housing (Table 1). Second, the LIHTC’s federal tax expenditure understates its total cost: The developments often also receive subsidized debt, state-level assistance, and reductions in local property taxes. Third, vouchers have administrative costs ignored here. Finally, under a social welfare function, the income differences between assisted tenants and nearby non-assisted households would expand the voucher advantage.

9 Conclusion

Governments provide housing assistance to lower-income people in three main ways: public housing, rent vouchers, and project-based subsidies. Understanding the relative costs and benefits of these options has been for decades a core question at the intersection of public finance and urban economics. Economists are predisposed to favor rent vouchers over project-based subsidies, but the modern evidence for those views is limited. How does the largest U.S. project-based subsidy—the LIHTC—affect housing markets, and who benefits from it?

I assess LIHTC’s impacts and incidence using new data and three quasi-experimental research designs to estimate a dynamic model of developer behavior. I find developer subsidies have limited effect on the overall size of the housing stock, but instead they reallocate new housing progressively. Developer competition does benefit households, but in the process it generates substantial entry costs, and subsidies remain highly incident on developers. In counterfactuals that compare the LIHTC to a stylized rent voucher, I find a modest advantage for vouchers. These findings may help to explain the coexistence of project-based and tenant-based assistance in U.S. housing policy.
References


Appendices for Online Publication

A Additional Tables and Figures 2

B Supplemental Information 36

C Theoretical and Structural Appendix 46
A Additional Tables and Figures

Figure A1: Data Coverage

Notes: This figure shows the coverage of the LIHTC application data. Blue indicates coverage, red indicates non-coverage, and gray indicates that the state did not hold an LIHTC round in that year.
Figure A2: Neighborhood Effects of Tax Credit Awards: Controlling for Application Characteristics

Change in Occupied Housing Stock (p.p.)

Notes: The figure plots win and lose effects on the occupied housing stock (change in percentage points), as in Figure 7. The specification is an event-study analog of Equation 9 that includes non-applicant tracts: \( \Delta \log E[Y_{i,t+k}] = \alpha \text{Win}_{it} + \beta \text{Lose}_{it} + X_{it} \gamma_k \), for \( k \) quarters or years from the award. It is estimated via Poisson regression. The key distinction from Figure 7 is that it includes in controls \( X_{it} \), the two application characteristics on which there is significant imbalance: whether the application is entirely new-construction and the leave-out win rate of the developer. The color bands depict pointwise 95-percent confidence intervals, and standard errors are clustered at the tract level.
Figure A3: Neighborhood Effects of Tax Credit Awards, Instrumenting for Actual Wins with Simulated Wins

Change in Occupied Housing Stock (p.p.)

Notes: This figure plots win and lose effects on the occupied housing stock (change in percentage points). The specification is an instrumental-variables (IV) version of Figure 7, where I instrument for the actual tax-credit assignment of an application with its simulated assignment. I implement the IV within the Poisson regression using a control function. The color bands depict pointwise 95-percent confidence intervals, and standard errors are clustered at the tract level.
Figure A4: Parcel-Level Effects of Tax Credit Awards (Weighted)

Panel A: Total Development

Notes: Panel A plots win and lose effects (on left) and the win–lose difference (on right), of tax credit awards on the probability that construction is recorded as completed in a given year. Panel B decomposes these development impacts into LIHTC, other-subsidy, and unsubsidized development. Multiple parcels associated with a single application are aggregated using floor-space weighting. Standard errors in Panel A are clustered by the consolidated parcel.
Figure A5: Parcel-Level Effects of Tax Credit Awards: Heterogeneity by Win Probability

Panel A: Event Study

Panel B: Displacement Rate Summary

Notes: Panel A plots effects of winning versus losing a tax credit award on the probability that construction is recorded as completed in a given year. The figure reports these estimates separately by splitting the sample into quarter intervals of the application’s win probability. Panel B reports the 10-year displacement rate by interval of the win probability. Standard errors in Panel A are clustered by the consolidated parcel.
Figure A6: Winners-Versus-Losers Comparison in Neighborhood Event Study

Notes: This figure plots win–lose difference coefficients in an event-study comparison of Census tracts with LIHTC applications. The blue line shows baseline estimates, which only include county–year fixed effects as controls. The orange line augments this specification with the flexible control for win probability. The black line adds controls for Census tract pre-award characteristics, as listed in Section 4. The color bands depict pointwise 95-percent confidence intervals, and standard errors are clustered at the tract level.
Figure A7: Detailed Land-Use Effects of Tax Credit Awards

Panel A: Single- Versus Multi-Family

![Graph showing the number of households over time for likely multi-family and single-family households.]

Panel B: Rental Versus Owner-Occupied

![Graph showing the number of households over time for likely renters and owner-occupiers.]

Notes: Panel A plots win and lose effects on the number of households (change in percentage points) in the same Census block group as the LIHTC application who reside in multi-family or single-family residences. Panel B reports the same effects but for households who are likely renters versus likely owner-occupants. In both, the specification is an event-study analog of Equation 9 that includes non-applicant tracts: $\Delta \log E[Y_{i,t+k}] = \alpha_k \text{Win}_{it} + \beta_k \text{Lose}_{it} + X_{i,t} \gamma_k$, for $k$ quarters or years from the award. The specification is estimated via Poisson regression, and it includes controls $X_{i,t}$ for the win probability, pre-award tract characteristics, and the baseline level of the outcome. In both panels, the color bands depict pointwise 95-percent confidence intervals, and standard errors are clustered at the tract level.
Notes: This figure plots win and lose effects on the number of households by income decile in the same Census block group as the LIHTC application. The time horizon is 10 years after the award, and I rescale the effects so that they represent contributions to the aggregate percentage change in the block-group household count. The specification is an event-study analog of Equation 9 that includes non-applicant tracts: \( \Delta \log E[Y_{i,t+k}] = \alpha_k \text{Win}_{it} + \beta_k \text{Lose}_{it} + X_{it} \gamma_k \), for \( k \) quarters or years from the award. The specification is estimated via Poisson regression, and it includes controls \( X_{it} \) for the win probability, pre-award tract characteristics, and the baseline levels of the outcomes. That is, to address measurement error in household income, all specifications includes control for all baseline income-decile household counts. The gray bars depict 95-percent confidence intervals, and standard errors are clustered at the tract level.
Figure A9: Effects of Tax Credit Awards on the Subsidized Housing Stock

Notes: This figure plots win and lose effects on four categories of the subsidized housing stock. The outcome is defined by taking annual HUD counts of subsidized units by Census tract and dividing it by the count of households in 2000. The numerator is from the Picture of Subsidized Households database, and the denominator is from the 2000 Census. The specification is an event-study analog of Equation 9 that includes non-applicant tracts: \( \Delta \log E[Y_{t,t+k}] = \alpha_k \text{Win}_{it} + \beta_k \text{Lose}_{it} + X_{it} \gamma_k \), for \( k \) quarters or years from the award. The specification is estimated via Poisson regression, and it includes controls \( X_{it} \) for the win probability, pre-award tract characteristics, and the baseline level of the outcome. In both panels, the color bands depict pointwise 95-percent confidence intervals, and standard errors are clustered at the tract level.
Figure A10: Replication of Tract-Level Event Study (Data Axle)

Notes: This plots win and lose effects effects on the change (in percentage points) in the household count in the LIHTC application’s Census tract. The specification is intended for comparison to Figure 7, and it includes controls for the win probability and pre-award tract characteristics. The color bands depict pointwise 95-percent confidence intervals, and standard errors are clustered at the tract level.
Figure A11: Neighborhood Effects of Tax Credit Awards, No Win-Probability Control

Panel A: Tract (USPS)

Change in Occupied Housing Stock (p.p.)

Years from LIHTC Competition

Win
Lose

Panel B: Block Group (Data Axle)

Same Block Group
Change in Households (p.p.)

Same Tract, Different Block Group

Years from LIHTC Competition

Win
Lose

Notes: Panel A plots win and lose effects on the occupied housing stock (change in percentage points) from a quarterly event-study analog of Equation 9 that includes non-applicant tracts. Panel B plots win and lose effects effects on the change (in percentage points) in the household count in the same Census block group as the LIHTC application, or in the same tract but other block groups. All specifications include controls for pre-award tract characteristics but exclude the win-probability control. See Figure 7 for comparison. In both panels, the color bands depict pointwise 95-percent confidence intervals, and standard errors are clustered at the tract level.
Notes: This figure plots estimates of cohort-specific quarterly coefficients $\beta_{k,y}$ from an event-study specification that is estimated on the subsample of tracts that have application in year $y$ or that never apply. Estimated via a Poisson regression, the specification is $\Delta \log E[Y_{i,t+k}] = \alpha_{k,y} \text{Win}_{it} + \beta_{k,y} \text{Lose}_{it} + X_{it} \gamma_{k,y}$ for each time horizon $k$ and cohort $y$. Each grey line traces a path of coefficients for the same cohort $y$ over the quarters $k$ relative to the quarter of application. The solid black line displays the precision-weighted averages of these cohort-specific coefficients at each event time horizon.
Figure A13: Application Supply Responses to Changes in Eligibility for Basis Boost: Falsification Test

Notes: This figure plots event-study coefficients from Equation 10 for various application supply outcomes. The “actual” coefficients plot responses for actual application-supply outcomes. The “predicted” coefficients plot responses for an outcome constructed as the predicted values of year-specific regressions on the QCT running variables. Entry and exit from boost are assumed to have symmetric effects. All specifications are estimated by Poisson regression and include state–year fixed effects as controls. The bands show pointwise 95-percent confidence intervals, with standard errors clustered by tract.
Figure A14: Application Supply Responses to Changes in Eligibility for Basis Boost: Secondary Outcomes

Notes: This figure plots event-study coefficients from Equation 10 for various application supply outcomes. Entry and exit from boost are assumed to have symmetric effects. All specifications are estimated by Poisson regression and include state–year fixed effects as controls. The bands show pointwise 95-percent confidence intervals, with standard errors clustered by tract.
Figure A15: Application Supply Responses to Changes in Eligibility for Basis Boost

Notes: This figure plots event-study coefficients from Equation 10 for various application supply outcomes. Entry and exit from boost are assumed to have symmetric effects. In the baseline specification, I include state–year fixed effects. The other specifications introduce county–year fixed effects or tract controls interacted with year indicators. The tract controls are decile-group indicators for the tract’s poverty rate, population density, and cumulative change in population from 1990 to 2000. The bands show pointwise 95-percent confidence intervals, with standard errors clustered by tract.
Figure A16: Application Supply Responses to Changes in Eligibility for Basis Boost: Include Always-Boosted Tracts

Notes: This figure plots event-study coefficients from Equation 10 for various application supply outcomes. The sole difference from Appendix Figure A14 is that here I include always-boosted tracts in the sample. Entry and exit from boost are assumed to have symmetric effects. All specifications are estimated by Poisson regression and include state–year fixed effects as controls. The bands show pointwise 95-percent confidence intervals, with standard errors clustered by tract.
Figure A17: Application Supply Responses to Changes in Eligibility for Basis Boost: QCT and DDA Effects

Notes: This figure plots event-study coefficients from Equation 10. Entry and exit from boost are assumed to have symmetric effects, but I allow the two causes of basis boosts (Qualified Census Tract [QCT] and Difficult Development Area [DDA]) to have different effects on application supply. The specification is estimated by Poisson regression and includes state–year fixed effects as controls. The bands show pointwise 95-percent confidence intervals, with standard errors clustered by tract.
Figure A18: LIHTC Application Volume Around the Qualified Census Tract Threshold: Alternative Outcomes

Notes: This figure shows the impact of the Qualified Census Tract (QCT) threshold on four measures of tract-level LIHTC application volume. In all panels, the running variable is a distance defined with respect to tract rank within its metropolitan area. All panels split the data into 15 equal-interval bins on either side of the QCT threshold.
Figure A19: Covariate Smoothness Around the Qualified Census Tract Threshold

Notes: This figure examines the conditional expectation function for six covariates around the Qualified Census Tract (QCT) threshold. In all panels, the running variable is a distance defined with respect to tract rank within its metropolitan area. All panels split the data into 15 equal-interval bins on either side of the QCT threshold. The “relative income” measure is an index, standardized to zero mean and unit standard deviation, used in QCT assignment that compares a tract’s median household income to its metro-area median.
**Figure A20: Application Characteristics Around the Qualified Census Tract Threshold**

*Notes:* This figure shows average application characteristics around the threshold of Qualified Census Tract (QCT). In all panels, the running variable is a distance defined with respect to tract rank within its metropolitan area, and the data is split into 15 equal-interval bins on either side of the QCT threshold. The RD estimate and standard error corresponding to the plotted outcome is reported in the upper-right corner of each panel. LIHTC request per unit is in thousands of dollars.
Figure A21: Nonparametric Estimates of Bunching at Rent Score Kink

Notes: This figure depicts estimated effects of kinked incentives in Qualified Allocation Plans for setting rents in LIHTC units. The coefficients come from a Poisson regression \( \log E[A_{rm}] = \alpha_r + \alpha_m + \sum_s \beta_s [r - m = s] \), where \( A_{rm} \) is the count of applications with rent level \( r \) given that the score-maximizing rent level is \( m \). The coefficients \( \alpha_r \) and \( \alpha_m \) are respectively fixed effects for an application’s rent level and for an application round’s score-maximizing rent level. “FE” bunching coefficients \( \beta_s \) are identified from heterogeneity in the score-maximizing rent level. The “naive” specification omits the fixed effects \( \alpha_r \) and \( \alpha_m \), pooling variation. An application’s rent level is defined relative to the local annual median income (AMI) of the household to whom its units would be “affordable,” that is, no greater than 30 percent of income. To obtain counts, I bin the application data by rounding to full percentage points of AMI. To facilitate interpretation, the Poisson coefficients are exponentiated to \( \exp(\beta_s) - 1 \).
Figure A22: Heterogeneity in the Trade-Off Between Win Probability and Rental Income

Notes: The left panel is a histogram of the marginal rate of transformation (MRT), local to the developer’s actual choice. The MRT is defined as a ratio of the percentage-point change in win probability to the change in the present discounted value of rental income per unit, measured in thousands of dollars. To aid visualization, I winsorize the distribution at MRT < −10, and I drop the all observations with MRT = 0. The shares reported on the vertical axis are thus conditional on a non-zero MRT. The right panel is a binned scatterplot of the probability the developer accepts a choice giving them a higher rental income but lower win probability, scattered with respect to the MRT associated with this choice.
Figure A23: Income Comparison of LIHTC Residents to Other Locals

Notes: This figure plots the distribution of a ratio of median incomes. The ratio compares the median household income of LIHTC residents at the property level to the median in the corresponding Census tract. Median-income data for LIHTC residents is as of 2019; the tract data come from the ACS centered on 2019. The histogram is weighted by each LIHTC property’s total unit count.
Figure A24: Map of LIHTC Applications

Notes: This figure displays a map of the tax-credit applications. States are shaded in grey if no data are available.
Figure A25: Counts of Applications and Proposed Units by Year and Outcome

Panel A: Applications

Panel B: Proposed Units

Notes: This figure plots, in Panels A and B respectively, the counts of applications and proposed total units (both income-restricted and unrestricted) in the data for each year and application outcome.
Figure A26: Round-Level Distributions of Average Simulated Win Probability and Explained Share of Variance in Wins

Notes: This figure displays, in the left panel, the distribution of average simulated win probabilities at the level of tax-credit competition round. In the right panel, the figure displays the round-level distribution of the share of variance of wins that is explained by variance in simulated win probabilities, that is, \( \frac{\text{Var}(\hat{\beta}_i)}{\text{Var}(\text{Win}_i)} \). The vertical lines denote unweighted averages of the average simulated win probability (left) and the explained share of variance in wins (right).
Figure A27: Analysis of Self-Scores

Notes: This figure displays, in the left panel, a binned scatterplot of actual application scores versus self-scores. I transform scores into percentile ranks within the distribution of actual scores. In the right panel, the figure displays a binned scatterplot comparing estimates of the win probability using actual and self-scores. The procedure to compute win probabilities using self-scores is described in Appendix B.
Table A1: Heterogeneous Application Supply Responses to Qualified Census Tract Threshold

<table>
<thead>
<tr>
<th></th>
<th>Main Effect</th>
<th>Interaction Effect</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est. (1)</td>
<td>SE (2)</td>
<td></td>
</tr>
<tr>
<td>QCT Threshold</td>
<td>0.592***</td>
<td>(0.084)</td>
<td></td>
</tr>
<tr>
<td>Panel A: High Population Growth</td>
<td></td>
<td>-0.291***</td>
<td>0.043</td>
</tr>
<tr>
<td>QCT Threshold</td>
<td>0.435***</td>
<td>(0.100)</td>
<td></td>
</tr>
<tr>
<td>Panel B: High Rental Vacancy</td>
<td></td>
<td>0.019</td>
<td>0.063</td>
</tr>
<tr>
<td>QCT Threshold</td>
<td>0.467***</td>
<td>(0.084)</td>
<td></td>
</tr>
<tr>
<td>Panel C: High Population Density</td>
<td></td>
<td>-0.062</td>
<td>3.85</td>
</tr>
<tr>
<td>QCT Threshold</td>
<td>0.501***</td>
<td>(0.116)</td>
<td></td>
</tr>
<tr>
<td>Panel D: High Poverty Rate</td>
<td></td>
<td>-0.156</td>
<td>0.230</td>
</tr>
<tr>
<td>QCT Threshold</td>
<td>0.383***</td>
<td>(0.090)</td>
<td></td>
</tr>
<tr>
<td>Panel E: High Non-Hispanic White Share</td>
<td></td>
<td>0.123</td>
<td>0.539</td>
</tr>
</tbody>
</table>

Notes: This table estimates heterogeneous effects of the Qualified Census Tract (QCT) threshold on application supply according to Census tract characteristics. In all panels, the outcome measure is the application count in that tract–year. Effects are estimated using a local-linear Poisson regression with a triangular kernel of bandwidth 0.2 around the QCT threshold. Main effects of the QCT threshold are reported in Columns 1 and 2, and interaction effects are reported in Columns 3 and 4. In estimating heterogeneous effects, I split Census tracts at the median of the named variable, with the median reported in Column 5. * = p < 0.10, ** = p < 0.05, *** = p < 0.01.
Table A2: Estimating Developer Win Valuations from Bidding Behavior: Binding Subsample

<table>
<thead>
<tr>
<th></th>
<th>(1) OLS</th>
<th>(2) IV</th>
<th>(3) Cond. Logit.</th>
<th>(4) + Ctrl. Funct.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Win Probability</td>
<td>0.407***</td>
<td>1.801***</td>
<td>1.199***</td>
<td>6.277***</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.057)</td>
<td>(0.091)</td>
<td>(0.240)</td>
</tr>
<tr>
<td>Log Average Rent</td>
<td>1.029***</td>
<td>3.667***</td>
<td>2.970***</td>
<td>11.956***</td>
</tr>
<tr>
<td></td>
<td>(0.073)</td>
<td>(0.144)</td>
<td>(0.216)</td>
<td>(0.548)</td>
</tr>
<tr>
<td>Applications</td>
<td>4.274</td>
<td>4.274</td>
<td>4.268</td>
<td>4.268</td>
</tr>
<tr>
<td>Marg. Rate of Substitution</td>
<td>1.590</td>
<td>1.282</td>
<td>1.558</td>
<td>1.199</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.029)</td>
<td>(0.057)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Mean Win Value Per Unit</td>
<td>$32,426</td>
<td>$40,235</td>
<td>$33,090</td>
<td>$43,019</td>
</tr>
<tr>
<td></td>
<td>(1,244)</td>
<td>(971)</td>
<td>(1,247)</td>
<td>(931)</td>
</tr>
<tr>
<td>Developer Incidence Share</td>
<td>0.255</td>
<td>0.317</td>
<td>0.260</td>
<td>0.338</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.009)</td>
<td>(0.011)</td>
<td>(0.009)</td>
</tr>
</tbody>
</table>

Notes: This table reports estimates of coefficients in Equation 13. I use the subsample of the application data in which I estimate that rent regulations bind. In Column 4, I take a control-function approach to instrument for the win probability in the conditional logit. Marginal rates of substitution are calculated as a ratio of coefficients; see the text for detail on the other calculations. Standard errors are clustered by application. * = \( p < 0.10 \), ** = \( p < 0.05 \), *** = \( p < 0.01 \).
Table A3: Structural Parameter Estimates

<table>
<thead>
<tr>
<th>Group</th>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outside Option</td>
<td>Intercept</td>
<td>-0.064</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Poverty Rate</td>
<td>-0.029</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Log Population Density</td>
<td>0.197</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Log Unit Count</td>
<td>0.277</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Log Credits Per Unit</td>
<td>-0.347</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Permanent Unobs.</td>
<td>0.924</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>Reapplicant</td>
<td>0.396</td>
<td>0.003</td>
</tr>
<tr>
<td>Net Value of Win</td>
<td>Intercept</td>
<td>0.388</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>Poverty Rate</td>
<td>-0.465</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>Log Population Density</td>
<td>0.171</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Log Unit Count</td>
<td>-0.263</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Log Credits Per Unit</td>
<td>-0.574</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>Permanent Unobs.</td>
<td>0.731</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>Reapplicant</td>
<td>0.513</td>
<td>0.007</td>
</tr>
<tr>
<td>Entry Cost</td>
<td>Intercept</td>
<td>0.584</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Poverty Rate</td>
<td>-0.212</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Log Population Density</td>
<td>-0.015</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Log Unit Count</td>
<td>-0.203</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Log Credits Per Unit</td>
<td>-0.103</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Permanent Unobs.</td>
<td>-0.073</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Reapplicant</td>
<td>-0.386</td>
<td>0.004</td>
</tr>
<tr>
<td>Win Probability</td>
<td>Parameter 1</td>
<td>1.270</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>Parameter 2</td>
<td>0.686</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Parameter 3</td>
<td>0.759</td>
<td>0.008</td>
</tr>
<tr>
<td>Temp. Unobs. Dispersion</td>
<td>Application Cost</td>
<td>0.005</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Outside Option</td>
<td>0.011</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Bidding</td>
<td>0.097</td>
<td>0.007</td>
</tr>
<tr>
<td>Log Units</td>
<td>Mean</td>
<td>3.804</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.497</td>
<td>0.006</td>
</tr>
<tr>
<td>Log Credits Per Unit</td>
<td>Mean</td>
<td>12.142</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.492</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Notes: This table reports estimates of the model structural parameters, along with standard errors. * = \( p < 0.10 \), ** = \( p < 0.05 \), *** = \( p < 0.01 \).
| Group                        | Moment                  | Data  | Simulation | $|t|$ |
|-----------------------------|-------------------------|-------|------------|-----|
|                             |                         | Estimate | SE | Estimate | SE |       |
| Application: Cross-Sectional Regression | Intercept               | 0.030    | 0.001     | 0.036   | 0.003   | 2.07  |
|                             | Poverty Rate            | 0.146    | 0.007     | 0.000   | 0.022   | 6.37  |
|                             | Log Population Density  | -0.000   | 0.000     | 0.007   | 0.001   | 5.30  |
| Application: Average Derivative | Application Probability | 0.014    | 0.002     | 0.033   | 0.002   | 6.17  |
|                             | Win Probability         | 0.003    | 0.030     | -0.053  | 0.003   | 1.84  |
|                             | Log Request             | -0.010   | 0.043     | -0.008  | 0.007   | 0.04  |
|                             | Log Units               | -0.030   | 0.044     | 0.079   | 0.009   | 2.43  |
| Post-Loss Behavior          | Intercept               | 0.254    | 0.004     | 0.235   | 0.005   | 2.88  |
|                             | Win Probability         | 0.024    | 0.013     | 0.130   | 0.037   | 2.72  |
|                             | Poverty Rate            | -0.031   | 0.029     | 0.110   | 0.042   | 2.77  |
|                             | Log Population Density  | 0.030    | 0.003     | 0.023   | 0.004   | 1.70  |
|                             | Log Units               | -0.022   | 0.009     | 0.042   | 0.015   | 3.62  |
|                             | Log Credits Per Unit    | -0.048   | 0.009     | -0.040  | 0.013   | 0.48  |
|                             | Reapplicant             | -0.027   | 0.009     | -0.028  | 0.014   | 0.03  |
| Reapplication Choice        | Intercept               | 0.367    | 0.004     | 0.408   | 0.006   | 5.36  |
|                             | Win Probability         | 0.040    | 0.015     | 0.252   | 0.042   | 4.75  |
|                             | Poverty Rate            | -0.017   | 0.032     | 0.072   | 0.048   | 1.53  |
|                             | Log Population Density  | 0.013    | 0.003     | 0.032   | 0.004   | 3.65  |
|                             | Log Units               | -0.074   | 0.010     | -0.043  | 0.017   | 1.56  |
|                             | Log Credits Per Unit    | -0.020   | 0.010     | 0.032   | 0.015   | 0.68  |
|                             | Reapplicant             | 0.117    | 0.010     | -0.006  | 0.014   | 6.51  |
| Bid Choice                  | Win Probability         | 0.386    | 0.025     | 0.353   | 0.011   | 1.19  |
|                             | Log Average Rent        | 0.755    | 0.079     | 0.817   | 0.028   | 7.49  |
|                             | Win Prob. × Poverty     | -0.361   | 0.171     | -0.464  | 0.067   | 0.56  |
|                             | Log Avg. Rent × Poverty | -1.092   | 0.553     | 0.008   | 0.206   | 1.86  |
|                             | Win Prob. × Density     | -0.059   | 0.014     | -0.021  | 0.007   | 2.43  |
|                             | Log Avg. Rent × Density | -0.097   | 0.043     | -0.168  | 0.019   | 1.53  |
| Distributional Parameters   | Win Probability Mean    | 0.436    | 0.003     | 0.412   | 0.026   | 0.91  |
|                             | Win Probability SD      | 0.387    | 0.002     | 0.347   | 0.019   | 2.11  |
|                             | Log Units Mean          | 4.005    | 0.004     | 4.011   | 0.047   | 0.13  |
|                             | Log Units SD            | 0.511    | 0.002     | 0.617   | 0.033   | 3.16  |
|                             | Log Request Mean        | 12.280   | 0.004     | 12.282  | 0.039   | 0.04  |
|                             | Log Request SD          | 0.573    | 0.003     | 0.508   | 0.027   | 2.35  |
|                             | Win Probability AR1 Intercept | 0.274   | 0.007     | 0.273   | 0.007   | 0.14  |
|                             | Win Probability AR1 Slope | 0.393   | 0.018     | 0.394   | 0.018   | 0.07  |

Notes: This table reports estimates and standard errors for the data moments to be matched in the structural model, alongside their simulation analogs. The final column reports the absolute value of the $t$-statistic testing equality of each data and simulation moment. $^* = p < 0.10$, $^*^* = p < 0.05$, $^*^*^* = p < 0.01$. 


Table A5: Sensitivity Analysis

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outside Option</td>
<td>-0.264</td>
<td>0.013</td>
<td>0.009</td>
<td>-0.561</td>
<td>-0.169</td>
<td>0.045</td>
</tr>
<tr>
<td>Win Value</td>
<td>-0.162</td>
<td>-0.331</td>
<td>-0.045</td>
<td>0.081</td>
<td>0.150</td>
<td>0.074</td>
</tr>
<tr>
<td>Entry Cost</td>
<td>-0.006</td>
<td>-0.083</td>
<td>-0.007</td>
<td>0.095</td>
<td>0.026</td>
<td>-0.003</td>
</tr>
</tbody>
</table>

Panel A: Model Primitives

Panel B: Incidence

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household Share</td>
<td>-0.041</td>
<td>0.035</td>
<td>-0.018</td>
<td>-0.229</td>
<td>-0.032</td>
<td>-0.009</td>
</tr>
<tr>
<td>Developer Share</td>
<td>0.021</td>
<td>-0.003</td>
<td>0.017</td>
<td>0.122</td>
<td>-0.005</td>
<td>0.013</td>
</tr>
<tr>
<td>Entry Cost Share</td>
<td>0.019</td>
<td>-0.032</td>
<td>0.002</td>
<td>0.108</td>
<td>0.037</td>
<td>-0.003</td>
</tr>
</tbody>
</table>

Panel C: Impacts

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Displacement Rate</td>
<td>-0.016</td>
<td>0.007</td>
<td>0.002</td>
<td>-0.039</td>
<td>-0.001</td>
<td>-0.010</td>
</tr>
<tr>
<td>Marginal Cost Per Unit</td>
<td>-0.5</td>
<td>1.4</td>
<td>-3.0</td>
<td>-1.1</td>
<td>8.9</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Notes: This table reports the results of a sensitivity analysis following Andrews et al. (2017). I consider centered one-percentage-point perturbations to the regression coefficients used as empirical moments. I re-estimate the structural model with 0.5-p.p. upward and downward perturbations to each moment. This centers the sensitivity analysis on the actual estimates and circumvents non-smoothness concerns. Marginal costs per unit are in thousands of dollars. * = p < 0.10, ** = p < 0.05, *** = p < 0.01.
Table A6: More Estimates of Model Primitives and Potential-Applicant Characteristics

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Winners</td>
<td></td>
<td></td>
<td>Losers</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>P25</td>
<td>P75</td>
<td>Median</td>
<td>P25</td>
<td>P75</td>
</tr>
<tr>
<td>Ex-Ante Value</td>
<td>1.51</td>
<td>1.16</td>
<td>1.80</td>
<td>1.34</td>
<td>1.05</td>
<td>1.71</td>
</tr>
<tr>
<td>Outside Option</td>
<td>1.35</td>
<td>0.93</td>
<td>1.67</td>
<td>1.22</td>
<td>0.82</td>
<td>1.62</td>
</tr>
<tr>
<td>Win Value</td>
<td>1.72</td>
<td>1.43</td>
<td>2.21</td>
<td>1.23</td>
<td>0.64</td>
<td>1.62</td>
</tr>
<tr>
<td>Application Cost</td>
<td>0.17</td>
<td>0.10</td>
<td>0.25</td>
<td>0.25</td>
<td>0.03</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Panel A: Model Primitives as a Share of Basis

Panel B: Model Primitives in Thousands of Dollars Per Unit

Notes: This table reports estimates of the model primitives, the ex-ante value of the application value function \( V^A(s_{it}, z_{it}) \), the outside option \( \pi_0(s_{it}) \), the win value \( \pi_1(s_{it}) \), and the application cost \( \kappa(s_{it}) \).

Table A7: Selectivity Correction for New-Unit Rent Hedonic Estimation

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Median Rent,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Units</td>
<td>0.967</td>
<td>0.997</td>
<td>0.887</td>
<td>0.864</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.008)</td>
<td>(0.013)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>Selection Correction</td>
<td>0.076</td>
<td>0.069</td>
<td>0.055</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.009)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table examines the bias from sample selection in the estimation of willingness to pay for new rental units by Census tract. The specification estimated is Equation 18. Bootstrap standard errors account for first-stage estimation of the selectivity correction. * = \( p < 0.10 \), ** = \( p < 0.05 \), *** = \( p < 0.01 \).
Table A8: Application Data Coverage: By Property and State-Year

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Diff. (SE)</td>
<td>Yes</td>
<td>No</td>
<td>Diff. (SE)</td>
</tr>
<tr>
<td>Unit Count</td>
<td>60.85</td>
<td>66.45</td>
<td>-5.60*** (0.90)</td>
<td>65.67</td>
<td>61.10</td>
<td>4.56*** (0.93)</td>
</tr>
<tr>
<td>% Low-Income Units</td>
<td>89.7</td>
<td>89.9</td>
<td>-0.3 (0.6)</td>
<td>88.7</td>
<td>91.8</td>
<td>-3.1*** (0.6)</td>
</tr>
<tr>
<td>Monthly Rent</td>
<td>1,672.97</td>
<td>1,693.44</td>
<td>-20.48*** (8.70)</td>
<td>1,684.57</td>
<td>1,684.08</td>
<td>0.49 (9.08)</td>
</tr>
<tr>
<td>Rents Below Fed. Max.</td>
<td>75.8</td>
<td>54.9</td>
<td>20.9*** (0.9)</td>
<td>70.5</td>
<td>53.2</td>
<td>17.3*** (0.9)</td>
</tr>
<tr>
<td>% New Construction</td>
<td>68.9</td>
<td>60.4</td>
<td>8.5*** (0.8)</td>
<td>67.3</td>
<td>58.5</td>
<td>8.8*** (0.9)</td>
</tr>
<tr>
<td>Family</td>
<td>61.7</td>
<td>55.6</td>
<td>6.1*** (1.0)</td>
<td>58.9</td>
<td>56.8</td>
<td>2.1** (1.0)</td>
</tr>
<tr>
<td>Elderly</td>
<td>44.3</td>
<td>33.2</td>
<td>11.2*** (1.1)</td>
<td>38.9</td>
<td>35.8</td>
<td>3.1*** (1.1)</td>
</tr>
<tr>
<td>Other</td>
<td>15.1</td>
<td>18.0</td>
<td>-2.9*** (0.6)</td>
<td>17.3</td>
<td>15.8</td>
<td>1.6** (0.6)</td>
</tr>
<tr>
<td>PDV Tax Credits Per Unit</td>
<td>278,327</td>
<td>161,329</td>
<td>116,997*** (9,014)</td>
<td>239,807</td>
<td>165,749</td>
<td>74,059*** (9,241)</td>
</tr>
<tr>
<td>Nonprofit</td>
<td>27.6</td>
<td>33.5</td>
<td>-5.9*** (0.8)</td>
<td>29.2</td>
<td>34.3</td>
<td>-5.1*** (0.9)</td>
</tr>
</tbody>
</table>

Panel A: Property Characteristics

Panel B: Location Characteristics

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>Diff. (SE)</th>
<th>(4)</th>
<th>(5)</th>
<th>Diff. (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Income Per Capita</td>
<td>27,539</td>
<td>27,554</td>
<td>-15 (330)</td>
<td>27,366</td>
<td>27,886</td>
<td>-520 (344)</td>
</tr>
<tr>
<td>% Poor</td>
<td>77.1</td>
<td>75.3</td>
<td>1.8*** (0.3)</td>
<td>76.3</td>
<td>75.8</td>
<td>0.5 (0.4)</td>
</tr>
<tr>
<td>% Less than HS</td>
<td>15.2</td>
<td>15.7</td>
<td>-0.4* (0.2)</td>
<td>15.5</td>
<td>15.3</td>
<td>0.2 (0.3)</td>
</tr>
<tr>
<td>% HS Graduate</td>
<td>32.0</td>
<td>31.5</td>
<td>0.5** (0.3)</td>
<td>32.1</td>
<td>30.9</td>
<td>1.2*** (0.3)</td>
</tr>
<tr>
<td>% Some College</td>
<td>29.5</td>
<td>28.7</td>
<td>0.8*** (0.2)</td>
<td>29.3</td>
<td>28.6</td>
<td>0.6*** (0.2)</td>
</tr>
<tr>
<td>% College Graduate</td>
<td>14.8</td>
<td>14.9</td>
<td>-0.1 (0.2)</td>
<td>14.6</td>
<td>15.3</td>
<td>-0.7*** (0.2)</td>
</tr>
<tr>
<td>% More than College</td>
<td>8.5</td>
<td>9.3</td>
<td>-0.8*** (0.2)</td>
<td>8.5</td>
<td>9.8</td>
<td>-1.3*** (0.2)</td>
</tr>
<tr>
<td>% Non-Hispanic White</td>
<td>54.7</td>
<td>51.5</td>
<td>3.2*** (0.8)</td>
<td>52.6</td>
<td>53.5</td>
<td>-0.9 (0.8)</td>
</tr>
<tr>
<td>% Non-Hispanic Black</td>
<td>23.1</td>
<td>27.7</td>
<td>-4.6*** (0.7)</td>
<td>25.6</td>
<td>25.8</td>
<td>-0.2 (0.8)</td>
</tr>
<tr>
<td>% Hispanic</td>
<td>14.8</td>
<td>13.8</td>
<td>1.0** (0.5)</td>
<td>14.6</td>
<td>13.5</td>
<td>1.1** (0.5)</td>
</tr>
<tr>
<td>% Asian</td>
<td>2.4</td>
<td>2.4</td>
<td>-0.0 (0.1)</td>
<td>2.4</td>
<td>2.4</td>
<td>0.0 (0.1)</td>
</tr>
<tr>
<td>Pop. Density (per sq. mi.)</td>
<td>3,708</td>
<td>5,252</td>
<td>-1,544*** (193)</td>
<td>4,155</td>
<td>5,349</td>
<td>-1,194*** (202)</td>
</tr>
<tr>
<td>% Rentals Vacant</td>
<td>6.01</td>
<td>6.00</td>
<td>0.00 (0.15)</td>
<td>5.97</td>
<td>6.07</td>
<td>-0.11 (0.16)</td>
</tr>
</tbody>
</table>

Observations: 6,334 8,528 9,568 5,294
P-val of Balance Test: 0.000 0.000

Notes: This table uses HUD’s LIHTC property database to compare properties with matches in my application data to unmatched properties. I use HUD data from 2005 to 2019 for all U.S. states. Location characteristics are from the five-year ACS centered on 2019. Columns 1 and 2 report means of the matched and unmatched samples at the property level, and Column 3 reports the difference in means. Columns 4–6 proceed in parallel, but defining the match at the state-year level, so as to distinguish between match failures and sample coverage. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$. 
B Supplemental Information

B.1 Data Construction

This appendix section explains the construction of data used in the analysis.

Data Coverage. Appendix Figure A24 shows a map of applications in my data. The key populous states my data miss are Illinois (IL), Massachusetts (MA), and New York (NY). Appendix Figure A25 shows application counts and their total proposed units by the year of application and by their grant assignment. Nationwide, the actual number of funded LIHTC units has slightly declined over this period. Data coverage therefore improves considerably over time.

Appendix Table A8 examines the coverage of my application data by matching winning applications to HUD’s LIHTC property database. This table excludes winning applications that are not successfully matched into the HUD data. The HUD data suffers from significant flaws: Its own coverage is far from complete, and properties are often miscoded in terms of the type of credit they receive (4% or 9%), or are allocated funding in years other than the one in which they apply. With these caveats in mind, however, it still provides a useful way to examine the consequences of the incomplete state–year coverage of my application data. Columns 4–6 show that state–years covered in my application data do differ on observables from all LIHTC state–years from 2005 to 2019. They tend to have larger projects in unit count and in tax credits per unit, and they are more likely to set rents below the federal maximum. The differences on location characteristics are smaller, though my application data tends to overweight state–years with lower population density.

Appendix Figure A26 provides some histograms at the application-round level. In particular, its left panel shows that the win probability in the median round is about 40 percent, and there is substantial variation in win probabilities across rounds. Its right panel shows the distribution of a measure of a round’s ex-ante probability, the “explained share” \( \frac{\text{Var}(\hat{p}_t)}{\text{Var}(W_t)} \). When this explained share is one, then applicants face no uncertainty as to whether they would win or lose. When this explained share is zero, then the competition is equivalent to a uniform lottery. Intermediate cases are weighted lotteries. The histogram finds that applicants in most LIHTC rounds feature substantial ex-ante uncertainty about their grant assignment.

Standardizing Geography Definitions. This paper almost exclusively uses the 2000 definition of Census tracts, block groups, and blocks. The rationale for this choice is that 2010 and 2020

\[37\text{Data were unavailable in IL due to the state’s public-records law protecting the contact information of developers. The MA and NY housing finance agencies informed me they do not store some key variables for my analysis outside of the paper applications submitted by developers. Among smaller states, Nevada, Kansas, and Louisiana all cited record-keeping issues and were unable to entirely meet my records requests. A state law in Arkansas explicitly forbids their agency to release LIHTC applications. Missouri and Vermont do not use numerical rules to score applications. In the District of Columbia and South Dakota, the agencies refused my records requests, and my submissions to their state appeal boards were unsuccessful.}\]
geography definitions are potentially endogenous to the LIHTC, in that developments can require these geographies to be redrawn. The exceptions regarding 2000 Census geographies are the RDD, rent discount, and counterfactual tenant income analyses.


Another instance where I use geographic crosswalks is to reconcile HUD’s Difficult Development Area (DDA) boundaries to lower-level Census geographies. First, I digitized lists of DDAs for the application years 2002 through 2016. Over time, these DDAs have been defined at varying combinations of geographic levels: county, county subdivision / town / place, Zip Code Tabulation Area (ZCTA), and several metropolitan area concepts (MSA/PMSA/CBSA). Once digitized, I used probabilistic crosswalks from Geocorr to map DDAs to the tract level. Finally, I harmonize the tracts to the 2000 Census definition as above.

Geocoding Addresses. I used a combination of the Google Maps API, Geocodio, and OpenStreetMap to geocode street addresses in applications. Results were first checked for basic consistency with rules. For instance, the geocoded state should match the state in which the application is submitted. Similarly, the geocoded county, town, and zip code should generally match information submitted on the application. Inconsistencies were then hand-reviewed. Geocoding results were also carefully hand-reviewed whenever these services flagged the geocode as inexact.

Some applications provide the address information in formats that do not lend themselves to easy geocoding. This often occurs when a project is proposed in newer neighborhoods, or low-density areas, where exact street numbers have not been assigned to every parcel. For instance, the street-address field provided is sometimes an intersection, or with reference to an intersection (e.g. “A St, 500 ft S of A St and B Ave” or “SEC [southeast corner] of A St and B Ave”). Such addresses required extensive hand-review to assign geocodes.

A notable co-benefit of the data linkages to the NHPD and to CoreLogic was an opportunity to detect and resolve other geocoding errors. Inconsistencies were also hand-reviewed. Through these steps, I concluded my geocoding of LIHTC applications is likely to be highly accurate. Wilson et al. (2022) document that the HUD LIHTC database has significant geocoding errors, motivating

38The originals are available in PDFs on the HUD website: https://www.huduser.gov/portal/datasets/qct.html.
39Geocorr is maintained by the Missouri Census Data Center and is available at https://mcdc.missouri.edu/applications/geocorr.html.
40In surprisingly many cases, the application submission is incorrect, and then I keep the geocoding output version (e.g., the town is misspelled, the zip code is incompatible with other information).
my attention to this issue.

**Detecting Reapplicants.** To identify applications that are potentially reapplications, I first use rule-based record-linkage methods (dtalink in Stata) to link my data to *itself* (i.e., duplicating the database and treating the duplicate as if it were another file). After dropping cases where an observation links to itself, using a unique application identifier that I assigned at data entry, the data contain applications linked to their potential reapplications. Due to symmetry in the record-linkage rules, all pairs necessarily appear in duplicate (A links to B, so B also links to A), which is easily resolved by dropping one of each such pair.

Variables used in the record linkage included substrings of the project name and address, geography (city, county, geographic coordinates in degree-wise Manhattan distance), developer contact information, and project variables (unit count, funding request). Record-linkage programs often assign “scores” based on the extent of matching criteria. I used this score to divide potential links into extremely likely or unlikely reapplications and marginally-likely reapplications. I hand-reviewed all applications of the marginal group and auto-coded those at the extremes.

**Data Axle (Infogroup).** The Data Axle files are annual cross-sections from their “Consumer Historical” database. I recode several variables for my analysis. First, the variable `owner_renter_status` provides a scale of 0 to 9 for the likelihood the resident is a renter or owner-occupant. I split this scale at the midpoint in identifying likely renters and owner-occupants. Second, the land-use variable is `location_type`. I exclude any households coded as living in nursing homes, retirement homes, trailers, or undefined location types in the single-family versus multi-family analysis. Third, the variable `find_div_1000` is the household’s predicted income in thousands of current dollars. I map this to national household deciles in each year using Table A-4a of Semega and Kollar (2022).

**Competition Dates.** Analyses in Section 4 use quarterly data on the award dates of LIHTC competitions. I collected these by hand in several steps. First, I looked for related documents that had datestamps on them, such as award announcement press releases, the lists of winners, or schedules included in the QAP. Second, I used the metadata on spreadsheets or other public records shared with me, which was often not wiped in the public-records release process. These spreadsheets often contain a “last saved” or “last printed” date which is the award date or close to this date. Third, states typically follow the same annual or biannual schedule, so I interpolated missing years.

**National Housing Preservation Database (NHPD) Linkage.** I use the NHPD to establish whether, in the absence of a LIHTC award, a losing application builds subsidized housing via another subsidy. I use Stata’s `dtalink` to match the application data to the NHPD, using the same variables listed in “Detecting Reapplicants.” I also then hand-review marginal potential matches.
CoreLogic Linkage. I link both winning and losing applicants to the CoreLogic data. For losers, this is to establish whether any development occurs. For winners, this is both to confirm development and to obtain additional data (for use in potential future analyses).

The challenge in matching the CoreLogic data is its scale, requiring me to carefully control false positives, as most parcels are not LIHTC parcels. I begin the match process by identifying a subset of near-sure matches by (1) matching substrings of the LIHTC project name to the parcel’s first-listed LLC owner, (2) the exact street address, or (3) a very small distance between my and CoreLogic’s geocoded latitudes and longitudes. I hand-reviewed all such matches to confirm they were not false positives.

I then use a regression model in the CoreLogic to predict likely matches based on a larger set of variables. I do this iteratively in rounds of matching: After hand-reviewing the latest set of matches, I add them to a file of confirmed matches used in predicting other likely matches. Over many rounds of matching, I varied the set of variables, but they broadly fall into two classes. First are common variables across the two datasets, like those used to identify reapplicants or perform the NHPD match. Second are CoreLogic-only variables only measured in CoreLogic but that I find to be highly effective in avoiding false positives, such as lot size and land-use classification.

B.2 Running Variable Definition in QCT RDD

This appendix section explains the differences from Baum-Snow and Marion (2009) and Davis et al. (2019) in defining the running variable for the Qualified Census Tract (QCT) regression discontinuity design.

The basic criteria to be a QCT are (1) a tract poverty rate above 25 percent and (2) a low tract median household income. The definition of the second criterion is a ratio of tract median income to its metropolitan area’s median income, adjusted for household size. A tract may qualify if more than 50 percent of the tract earns less than 60 percent of the adjusted metro-area median. Baum-Snow and Marion (2009) use only the latter threshold, whereas Davis et al. (2019) uses the minimum of the distances to each threshold to collapse the two-dimensional discontinuity into a single running variable.

HUD’s implementation of these thresholds is more complex than this summary. Among the complexities are: (1) an adjustment to ensure that no more than 20 percent of a metropolitan area is classified as a QCT, (2) disqualifications for tracts with high sampling error in the ACS, (3) unusual methods of averaging over multiple American Community Survey years, and (4) a tiered ranking based on whether the tract meets one or both of the criteria. These aspects all otherwise reduce the first-stage coefficient in the QCT RDD.

From this complexity, a single priority ranking of tracts within each metro area nevertheless does arise. That is, each tract $i$ in each metro area $m$ has a well-defined rank $r^m_i$, with a threshold
rank \( \tilde{r}^m \) such that—with one exception of which I am aware—a tract’s QCT assignment is given by

\[ D_i = 1[\tilde{r}^m_i < \tilde{r}^m]. \]

I use \( r^m_i \) as my running variable, rescaled into a metro-area percentile rank.

The exception is that HUD classifies small tracts as QCTs that are beyond the metro-area threshold rank \( \tilde{r}^m \), so that it can use up any “spare” capacity below the threshold of 20 percent of metro-area population. It does this in rank order, explaining why my running variable still has some tracts on the wrong side of the threshold still being QCTs.

Table 2 shows a first-stage coefficient of 0.62—that is, being just to the right of the QCT threshold on the running variable raises the probability of being a QCT by 62 percentage points. Baum-Snow and Marion (2009) do not provide an equivalent estimate, and the first stage in Davis et al. (2019) is approximately 0.4. There is thus a considerable power gain to be extracted from precisely replicating HUD’s QCT assignment formula.

### B.3 Sampling Variation in Basis Boost

My analysis of the role of sampling variation in the basis boost is greatly facilitated by the HUD Qualified Census Tract data files.\(^{41}\) These files include the official Census estimates of the margin of error for each observation. HUD includes these values in the data files because the QCT rules exclude tracts from QCT eligibility if these margins of error are too large relative to the estimate.

Census reports margins of error in the American Community Survey (ACS), the source of the QCT variables since 2013, by multiplying estimated standard errors by exactly 1.645.\(^{42}\) For each observation \( i \), I took two random draws from the normal distribution \( N(\mu_i, \sigma_i) \), where \( \mu_i \) was the Census estimate for that observation and \( \sigma_i \) is its implied standard error.

Using the same programs as to measure the QCT running variable, I then calculate whether a tract would have been assigned as a QCT. I hold DDA status as given, and so tracts that are already in DDAs will be boosted irrespective of the random draws.

### B.4 Construction of Hypothetical Applications

The following states are included in the subsample for which I can analyze the trade-off between win probability and rental income: AZ, CA, CO, GA, IA, IN, NM, OH, OK, TX, UT, WA, WI. In these states, I have data on proposed rents in applications, and the QAP scoring rule features some bonus for lower rents.

After reviewing these states’ QAPs in each year, I calculated how many rent-related points each application received, given their current application. I then simulated alternative applications by

---

\(^{41}\)Available at [https://www.huduser.gov/portal/datasets/qct.html](https://www.huduser.gov/portal/datasets/qct.html).

recalculating the rent-related points and swapping out their actual score points for this alternative.

For applications at the federal maximum rent (60 percent of area median income), I did not simulate an “up” deviation to a higher rent, as this would disqualify their application. For applications at the score-maximizing kink, I simulated a “down” application with a modestly lower rent but which received no additional QAP points.

I follow Section 2.3 to map from scores to win probabilities. I only allow one of the $K$ “folds” of the applications to deviate to an alternative rent ("up" or "down") at a time. Alternative win probabilities are calculated when applications are in the deviation fold. Within the deviation fold, it is randomly assigned in each bootstrap whether an application will deviate up or down.

B.5 Robustness Checks

This appendix summarizes robustness checks for analyses presented in the main text of the paper.

**Section 4.** Appendix Figure A10 replicates the USPS tract-level event study in the Data Axle data, finding considerable attenuation, which explains why the block-group and tract impacts in Figure 7 are of roughly the same magnitude. Appendix Figure A11 shows that the results of Figure 7 change little when I omit the win-probability control. This implies winners are not highly selected on their outside option relative to losers. Appendix Figure A12 presents event studies estimated separately for each treatment-year cohort, ruling out the possibility my results are driven by improper comparisons between early-treated and late-treated cohorts.

**Section 5.** Appendix Figure A14 shows estimated boost effects on a broader set of outcomes, such as the probability of at least one application from a Census tract in that year, counts of winning and losing applications, and counts of funded units. Appendix Figure A15 shows the estimates are robust to county–year fixed effects, as well as to controlling for tract characteristics with time-varying coefficients. Appendix Figure A16 includes always-boosted tracts. Appendix Figure A17 separately estimates impacts for the two boost triggers (QCT and DDA).

B.6 Analysis of Self-Scores

A feature of the LIHTC is that, in some states, developer applicants are required or recommended to submit “self-scores,” wherein they fill out the QAP scoring rubric before their application is reviewed by the government. The self-score data allow me to explore developer’s elicited expectations, similar to the survey in Kapor et al. (2020). The main takeaway from the self-score data is that the typical developer essentially knows their application’s score when they apply.

I begin by comparing actual and self-scores in rank terms, ranking applications within their rounds. Self-scores are highly informative about actual scores (Spearman rank correlation coefficient of 0.73) and in fact understate actual scores on average. The left panel of Appendix Figure A27
displays a binned scatterplot of actual scores ranks versus self-scores. I transform both self-scores and actual scores into ranks in the distribution of actual scores. On average, an applicant whose self-score implies the top-rank score in their round has a 80th-percentile actual score, whereas a bottom self-scoring applicant has a 5th-percentile actual score.

I then re-estimate applicants’ win probabilities using their self-scores. Following Hendren (2013), my approach allows for developers to strategically report self-scores that are not their true beliefs. To summarize the approach, I estimate the conditional distribution of actual scores given self-scores, and then I compute the developer’s win probability by sampling from this conditional distribution. The key implementation steps are: (1) pool scores across rounds by transforming them into percentile ranks, and (2) use kernel-density methods to nonparametrically estimate the copula between actual score percentile and self-score percentile. This approach assumes that applicants know the distribution of their potential rivals but not the draw.

Similar to the results for self-scores, the new estimates of win probabilities show developers are highly informed. The $R^2$ of the win-probability measures constructed using actual scores and self-scores is 0.94. The right panel of Appendix Figure A27 shows the binned scatterplot of these two estimates of win probabilities. Perhaps surprisingly, there is substantial mass of the self-score-based win probabilities near zero and one. A lack of knowledge about one’s own score is thus unlikely to explain why I observe so many developer applications with little chance of winning. A remaining potential explanation is that applicants are less informed about their potential rivals than my simulations assume.

**B.7 Estimating the Marginal Willingness to Pay for New Rental Housing**

This section introduces an econometric strategy for estimating the marginal willingness to pay (MWTP) for new rental housing by location. It addresses a sample-selection problem as in Heckman (1979): Rents on new units are only observed when new units are built, and construction occurs when potential rents exceed construction costs.

Such sample selection introduces a bias in hedonic regressions of new-unit rents on location characteristics. Failing to account for sample selection will cause hedonic regressions to systematically overstate potential new-unit rents in neighborhoods where no new rental units were built.

Addressing such sample-selection concerns is of potentially substantial importance, as new construction of rental units occurs in only 17 percent of Census tracts from 2010 to 2019. By implication, the researcher must impute new-unit rents in 83 percent of tracts.

**Selection Model.** Following Heckman (1979), I derive a bivariate-normal sample-selection model. I then discuss the use of lagged construction activity as a proxy for local construction costs,
motivating its use as an instrument for the selection correction. Consider the following model of the MWTP for rental housing of age $a$ in location $i$:

$$R_i(a) = x_i \beta(a) + \delta(a) \xi_i + \varepsilon_i(a),$$

where $x_i$ contains observable location characteristics, $\xi_i$ is a time-invariant scalar unobservable for the location, and $\varepsilon_i(a)$ is an i.i.d. normal error term across locations and unit ages. The outcome $R_i(a)$ is the (log) rent for units of age $a$ in location $i$. The terms $\beta(a)$ and $\delta(a)$ are age-specific coefficients on location characteristics and the time-invariant unobservable respectively.

Suppose there are two ages of units, new $(a = 1)$ and old $(a = 0)$. I implicitly remove the time-invariant location unobservable term $\xi_i$ by controlling for old-unit rents in the same location.\(^{43}\) This yields a MWTP for new units of

$$R_i(1) = x_i \hat{\beta} + \delta R_i(0) + \bar{\xi}_i.$$ (16)

I now show the selection problem. The MWTP for new rental housing $R_i(1)$ is observed if and only if it exceeds construction cost:

$$B_i(1) = 1[R_i(1) \geq C_i(1)],$$

as otherwise there is no new rental development in location $i$. I model construction costs by age and location as:

$$C_i(a) = x_i \gamma + \rho(a) \eta_i + u_i(a),$$

where $\eta_i$ is similarly a time-invariant cost of the location and $u_i(a)$ is an i.i.d. normal error term across locations and unit ages. This can be similarly rewritten as

$$C_i(1) = x_i \bar{\gamma} + \rho C_i(0) + \bar{u}_i.$$ (17)

Combining Equations 16 and 17, I obtain that the gap between rents and costs is

$$R_i(1) - C_i(1) = x_i \theta_1 + \delta R_i(0) - \rho C_i(0) - \nu_i.$$ 

Using the normality of the error terms, the expected value of this gap, conditional on new construc-

---

\(^{43}\)This selection-correction approach will therefore only yield estimates of the MWTP in locations with some old rental housing, as such rents are otherwise unobserved. This covers the vast majority of locations of interest. I impute rents from nearby Census block groups for remaining missing observations.
tion occurring in the location, is

\[ E[v_i | B_i(1) = 1] = \frac{\phi(\tilde{\rho}C_i(0) + x_i\tilde{\theta} + \delta R_i(0))}{\Phi(\tilde{\rho}C_i(0) + x_i\tilde{\theta} + \delta R_i(0))} \]

where coefficients with tildes indicate normalization by the selection model’s scale coefficient.

This model predicts that, observed new-unit rents are positively selected on the error term when new rental units are observed. Such sample-selection bias will be strongest when a location has characteristics that predict the absence of construction. It is thus of particular relevance for estimating the MWTP for subsidized development, which is less constrained by ex-ante profitability.

Using bivariate normality, we obtain that

\[ E[R_i(1)|x_i, R_i(0), C_i(0)] = x_i\tilde{\beta} + \delta R_i(0) + \lambda \frac{\phi(\tilde{\rho}C_i(0) + x_i\tilde{\theta} + \delta R_i(0))}{\Phi(\tilde{\rho}C_i(0) + x_i\tilde{\theta} + \delta R_i(0))}, \]

where \( \lambda \) is a coefficient on the selection-correction term.

**Results.** Nonparametric identification of the parameters \((\beta, \delta, \lambda)\) requires a selection instrument. Supply-side variables are natural instruments in this context, and the model above motivates lagged construction costs \(C_i(0)\) for this purpose. In equilibrium, such costs should affect rents on new units only as a proxy for current construction costs. Formally, the instrument requires the exogeneity assumption that

\[ E[\tilde{\epsilon}_i(1)u_i(0)|x_i, R_i(0)] = 0, \]

implying that lagged construction costs are unrelated to current MWTP for new units conditional on location characteristics \(x_i\) and current rents \(R_i(0)\) on old units. This instrument will be relevant to the extent that there is a persistent component of construction cost, as in the model above. In particular, I use the log count of rental units built before 1989 in the locality.

I estimate Equation 18 by the Heckman (1979) two-step approach, bootstrapping over both steps so as to obtain standard errors of correct size. I use data at the Census tract level. For greater data availability, I use log median rents across all units in the tract, rather than exclusively old units as in \(R_i(0)\). I instrument for the selection-correction term with the following four variables: the shares of rental units built in the tract from 1990 to 1999 and from 2000 to 2009, along with binary indicators for whether any rental units in the tract were built in these intervals.

Appendix Table A7 reports the results. Column 1 shows that a tract’s new unit rents are closely predicted by the rents on older units in the same tract. Column 2 includes the selection correction, instrumented using lagged construction. Columns 3 and 4 augment this specification with tract characteristics and with state fixed effects. In particular, the characteristics are: log population
density, age shares of population (less than 18, 18–34, 35–64, 65 and over), racial/ethnic shares of population (non-Hispanic white, non-Hispanic black, Hispanic, Asian, other), and educational-attainment shares of population (less than high school, high-school graduate and some college, bachelor’s degree and up), log median household income, and the housing-unit vacancy rate.

Across specifications, the coefficient on selection-correction term is statistically significant, and the instruments are strong. The estimates imply an average selectivity bias of 6 to 10 percentage points in extrapolating from tracts with new rental units to tracts without them. The selectivity bias falls sharply when I include a powerful predictor of selection as a control: the log count of rental units built in the tract before 1989. Overall, the results suggest that sample-selection concerns are of considerable importance in estimating MWTP for new rental housing. Failing to account for sample selection would have otherwise led to overestimates of MWTP and therefore of tenant incidence.

**Regulated Rents.** To calculate LIHTC rent discounts, I also require regulated rents. Federal regulations specify maximum rents in terms of fractions of AMI: For instance, the federal maximum rent for LIHTC units is that they are “affordable” (i.e., no more than 30 percent of income) for a household making 60 percent of the area median. The maximum monthly rent is thus 1.5 percent (0.6 · 0.3/12 = 0.015) of AMI. I use HUD data on Area Median Incomes (AMI) in 2019, aligning with the central ACS year for market rents.

There is a specific AMI for each household size, whereas rents are set in terms of a unit’s number of bedrooms. HUD’s conversion rule is 1.5 persons per bedroom: A one-bedroom unit’s regulated rent is computed by averaging the AMI levels of a one- and two-person household, for instance, whereas a two-bedroom unit’s rent reflects the AMI level of a three-person household. These steps yield rent levels for units by numbers of bedrooms in each HUD geography. I project the data to the tract level using ACS data on the distribution of numbers of bedrooms in rental units by tract. I use this approach to calculate an average federal maximum rent in LIHTC units by tract that would be allowed at 100% AMI, and then I adjust using the actual limits in my application data, which are specified as fractions of AMI.

**B.8 Imputing Counterfactual Tenant Incomes**

I estimate the distribution of LIHTC tenant incomes using property-level data from HUD (Table 8 of the 2019 LIHTC Tenant Data release). These data contain property-level medians as well as household shares with annual incomes in the following ranges: $0 to $5,000; $5,000 to $10,000; $10,000 to $15,000; $15,000 to $20,000; $20,000 and up. The income shares are extensively censored for privacy.

---

44See [https://www.huduser.gov/portal/datasets/il.html#faq_2023](https://www.huduser.gov/portal/datasets/il.html#faq_2023).
I model the property-level income distributions as normal: \( y_{ib} \sim \mathcal{N}(\mu_b, \sigma_b) \). In particular, I use median incomes to estimate the two parameters:

\[
\begin{align*}
\mu_b &= \alpha + \beta \cdot \text{MedInc}_b, \\
\sigma_b &= \gamma + \delta \cdot \text{MedInc}_b.
\end{align*}
\]

I obtain estimates of the parameters by method of moments, in particular by minimizing the sum of square residuals between actual and predicted shares when reported.

For counterfactual tenants, I take the estimated market rents from above, and I find the household income distribution in ACS data that corresponds to that rent. In particular, I use ACS Table B25122 from the 2021 5-year ACS tabulation, which is centered on 2019. This table contains the joint distribution of household income and gross rent by Census tract. I am therefore able to leverage geography as well as the rent level to construct a counterfactual.

C Theoretical and Structural Appendix

C.1 Identification

In the model, nonparametric identification of developer primitives poses two fundamental issues. First, grant winners and losers are selected at two stages: self-selection into application, and selection by the grant administrator. This identification issue is one of Roy selection, as development choices are observed only for losing applicants, not also winners and non-applicants. The second issue is that the act of applying is both selection and “treatment.” For instance, sinking predevelopment costs when applying may, upon losing, make private development more attractive. Addressing selection is, consequently, insufficient to identify the primitives.

We have the following data for each application: the grant assignment \( W_i \), the win probability \( p_{it} \), characteristics \( x_{it} \), and the development choice \( B_{it} \) for losers. Non-applicants are not observed at the (potential) application level; we are limited to counts of applications by characteristics. Identification relies on two instruments: first for the win value \( \pi_1(s_{it}) \) to address self-selection, and second in the grant assignment \( W_{it} \) conditional on the win probability \( p_{it} \) to address selection by the administrator. I will also require two large-support assumptions: first for the win-value instrument, so that some developers are certain to apply; and second for some application characteristics, so that some developers are certain not to build privately if they apply and lose.

The first step of the identification argument is to establish that application-supply responses to changes in win value pin down the distribution of the net value of applying. To show this, I obtain
the application condition from Equations 2 and 3,

$$\Delta V^A(s_{it}, \varepsilon_{it}) = \Delta \Pi^A(s_{it}) + \beta \cdot \Delta \mathbb{E}[V^A(s_{it+1}, \varepsilon_{it+1}) \mid s_{it}] + \Delta \varepsilon^A_{it} \geq 0$$

where the differences above are $\Delta \mathbb{E}[V^A(s_{it+1}, \varepsilon_{it+1}) \mid s_{it}] = \mathbb{E}[V^A(s_{it+1}, \varepsilon_{it+1}) \mid A_{it} = 1, s_{it}] - \mathbb{E}[V^A(s_{it+1}, \varepsilon_{it+1}) \mid A_{it} = 0, s_{it}]$, $\Delta \Pi^A(s_{it}) = \Pi^A(1, s_{it}) - \Pi^A(0, s_{it})$, and $\Delta \varepsilon^A_{it} = \varepsilon^A_{it}(1) - \varepsilon^A_{it}(0)$.

I use this difference to derive the observable application probabilities. Let the average difference in value functions be $\Delta \hat{V}^A(s_{it}) = \mathbb{E}[\Delta V^A(s_{it}, \varepsilon_{it}) \mid s_{it}]$, integrating over the difference in shocks $\Delta \varepsilon^A_{it}$. Furthermore, let $\bar{s}_{it}$ denote the observable component of the developer’s state, integrating over the distribution of persistent unobservables $G(\eta_i)$. Assume for now that $G(\eta_i)$ is known. The application probability for a developer whose observable state is $\bar{s}_{it}$ is

$$\Pr(A_{it} \mid \bar{s}_{it}) = \int F^A(\Delta \hat{V}^A(s_{it})) dG(\eta_i), \tag{19}$$

where $F^A$ is the marginal distribution of the shocks $\Delta \varepsilon^A_{it}$. By shifting the net value of applying $\Delta \hat{V}^A(s_{it})$, the win-value instrument identifies $\Delta \Pi^A(s_{it})$ and traces out the distribution $F^A$, given knowledge of the persistent-unobservable distribution $G$.

Identification of $G$ now follows from reapplication behavior. In particular, the reapplication probability of a developer with an observable state $\bar{s}_{it+1}$ is

$$\Pr(A_{it+1} \mid \bar{s}_{it+1}, A_{it} = 1, W_{it} = 0) = \int F^A(\Delta \hat{V}^A(s_{it+1})) dG(\eta_i \mid A_{it} = 1, W_{it} = 0). \tag{20}$$

Reapplicants are selected on $\eta_i$. Conditional on $\eta_i$, reapplication behavior remains governed by $\Delta \Pi^A$ and $F^A$. Thus, Equations 19 and 20 jointly identify $\Delta \Pi^A$, $F^A$, and $G$. At this point, however, we have not identified $\Pi^A$ in levels, nor the primitives $\pi_1$ or $\kappa$.

The second step of the identification argument is to identify the private-development payoffs $\Pi^B$ and the marginal distribution $F^B$ from developer choices upon losing. Following Heckman (1990), I use “identification at infinity”, invoking the large support of the win-value instrument.\(^4\)

In a limit set of developers who are certain to apply, we have that

$$\lim_{\Pr(A_{it} \mid \bar{s}_{it}) \to 1} \Pr(B_{it} \mid \bar{s}_{it}, A_{it} = 1, W_{it} = 0) = \int F^B(\Delta \Pi^B(s_{it})) dG(\eta_i). \tag{21}$$

Then, conditional on the win probability $p_{it}$, the distribution of $\eta_i$ is independent of characteristics $x_{it}$.\(^4\) By consequence, $\Delta \Pi^B$ and $F^B$ are identified among applicants (that is, given $h_{it} = 1$).

\(^4\)Newey (2007) proves nonparametric identification of nonseparable sample-selection models under these conditions. The necessary condition to apply his result in my setting is additive separability of $\pi_1$, $\pi_0$, and $\kappa$ from $\eta_i$ and $\varepsilon_{it}$.

\(^4\)Outside of this limit set, self-selection induces a correlation between $\eta_i$ and $s_{it}$ among applicants.
Naturally, pre-application outside options are not identified from post-application building choices.

The third step is to identify the primitives $\pi_0$, $\pi_1$, and $\kappa$ from $\Delta \Pi^A$, $\Delta \Pi^B$, $F^A$, $F^B$ and $G$. The question is how to move from differences in choice-specific values to the levels of these values. The solution is that flow payoffs from not applying and not building are normalized to zero in Equations 3 and 5. Importantly, exogenous variation in win probabilities $p_{it}$ is necessary to distinguish win values and entry costs; otherwise, only a combined value $p_{it} \pi_1(s_{it}) - \kappa(s_{it})$ is identified.\footnote{Alternative paths to distinguishing these terms are to observe applications with win probabilities near zero, or to observe aggregate-development responses (including non-applicants) to win value.}

There are two final points on identification. First, the joint distribution of the shocks $\Delta \varepsilon_{it}^A$ and $\Delta \varepsilon_{it}^B$ is identified by the private development choices of marginal applicants. Intuitively, the shocks are positively correlated if marginal applicants are less likely to develop privately upon losing than observably-similar infra-marginal applicants. Second, identifying the pre-application outside option follows from the large-support assumption on an application characteristic. When a developer will never build without subsidy, the impact on the value function from the shift in the outside option, $\pi_0(s_{it}(h_{it} = 1)) - \pi_0(s_{it}(h_{it} = 0))$, is zero. Outside of this limit set, developers can benefit from the effect of applying on their post-application outside option.

C.2 Structural Model Details

Sampling. A challenge in estimation is that applying for LIHTC is rare. Each year, approximately three percent of tracts have at least one application. To raise precision without an excessive number of observations, either simulated or actual sdata, I over-sample tracts with applications and then re-weight the sample.

For the actual data, I use choice-based sampling as in Manski and Lerman (1977). In particular, I configure the sample so that I retain all tract–years with applications, and I randomly sample tract–years without applications so that such observations represent one-fourth of the sample. In estimating the model, I always match data moments weighted to reflect the choice-based sample.

The same problem also arises in the simulated data. Here I use importance sampling to raise the application probability in the unweighted simulated data. To motivate my approach, note that applications with a strongly positive value for the persistent unobservable $\eta_i$ are likelier to apply (Appendix Table A3). Instead of the standard normal, I therefore use $\eta_i \sim \mathcal{N}(2, 1)$ and construct simulation weights $\omega_{it} = \phi(\eta_i) / \phi(\eta_i - 2)$, where $\phi(\cdot)$ denotes the standard normal density. In practice, I find this roughly triples the unweighted application probability in the simulation, improving model precision at lower counts of simulated observations.

Drawing Potential Applicants. In the simulation, each tract draws a potential applicant each year. I first sample with replacement from the empirical distribution of tract–years, providing me with a
simulated joint distribution of tract characteristics.

For each simulated potential application, I then take i.i.d. lognormal draws for the number of units and the tax credit amount per unit. The parameters of the lognormal distribution are estimated. This procedure abstracts away from correlation between application characteristics and tract characteristics. In my context, however, a negligible share of variation in application characteristics is explained by tract characteristics. This simplification is likely inconsequential.

I also randomly sample the persistent unobservable $\eta_i$ and an indicator for whether the potential applicant is a reapplicant. I set the share of potential reapplicants to 0.22, consistent with the empirical share of applicants that are reapplicants. As I do not target this share as a moment, this choice merely controls the relative precision of the data moments for application and reapplication. I thus ignore the “initial conditions” problem of Heckman (1981) in drawing independent samples of potential-reapplicant status, the persistent unobservable, and application and tract observable characteristics.

**Counterfactual 1: No LIHTC.** I calculate the incidence of the LIHTC by computing a counterfactual in which the LIHTC does not exist. I do so by setting win probabilities to zero and by changing the structural parameters so that entry costs are arbitrarily large and win values are zero.

I find the new equilibrium rents and housing quantity as follows. From the model’s demand side, I impose that $\Delta \log \bar{r} = \frac{1}{\varepsilon} \Delta \log H$, where $\bar{r}$ is the average rent, inclusive of the LIHTC rent savings, and $H$ is the total housing stock. The log change $\Delta \log \bar{r}$ therefore accounts for both rent savings and general-equilibrium effects.

On the supply side, I increment the intercept coefficient on the outside option until equilibrium is reached. An important detail in this approach is that the market rent $r^m$ prevails in the outside option, both before and after the LIHTC is eliminated. I therefore net out the value of the rent savings from $\Delta \log \bar{r}$ in adjusting outside options. This adjustment results in the market rent $r^m$ rising less in logarithmic terms than the average rent $\bar{r}$.

**Counterfactual 2: Stylized Voucher.** I eliminate the LIHTC as above. I then introduce the stylized voucher as a subsidy wedge between the rents facing the supply and demand sides of the model. This is accomplished by incrementing pre-voucher rents until enough housing is produced that, at the post-voucher rent, the model’s demand side wants to consume as much housing as it did under the LIHTC. This approach also restores the same average rent $\bar{r}$ as under the LIHTC. In the budget-balanced voucher counterfactual, I increment pre-voucher rents as above but stop when the fiscal cost of the voucher equals the fiscal cost of the LIHTC.
References for Appendices


