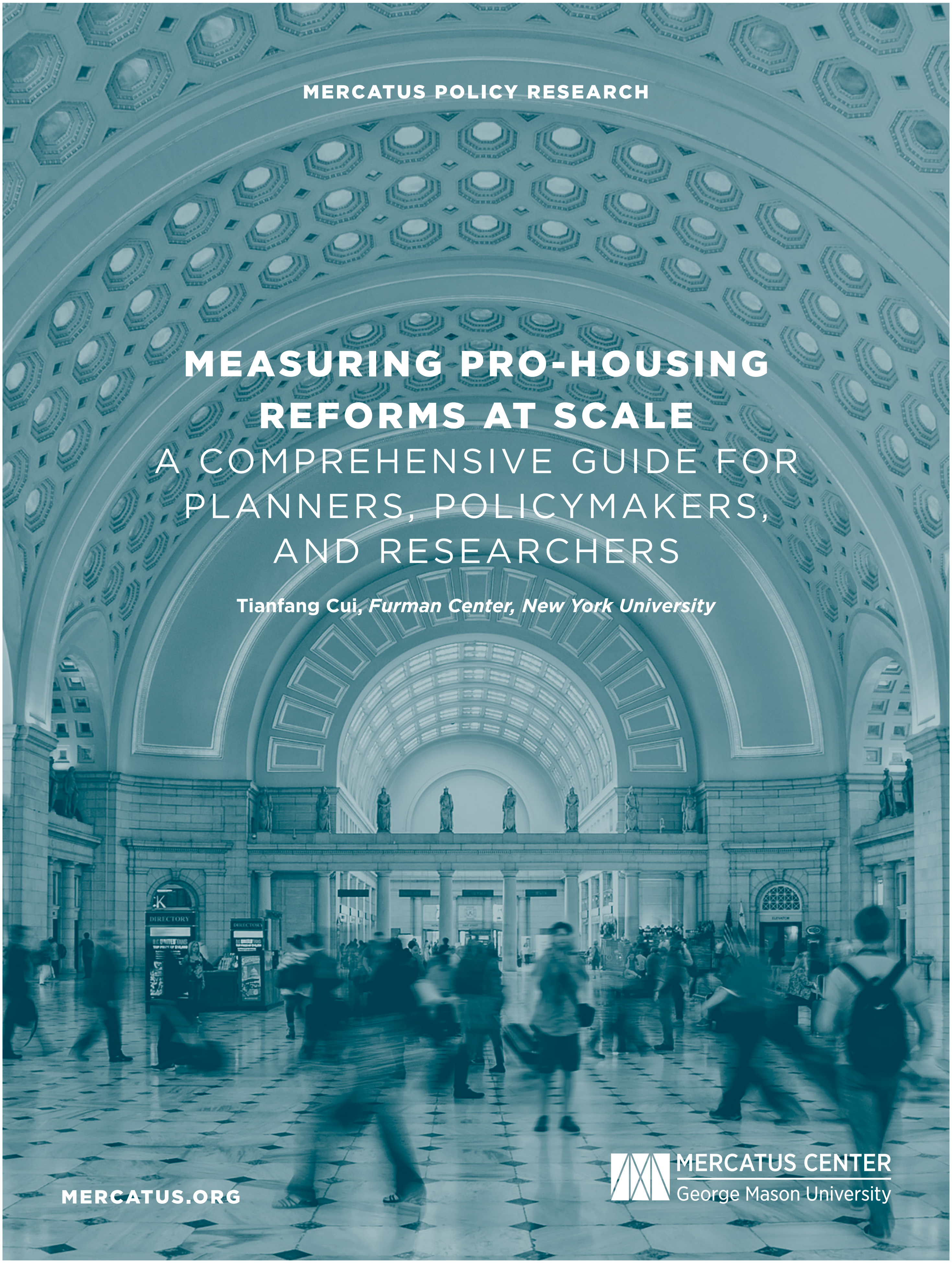


MERCATUS POLICY RESEARCH

# MEASURING PRO-HOUSING REFORMS AT SCALE

## A COMPREHENSIVE GUIDE FOR PLANNERS, POLICYMAKERS, AND RESEARCHERS

**Tianfang Cui**, ***Furman Center, New York University***

MERCATUS.ORG

MERCATUS CENTER
George Mason University

*Tianfang Cui. "Measuring Pro-housing Reforms at Scale: A Comprehensive Guide for Planners, Policymakers, and Researchers." Mercatus Policy Research, Mercatus Center at George Mason University, Arlington, VA, July 2024.*

## Abstract

To increase housing supply and housing affordability across the United States, state and local governments have started experimenting with a menu of pro-housing reforms. If politicians and planners adopt certain reforms, how can they best monitor their effectiveness at providing affordable housing? This paper outlines how those practitioners can use data for housing policy evaluation, from determining what to study to understanding the research methods used by outside experts. To illustrate recommendations, I use frontier empirical research in urban economics and recent data collection successes by government agencies. This paper offers a standardized reform evaluation plan that can speed up the production of knowledge about which reforms work in different market conditions.

*JEL* codes: R31, R52, R58

Keywords: land use, zoning, policy evaluation, housing supply

The agenda to reform housing development regulations is more popular now than it has been in half a century. In the 1970s, advocates opposed zoning restrictions that limit the types of homes produced in the United States. Over time, those laws restricted the number of residents who could afford to live outside the inner city. Since then, a body of research has confirmed that local regulations increase housing unaffordability (Gyourko and Molloy 2015; Freemark 2023).

With home prices and rents reaching new highs in the 2020s, it is critical to understand how policies can help increase new construction. Although the question is simple, it is difficult to draw clear lessons from many proposed reforms. In practice, policymakers are also balancing competing policy objectives. For example, an agenda to "increase affordable housing" might be about an increased supply of market-rate housing or about incentives to develop subsidized housing with below-market-rate rent. The debate about what pro-housing reforms should accomplish detracts from the search for policy changes that will have sizable effects.

This paper is for those who work in urban planning or local administration and want to use data to verify that reforms are working. It starts from the premise that policymakers should know *how* reforms are shifting the housing market before deciding *which* objectives the reforms should achieve. It then provides a road map to help practitioners produce a high-quality evaluation of reforms. Convincing results from a pilot phase can also empower decision makers to make policy changes permanent.

First, this paper explains what practitioners should know from existing studies and which questions are not settled. Section 1 argues that researchers have a consensus on many positive consequences of housing supply. The largest gap in knowledge is whether policy reforms alone can increase supply of more affordable housing. Recently passed reforms and permit data can help answer many questions on this pivotal front. Planners can help structure these data and measure the degree of policy change.

Section 2 begins to sketch the steps to develop a reform evaluation plan. Although the plan should be directed by a core group of practitioners, its success relies on many participants contributing domain knowledge. This section details a list of common reforms recently adopted in North America so that practitioners can reference which reforms are targeted to cultivate the types of housing units they want to build. Two case studies about permit data cleaning illustrate some benefits and costs that practitioners must balance with data collection.

The paper concludes with statistical techniques that are familiar to scholars and are used in the cited literature to estimate causal impacts of policies. If local experiments with reform evaluations are assessed with these techniques, they can form a body of knowledge of which reforms work and in which market contexts. Even if readers have no policy reforms they can evaluate, the final sections herein can help them critically evaluate well-publicized findings.

The focus in this paper is on pro-housing reforms with minimal costs. The examples are drawn from zoning changes, not property tax abatements and other subsidies given to developers. An evaluation of a subsidy or tax on housing outcomes can be only one part of a cost-benefit analysis: those analyses also require a model of the policy's opportunity cost. Understanding the steps to producing a high-quality policy evaluation should help practitioners become more capable of understanding which cost-benefit analyses are of the highest quality—that is, the ones based on the most credible predictions of policy impacts.

## 1. How Would Local Evaluations Improve on Existing Research

It can be difficult to understand the research on pro-housing reforms. In April 2023, there were conflicting reports on whether zoning reforms produce more affordable housing. Brey (2023) says, "It's true that zoning reform helps create more housing units, but there's no evidence it makes housing cheaper." Horowitz and Canavan (2023) say, "New zoning rules to allow more housing have helped curtail rent growth, saving tenants thousands of dollars annually." Whom should policymakers believe?

Table 1 summarizes the most relevant results from both studies, as well as providing a brief description of the analysis methods. Looking beyond the headlines, one can see that, while both studies share some outcomes, such as rent growth, they evaluate different reforms. The results of another study, by Stacy et al. (2023), are preliminary, especially because the effects they identify are not driven by larger-scale reforms. Conversely, the rent and supply effects in Horowitz and Canavan (2023) are so significant that shifting demand for urban living

**TABLE 1.** Comparing recent analysis on zoning reform effects

| Research citation | Stacy et al. (2023), as reported by Brey (2023) | Horowitz and Canavan (2023) |
|---|---|---|
| Method of analysis | Study uses difference-in-differences research design. Authors scoured newspapers to identify which cities were "treated" with zoning reforms. | Instead of looking nationwide, this study gathered rent changes from four jurisdictions that comprehensively revised their zoning ordinances. |
| Main finding | Compared to average "control" cities, where no zoning changes passed, an average reform city had 0.8 percent more occupied units and $60 lower monthly rent. | While rent growth nationwide grew by 31 percent over 2017–23, the four jurisdictions saw housing unit growth and cumulative rent growth no higher than 7 percent. |
| Concern 1 | Effect on lower monthly rents is statistically insignificant. If zoning reform has no effect on rents, the $60 estimate could still come out of pure chance because of factors unrelated to reform. | Were zoning changes adopted as the cities grew less desirable? If so, low rent growth is hiding rent growth driven by demand elsewhere in the metro area. |
| Concern 2 | Zoning reforms the authors observe were mostly not the citywide reforms debated today. Examples of reforms studied include reduced lot size requirements for new housing units and legalization of accessory dwelling units. | Are market conditions in case study cities—primarily cities such as Portland, Oregon, or high-income jurisdictions such as Tysons, Virginia—comparable to those in other parts of the country? |

after 2020 could be at play. It is unclear whether trend breaks in rent growth are due to policy changes or to the nature of the cities that passed them. Without more clarity on the reforms passed and market conditions, neither study is applicable to current debates.

By efficiently conducting their own evaluation plan, practitioners interested in pro-housing reforms can sidestep any confusion with interpreting these two studies. The body of this paper provides a guide to that end. It is also true that, as Freemark (2023, 4) concludes, between all the ways market rent could respond to a pro-housing reform, "[t]here is inadequate research to fully identify the degree to which any effect outweighs another." A more manageable evaluation, as this section argues, will be clear about how we are measuring improved housing supply and what reforms were passed to get there.

## 1.1 Evaluating reforms more precisely: Studying effects on housing supply

Rent is one kind of price signal that shows that developers in the market are responding to reforms, but it is not the only signal. Two recent studies use local data on new construction permits to show that developers respond to reforms by redesigning their projects in more cost-effective ways:

- Before and after Minneapolis eliminated parking requirements for multifamily buildings, the city published multifamily permits online that listed how many parking lots developers had proposed. Schieferdecker (2021) coded this information into a spreadsheet and showed that Minneapolis's elimination of parking-space minimums was followed by more buildings offering less than one parking spot per unit. Similar data should exist for other cities that removed parking requirements.

- Beyond the United States, the Indian city of Mumbai rolled out multiple "spot rezonings" on specific parcels in its urban core. Nagpal and Gandhi (2024) digitized permits issued by Mumbai officials that asked how much of a building's floorspace was used as a common area or for residential amenities. With these data, they showed that developers build taller buildings on upzoned parcels and that unit housing values decrease. Although the units in such buildings are smaller, tenants have access to more common space.

When it comes to measuring which cities restrain new construction the most, existing studies surveyed local officials on how many months it takes to process permits in their area. These data are only snapshots of the time that the survey rolled out (Gyourko, Hartley, and Krimmel 2021). More in-depth analysis exists for cities known for delays in issuing permits, such as San Francisco, where analysis of individual permit applications from 2009 through 2022 shows the average time to issue a building permit was two years at the sample's start, growing to three years.[1]

The three studies mentioned in this subsection show clear trends because they measure different outcomes—that is, whether it is profitable to build new kinds of housing typologies with standards different from existing built forms once reforms are passed. Even before units are completed, a change in design for permitted units is the first signal that reforms have changed developers' calculations. Localities can then track, following their usual permitting process, how long it takes for the permitted units to get completed.

Judging reform-based impacts on housing supply offers meaningful conclusions about housing affordability. A wider body of research tends to support the position that if new apartments are built, lower-income areas nearby will see a modest fall in rent. Table 2 shows that the magnitude of the fall varies from 1.2 percent to 7 percent for an average-market-rate building, depending on the rent data source and the time frame.[2] Full effects, though, take multiple years to appear.

1. Author's calculations using Goggin (2018), Gardiner and Neilson (2022), and California Department of Housing and Community Development (2023).
2. Rent data sources matter because ideal rent data—the contract rents offered in each unit in an area—have been nearly impossible to acquire at scale. Listed rent data on websites likely overrepresent new construction and underrepresent older buildings whose landlords cut rent as supply grows. Census median rents could better represent older buildings but cannot be produced in real time as rent contracts get renewed.

**TABLE 2.** Research on differential impacts of apartment supply on rent

<table>
<tr><th>Location of supply</th><th>Sample reform</th><th>Effect on local rent</th><th>Effect on metro area rent</th></tr>
<tr><td>High-income neighborhoods</td><td>Eliminating single-family zoning<br>Statewide upzoning</td><td>Studies show a decrease or are inconclusive.</td><td rowspan="2">Studies are inconclusive if supply is limited to a few neighborhoods.<br>Small supply growth can spur moving chains (high-income renters move out of lower-quality units), thereby lowering rent pressure.</td></tr>
<tr><td>Low-income neighborhoods</td><td>Higher density near commercial areas<br>Tax abatements for neighborhood redevelopment</td><td>Rent decreases according to most (but not all) of the literature.<br>Average building effects are a 1.2 to 7 percent rent decline over the medium term.</td></tr>
</table>

*Source*: Been, Ellen, and O'Regan 2023; calculations based on cited papers such as Asquith, Mast, and Reed 2023.

*Note*: The last column is shaded blue to indicate that the location of supply—high-income or low-income neighborhoods—does not influence the conclusions about the effect of supply on metro area rent. There is no systematic evidence that the amount of decrease in supply-driven rent on a metro area depends on the type of neighborhood.

As Been, Ellen, and O'Regan (2023) also conclude, the literature shows that if housing supply grows, housing somewhere in the city will become more affordable. Officials tracking how meaningful their reforms are would do well to strengthen the more uncertain part of the argument—specifically, how much policy can boost supply.

Existing data sources for American housing do not track design or housing type changes well.[3] The studies cited so far in this section used less detailed survey counts from the Census Bureau or collected unstructured local permitting documents and then converted the records into a structured file. Although detailed local data can help inform a city's reform agenda and the broad effectiveness of some reforms, the data cleaning process can pose delays to further analysis. Section 2 explains how to address this obstacle in the evaluation plan.

## 1.2 Evaluations with broad implications: Causal inference versus descriptive evidence

At a high level, pro-housing reforms are about "unlocking" development by giving developers more flexibility. One way to think of developers is as entrepreneurs. They take risks to assemble teams that build more housing to meet current market demand. In markets where land is scarce, reforms impose extra costs on developers if they limit how much floorspace it makes sense to build on a parcel. In markets with unaffordable homes but where land is abundant,

3. The Census Bureau primarily follows a classification from the 1940s, which separates most housing into single-family detached and multifamily homes of five or more units.

the most effective reforms may be those that legalize smaller and innovative home designs.

One way to think about a pro-housing reform is that it is a "treatment" for the parcels of land making up a city. The parcel is treated similarly to a person treated with a prescription drug. The drug causes a reduction in the outcome it claims to prevent, and it also has both beneficial and detrimental side effects. The problem of how pro-housing reforms affect supply, the subject of this paper, can be broken down into two research questions:

1. Did the recently passed reforms cause more housing construction? Units of housing supply are the policy evaluation's primary outcome.
2. Did the reforms, which included many policies designed to unlock certain housing typologies, cause developers to propose more units in those types? Measurements on the built form of permitted homes are the policy evaluation's secondary outcomes, along with other housing market outcomes such as lower rents.

A developer considers more than reforms when thinking about a project, including local housing demand, the financial macro environment, and other factors. While more housing could have been built after a reform was passed, it would be hasty to say the reform "caused" improved housing supply. What if the reform had no effect, but it was passed while local market conditions were already encouraging supply?

If that is true, the reform appears to matter only because of a coincidence in timing; its effects would not scale to other cities. Teasing apart these competing explanations using the data is what academics call the *causal inference problem*.

When the Food and Drug Administration demands a clinical trial for a medication, it wants evidence on reduced disease risks in a way that cannot be explained by differences in who took it and who did not. This is why clinical trials are randomized: if the "treated group" of medication takers did not receive any treatment at all, they would on average have similar disease risks to those who never received the medication. In other words, those who took the medication are compared to a control group, who are the counterfactuals for how those who took the medication would respond if they were never given the treatment.

More often than not, planners cannot randomize which parcels receive a zoning reform. Instead, they can list the confounders that drive the housing supply and that affect which reforms get passed in a specific place. Section 4 will discuss how to use local knowledge about these confounders in statistical methods to solve the causal inference problem.

Statistical methods work only when the confounders considered are data that can be measured and collected. It is reasonable to think that, despite best efforts, the extra steps involved for causal inference are not worth the cost, yet there is still value in looking at post-reform trends in the data by conducting a descriptive, not causal, analysis.

The Horowitz and Canavan (2023) report makes more sense as a descriptive analysis. In addition to the claim that pro-housing reforms can lower rents by a precise amount, the report also highlights jurisdictions that have implemented reforms that may have had significant impacts. Two of the four jurisdictions in that study were profiled more deeply in the Eastland (2023) and Hamilton (2023a) studies—New Rochelle, New York, and Tysons, Virginia, respectively. These studies described the actual reforms each made, offering more detail about specific reforms that could be used by other jurisdictions.

Another example of good descriptive work is by Brooks and Schuetz (2023), who tabulated census housing unit data across space to identify how concentrated new development was in several "high-growth neighborhoods." The descriptive trends cannot be explained by pro-housing reforms targeting those neighborhoods but are better explained by the neighborhoods' existing built environment and demographics. A similar neighborhood analysis before a zoning reform can inform policymakers about which areas to target in a pilot phase, along with providing a plan to control for known confounding factors when evaluating the policy.

## 2. A Road Map to an Evaluation Plan

### 2.1. Standardizing the workflow

Tracking policy reforms in real time does not have to be a costly effort, though it is most likely not an in-house one. Before starting to plan an evaluation, an appointed task force may have already agreed on which housing types are missing and which housing targets should be achieved. Ideally, the task force should follow the Montana model described in Hamilton (2023b); that task force included a broad range of interest groups and was transparent about why its members recommended certain policies over others.

Table 3 shows the steps involved in planning a reform evaluation once targets are set.[4] First, practitioners can decide whether they plan to adopt one of

4. One question beyond the paper's scope regards the details of implementation: If you're a state-level official trying to make local actors commit to a zoning reform, when should you offer "sticks" for non-

**TABLE 3.** Step-by-step process for pro-housing reform evaluation

| Workflow step | Who is involved? | How to proceed |
|---|---|---|
| *Setting up reforms* | | |
| Decide which housing types are undersupplied. | "Big tent" of pro-housing stakeholders | Read Hamilton (2023b) on the Montana Task Force. |
| Institute all reforms that could unlock those types. | Elected officials, local planners | Read subsection 2.2 on common reforms. |
| Summarize degree of zoning change following passage. | Local planners | Read subsection 2.3 on coding reform packages. |
| *Collecting and cleaning data* | | |
| Determine primary and secondary housing outcomes. | Planners, stakeholders | Read table 5, which breaks down outcomes from permit data. |
| Delegate those responsible for data collection. | Officials across several levels of government | Consider case studies in subsections 3.1 and 3.2. |
| Determine how to geocode permits. | Local planners | Use tools described in appendix A. |
| Determine how data will be accessed. | Planners, possibly local data team | Consider Seattle's example in subsection 3.1. |
| *Designing empirical analysis* | | |
| Descriptive or causal analysis? | Planners, any external researchers | Brainstorm confounders following subsection 4.2; determine whether causal analysis is worth the cost. |
| Can we exploit changes in reforms across space? | External researchers, using planners' local knowledge | Understand the assumptions in subsection 4.1, then determine whether they apply for the area being studied. |
| How much past data on outcomes do we have? | Local officials, planners | Explore the research designs in subsection 4.3; then check whether data are available. |
| Do we communicate our research design before the analysis? | External researchers | Describe design through a pre-analysis plan, as in subsection 4.4. |
| *Running the analysis* | | |
| Produce results following agreed-upon research design. | External researchers | Follow guidelines in section 4 for displaying results. |
| Compare results with other areas. | Planners, external researchers | Prepare data for meta-analysis, as detailed in subsection 4.5. |

several common reforms adopted elsewhere. Subsection 2.2 offers a list of those reforms and provides a standard table that can be used to document the intensity of reforms.

Section 3 uses two case studies to illustrate how data can be collected and cleaned in a way that can immediately be analyzed by external researchers. While it may be tempting to request data collection on many variables, any data task imposes some cost on personnel. It's important to note that some variables

compliance versus "carrots" for compliance? State legislatures are actively experimenting with this question, as described in Manji et al. (2023).

are more important for evaluation than others and that practitioners should prioritize according to their specific needs.

With the data in place, practitioners can agree with external researchers on a research design: a statistical model that employs detailed spatial data to estimate the success of policy effects beyond a single city. Even if practitioners do not conduct their own analysis, section 4 explains what those designs are, describes the pitfalls to avoid, and discusses when it is useful to collect more data to increase the precision of evaluation results.

## 2.2. Common reforms and their expected effects

To better distinguish the many reform opportunities cited in the literature, this paper groups reforms into six categories. These reforms will boost housing production if the marginal benefit of relaxing restrictions exceeds two marginal costs to developers: the construction cost and the cost of complying with any strings attached on where and how the reforms must be used.

1. ***Reform use restrictions.*** The history of American urban planning is tied with dividing land in the city on the basis of use zones, where only certain structures can be built freely. Almost unique among developed economies, the United States has zoning ordinances that create large sections of land exclusively for single-family detached homes. Simplifying use restrictions is then the first, but not the final, step to authorizing and providing housing options with more affordable units.

   **Recent reform 1:** California is one of the first states to pass laws giving ministerial approval to specific types of land use, effectively banning certain exclusionary uses. Assembly Bill 2299 in 2016 legalized accessory dwelling units below a certain size, and Senate Bill 9 in 2021 legalized converting all single-family lots to four-unit multiplexes.

   **Recent reform 2:** Since the 1990s, US cities have adopted form-based codes that take the focus away from use restrictions. New housing or establishments in residential neighborhoods can be built if the structure's dimensions are close to what is already there.

   **Could unlock:** Accessory dwelling units—that is, "granny flats" built to take up yard space—and varieties of manufactured housing.

2. ***Reform density restrictions.*** As a city grows, urban economists predict whether density will grow across all neighborhoods or whether regulations will limit density growth closer to the urban core, which induces

"horizontal" sprawl. American cities in the postwar years reinforced the latter pattern in neighborhoods that were built with low densities, using bulk regulations such as minimum lot sizes or minimum building sizes. Some US downtowns, such as San Jose, California, still cap density with height limits applied districtwide.

**Recent reform 1:** Houston, Texas, despite never adopting use restrictions, regulates built form through other ordinances. In 1999, minimum lot sizes that regulate single-family units were reduced from 5,000 to 1,400 square feet in many neighborhoods.

**Recent reform 2:** Large cities often experiment with increasing the floor area ratio (FAR), but these rezonings happen only in certain neighborhoods or with certain parcels. Recent proposals promote transit-oriented development (TOD) by allowing large FAR increases within so many feet of an existing transit station.[5]

**Could unlock:** Townhomes, as in Houston where they doubled units per single-family lot, or multifamily apartments, as discussed in Kulka, Sood, and Chiumenti (2023).

3. ***Remove parking-space minimums.*** In the car-dependent United States, it is easy to think that every new household needs an additional parking space. From the 1960s onward, cities have used parking-space minimums set by transportation engineers in the Institute of Transportation Engineers manual. The standards mandate a parking space for each apartment unit or for so many square feet of built area.

   The costs of parking-space minimums are not only in how much land is not used for housing on a parcel but also in how such minimums complicate building design—making certain multifamily housing types unfeasible. The inefficient use of land means more costs passed on to tenants as higher rent.

   **Recent reform:** Cities from Hartford, Connecticut, to Austin, Texas, have eliminated parking-space minimums citywide. Cities should set up apps that list parking rates and that process payments. A virtual platform makes it easy to price dynamically, with higher rates during peak traffic hours (Jordan 2019).

   **Could unlock:** Multifamily apartments.

5. Been, Jonlin, and Kazis (2023) compared existing TOD models in three US states.

4. ***Bypass discretionary review.*** Until the 1960s, US housing development was largely "by right": following the zoning ordinance's requirements was enough to get approval for development. Since then, officials in central cities and certain suburbs have exercised multiple steps of discretionary review before permits are approved.

   In cities such as New York and Los Angeles, this process could involve review by city planners on top of rezoning review by community boards or historic district commissions. In California, proposed apartment developments undergoing review can face years of delays because individuals can litigate how the developer should prepare an environmental review in court (Gray 2021). Delays translate into financial costs that restrict developers and limit housing supply.

   **Recent reform 1:** If allowing governmental approval for certain types of housing lacks political support, streamlining the development process can focus on specific bottlenecks. One example introduced in Montana and Washington State allows development to bypass state environmental review if it is already compliant with other planning laws (Kahn and Furth 2023).

   **Recent reform 2:** Instead of abolishing discretionary review, planners could fast-track the process with "pattern book zoning": multiplexes and small apartments can be built as long as they follow preapproved designs (Justus 2023).

   **Could unlock:** A variety of housing types, but particularly multiplexes and apartments that are targeted by local officials through discretionary review.

5. ***Restore past multifamily standards.*** A type of older housing in US history is the single-room occupancy (SRO) building—that is, hotels that rent out individual rooms to low-income populations. As those buildings depreciate and are replaced with new buildings, regulations can restrict new construction to follow higher standards but offer the same kind of micro-units.

   A closely related local regulation is occupancy limits for each housing unit, which bans homes rented out to more than a specified number of unrelated people. Reforms restoring multifamily standards are, in other words, rolling back restrictions that limit renter choice.

   **Recent reform 1:** Some cities, such as New York and Minneapolis, passed bans on new SRO permits entirely in the postwar decades. As part of its zoning reforms, Minneapolis repealed the ban and rewrote policy about where new SRO construction can take place.

**Recent reform 2:** Colorado Governor Jared Polis introduced a statewide pro-housing reform package in 2023 that would have overruled all occupancy limits by Colorado cities. Although the package failed, Boulder reversed a 1962 ordinance and authorized occupancy of up to five unrelated persons per unit.

**Could unlock:** Multiplexes and multifamily apartments designed differently from recent "five-over-ones," as tall as six floors and occupying large parcels.

6. ***Review building code standards.*** Architect Michael Eliason (2021) has argued for reforming "single stair" standards for apartments. He first notes that fire safety codes in the United States and Canada require two staircases for buildings taller than three stories. A floor plan that makes both staircases accessible to residents requires units that wrap around corridors. These floor plans are boxy and unable to fit into the narrow lots used to densify the urban core.
   The two-staircase requirement is one example of a wider issue: building codes in North America (i.e., the International Building Code) are implemented to reduce fire and disaster risks to real estate. But a tradeoff exists between minimization of risk and necessary design flexibility, such as constructing a new apartment building on a small parcel of urban land or building small homes with the least expensive materials.

   **Recent reform 1:** Advocates propose that the United States increase the minimum floor requirement for two staircases to eight floors, as is the case in New Zealand (Speckert n.d.). A compromise would be adopting Seattle or New York City standards that set the minimum at six floors.

   **Recent reform 2:** The state of Arkansas in 2019 passed Senate Bill 170, which banned local restrictions on vinyl siding or architectural style requirements. Housing that follows national standards for insurance eligibility are allowed "by right."

   **Could unlock:** Multifamily apartments designed differently from popular "five-over-ones," as argued in Smith (2023). Lower material costs could also make townhomes more feasible.

7. ***Strings attached.*** Two common conditions on housing reforms are location requirements and affordable housing requirements; the latter is

particularly common in the United States. First, more effective reforms tend to be allowed only where they have a planning use, such as dense TOD being used to increase the use and viability of rapid transit lines.

Second, numerous cities package housing reforms as part of a bonus in an "inclusionary housing" scheme. A new multifamily building has looser restrictions only if a specified fraction of its units have below-market rents. Under these arrangements, the marginal costs of these rent discounts could exceed the marginal benefits of reforms. When these programs are mandatory, buildings allowed using pre-reform standards become even less profitable to build.

## 2.3. Coding packages of common reforms

The political process results in many reforms across categories being passed at once, while some are passed with narrow scope. However, a bill that does not address *all* constraints that limit flexibility to build new housing types will lead to particularly uneven growth between types of housing. These points are apparent when comparing two citywide zoning reforms: the "upzoning" of Minneapolis, Minnesota, and Auckland, New Zealand.

The Minneapolis upzoning reform made headlines for its citywide use restriction changes, including the elimination of single-family zoning. However, table 4 shows that existing restrictions in three other categories were unchanged for those single-family neighborhoods. Minneapolis's story is therefore equally about divergence between single-family neighborhoods and areas next to transit or the urban core. Transit-rich areas saw a clear loosening of density restrictions, elimination of parking-space minimums, and restoration of multifamily housing standards. Policies in these categories were expected to unlock multifamily development, as can now be seen in citywide statistics (Maltman 2023).

The Auckland upzoning was a product of years of planning by the Auckland Council, which took the place of seven district authorities. The final plan encompasses everything within the city's borders, and it relaxed regulations on housing across nearly every reform category. Appendix table A1 highlights in green the categories in which reforms were put in place, both for apartment zones near existing transit and the three medium-density zones. Appendix table A2 lists a series of papers that found that the reform itself, not market demand, contributed meaningfully to Auckland's housing production and stabilized rents. The papers use standard statistical methods explained further in section 4.

**TABLE 4.** Reforms implemented during Minneapolis upzoning

| Policy | Reforms implemented | Pre-reform | Post-reform |
|---|---|---|---|
| Reform use restrictions | Single-family zoning abolished | Single-family zoning applied to 53 percent of land | All residential zones allow triplexes |
| Reform density restrictions | Inside transit zones: building height limits increased | 4 stories in downtown, up to 10 stories | 10 to 30 stories |
| | Outside transit zones: no major changes | 0.5 floor area ratio, minimum lot size ~5,000 square feet | 0.5 floor area, minimum lot size ~5,000 square feet |
| Remove parking-space minimums | Parking-space minimums eliminated citywide | Post-2015: minimums removed for < 50-unit multi-family homes near transit | No minimums; some parking-space maximums per unit |
| Bypass discretionary review | No major changes (Minneapolis has zoning variances and site plan review, but neither seems overly restrictive to development.) | | |
| Restore past multi-family standards | Remove occupancy limits for housing units<br>Codify single-room occupancy construction standards | | |
| Review building code standards | Not considered in reforms | | |
| **Strings attached** | **Main conditions** | | |
| Location requirements | Reforms concentrated in transit zones already in city plan: they cover downtown and blocks around light rail, bus infrastructure | | |
| Affordable housing requirements | More restrictive inclusionary zoning standards adopted; buildings with 20+ units must be paired with below-market-rate units | | |

*Sources*: Kuhlmann 2021; Neighbors for More Neighbors 2021.

*Note*: Rows shaded in blue indicate that the upzoning possibly removed binding zoning regulations. Rows shaded in yellow indicate no change in a possibly binding zoning regulation. Absence of shading means no prior regulation in the category was in effect.

What matters for evaluation is how reforms *changed* the status quo. For an outsider, knowledge of the degree of change can be difficult to discern—the status quo may have been complex or may not even have held constant between several competing jurisdictions. The benefit of having local planners on an evaluation team is to account for the status quo, ideally summarizing it through key statistics that easily fit into a table of reform intensity, such as changes in average floor area ratios.

## 3. Collecting Local Permit Data in Practice

### 3.1. Cleaning and geocoding data in a major city: Seattle's example

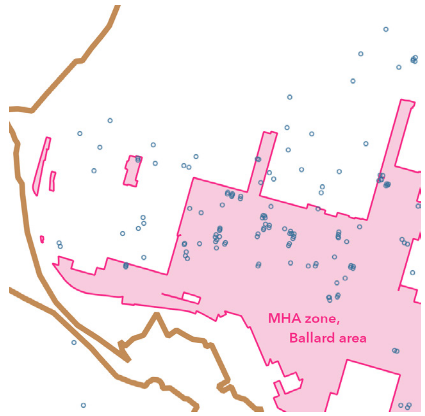
If resources were not a constraint, practitioners would do well to emulate the city of Seattle, which has had an open data portal since 2010, developed in close collaboration with Socrata, a Seattle software developer of national open data platforms (National League of Cities 2014).

**FIGURE 1.** How data ready for evaluation should look

a. Permit data as spreadsheet

| PermitNum | UnitsAdded | IssuedDate | OriginalAddress1 |
|---|---|---|---|
| 6761030-CN | NA | 12/4/19 | 947 NW 62ND ST |
| 6676461-CN | 7 | 3/18/19 | 917 NW 51ST ST |
| 6669315-CN | 0 | 6/17/19 | 6406 14TH AVE NW |
| 6622997-CN | 7 | 4/10/19 | 6300 26TH AVE NW |
| 6642811-CN | 7 | 2/19/20 | 2637 NW 59TH ST |
| 6670488-CN | 1 | 12/4/18 | 3013 NW 61ST ST |
| 6633419-CN | 6 | 3/5/19 | 6115 17TH AVE NW |
| 6641969-CN | 3 | 12/6/18 | 2442 NW 60TH ST |
| 6641970-CN | 0 | 12/6/18 | 2440 NW 60TH ST |
| 6667876-CN | 1 | 9/4/18 | 932 NW 61ST ST |
| 6632023-CN | 1 | 9/29/18 | 1539 NW 59TH ST |
| 6646497-CN | 1 | 11/7/18 | 932 NW 63RD ST |
| 6636586-CN | 1 | 3/8/19 | 2224 NW 63RD ST |
| 6601617-CN | 1 | 11/6/18 | 934 NW 63RD ST |
| 6633585-CN | 3 | 9/5/18 | 2656 NW 62ND ST |
| 6617665-CN | 4 | 10/10/18 | 901 NW 51ST ST |
| 6645801-CN | 1 | 9/3/18 | 314 B NW 41ST ST |
| 6645802-CN | 1 | 9/3/18 | 314 C NW 41ST ST |

b. Geocoded permits and zoning change




*Note*: Both panels display extracts from datasets available through Seattle Open Data. Both panels use permit data, while panel b combines geocoded permits with one neighborhood in Seattle that underwent a Mandatory Housing Affordability (MHA) zoning change.

Seattle's accessible data made it the most straightforward city to study in Krimmel and Wang (2023). Those authors used the city's zoning revisions to show that upzoned neighborhoods that have inclusionary zoning requirements for below-market-rent units end up with less new construction. While the authors considered several cities to analyze, they chose Seattle because its open data portal already had structured permit information and a shapefile of areas experiencing zoning change, generated from geographic information system (GIS) software.[6]

All permit and zoning data in Seattle comes from the Seattle Department of Construction and Inspections (SDCI), which processes permit applications through a website and digital vendor. Because the digital vendor likely already stores all permits in databases, SDCI probably extracts data from its digital database every day as a comma-delimited file (.csv) and then uploads it to Seattle's open data portal. Panel a of figure 1 provides a snapshot of the .csv file, while table 5 shows how nearly all the variables measuring the primary outcomes related to pro-housing reforms are included in SDCI data. Krimmel and Wang (2023) used those variables plus data on units removed to calculate net housing unit construction rates.

A researcher can learn how SDCI implemented its inclusionary zoning program from public documents, but more importantly, SDCI regularly uploads

6. Jake Krimmel and Betty Wang provided this context in interviews.

**TABLE 5.** Details of administrative permit data used by recent research

| | Krimmel and Wang (2023) | | Marantz, Elmendorf, and Kim (2023) | |
|---|---|---|---|---|
| Variable name | Recorded in Seattle data | Used in related study | Recorded in California HCD data | Used in related study |
| *Primary variables* | | | | |
| Site address | ✓ | ✓ | ✓ | ✓ |
| Building type | | | ✓ | ✓ |
| Number of units | ✓ | ✓ | ✓ | |
| Lot size | O | ✓ | | |
| Date of first entitlement approval | O | | ✓ | |
| Date of first building permit approval | ✓ | ✓ | ✓ | ✓ |
| Date of completion | ✓ | | ✓ | |
| *Recommended variables* | | | | |
| Site's geocoded location | ✓ | ✓ | | ✓ |
| Assessor's parcel number | | | ✓ | ✓ |
| Date of first permit submission | O | | O | |
| Estimated value | ✓ | ✓ | | |
| Project streamlined by state law | | | ✓ | |
| Units to be demolished on site | ✓ | ✓ | ✓ | |
| *Variables for further analysis* | | | | |
| Number of affordable units | | | ✓ | |
| Identity of developer | ✓ | | | |

*Note*: A check (✓) means the variable is included in the main dataset, while a circle (O) means the variable is in another public dataset that can be linked to the main one. HCD = Department of Housing and Community Development.

GIS shapefiles of zones where inclusionary zoning applies. Zone codes in the open data closely match what is described in policy documents.

Another critical factor is that Seattle geocodes all permit data in-house, matching observations to locations in space. It is then simple for outsiders to create spatial data, such as in panel b of figure 1, which shows an area experiencing a type of zoning change and surrounding permits. Appendix A outlines various methods to geocode data that are costly for enterprise uses, but on infrequent samples of permits, they can be applied cheaply or at no cost.

Overall, Seattle's open data are ready to use for outside evaluation because the data are structured to let enterprises—from real estate platforms such as Zillow to app developers—use the dataset easily.[7] Data structure for enterprise use can be easily adapted for policy evaluation. Unlike enterprises

7. In 2016, SDCI was an early adopter of a database standard for building permits—the Building & Land Development Specification. Data following the standard are machine readable by outside apps.

that demand data uploaded at daily or similar frequencies, housing supply outcomes adjust slowly enough that data cleaned monthly or quarterly can provide unique insights.

### 3.2. Data administration: Comparing Seattle with California's Department of Housing and Community Development

Another model for processing data is to have a state-level agency coordinate steps for data entry, with local authorities following that process. The state of California uses this model, and California's Department of Housing and Community Development (HCD) developed new monitoring powers for local land use in the late 2010s. The department now has around a hundred staff members, which is not so sizable given their responsibilities. Apart from tracking data, staff members must conduct reviews of the planning documents and ordinances of more than 500 local governments.

Because local authorities are still the entities that issue permits, they can hold up permitting and sidestep state laws that should legalize certain housing types statewide. But data tracked by HCD can also help the department's enforcement duties: Marantz, Elmendorf, and Kim (2023) proposed an automated method to detect localities that were holding up permitting of accessory dwelling units (ADUs). HCD should also audit local officials who issue fewer permits for ADUs per capita than what is predicted by the city's market conditions and neighborhood traits. HCD's system for data processing is standardized and comprehensive: local officials must report housing data through a spreadsheet template that HCD annotated with instructions.

Of the data that local officials report to HCD, the information that researchers likely care most about is shown in the spreadsheet's table A2. This form lists the status of buildings in a calendar year whose permits moved through a stage of the process, from the end of discretionary review to project completion. Local officials also categorize the building type and self-reported unit affordability, as well as whether state law prevents any discretionary review.

While HCD appears to use statewide parcel databases to geocode the data, permit coordinates are not available to external researchers. Table 5 shows the available variables and which ones were used by Marantz, Elmendorf, and Kim (2023); note that the study authors had to geocode permits themselves. In addition, HCD collects zoning changes from cities in the cities' housing element documents, which contain zoning maps difficult to manipulate by outsiders because they are provided as PDF (portable document format) files. Legal scholars have noted that the bottleneck in enforcement could be solved if HCD submissions included standard zoning shapefiles rather than PDF files (Elmendorf et al. 2021).

The HCD system's main goal remains to verify compliance with state laws. Its annual progress report process requires standardization and reporting beyond the variables most useful for enterprises and reform evaluation: HCD also requires self-reported affordability or categorization of each building on the basis of which set of state laws applies to it.[8] A tradeoff in understanding a reform's long-run impacts emerges between (a) reporting too many variables when a reform is piloted to see whether the reform has an effect and (b) reporting too few variables.

For practitioners starting to structure their administrative data, the Seattle and HCD case studies offer different paths to implementing data collection. HCD collects data with a setup that has lower fixed costs,[9] while Seattle's data are well maintained because a central team coordinates data structuring in cooperation with other departments (Seattle Information Technology 2016).

In both cases, the agencies collect more data than are necessary for evaluating reforms. One suggestion is to pilot the data collection to focus on what is necessary to produce new results. Table 6 offers a guide to crucial variables for evaluation and other variables recovered from permits that can offer market signals about reform impacts.

## 4. A Tour of Causal Inference Research Designs

This guide concludes with introductions to three research designs that are often used by academics to evaluate policies. Even though there could be many confounders related to when and where reforms were adopted, these designs deliver estimates that are not biased by how much those confounders matter—as long as the assumptions behind the designs hold. This section also describes how an evaluation team can determine whether the results are precise or uncertain and how to best situate results among many possible evaluations across jurisdictions.

8. City planners have raised concerns with HCD about how to collect these additional variables when the department cannot clearly guide the process. See the Q&A during the General Plan and Housing Element Annual Progress Reports Webinar, held by the Governor's Office of Planning and Research on October 6, 2022. The relevant portion can be found after minute 41:00 of https://www.youtube.com/watch?v=Hha9IQrEVqQ.

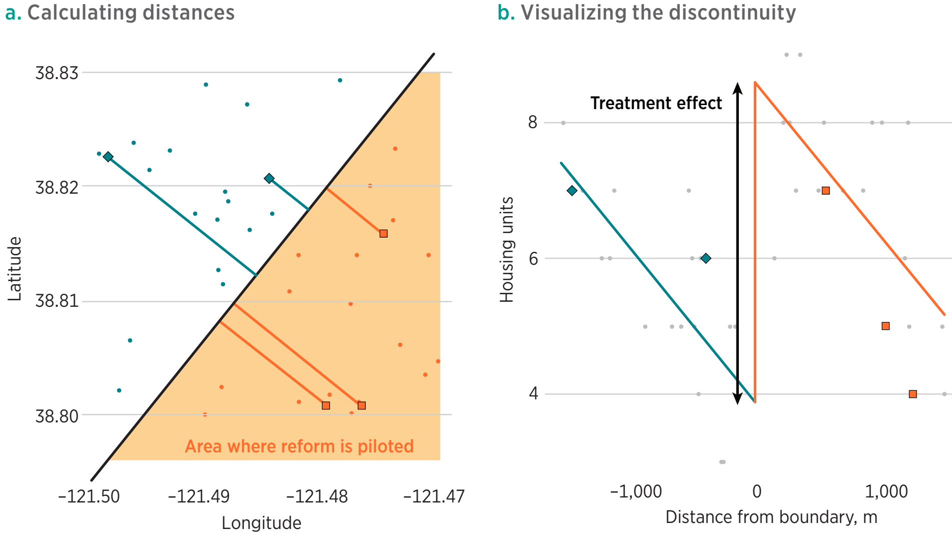
9. As of 2019, Seattle employs four people on its open data team, with an annual budget of $788,000. Costs likely come from salaries and licensing fees for the open data portal software.

**TABLE 6.** Suggested guidelines for reporting administrative permit data

| Primary variables | Guidelines |
|---|---|
| Site address | Standard post office format so that it can be geocoded |
| Building type | Selected from a list of predefined types to validate data more easily |
| Number of units | Self-reported by developer in submitted permit |
| Lot size | Ideally based on assessor records for the parcel or lot |
| Date of first entitlement approval | Based on when any land use or discretionary review was completed so the building permit agency can begin review; leave this column blank for by-right development |
| Date of first building permit approval | Based on building permit issuance |
| Date of completion | Based on when the building was approved for occupancy (from the building permit agency or another agency) |
| **Recommended variables** | **Guidelines** |
| Site's geocoded location | Expressed in coordinates and able to be visualized in a geographic information system |
| Parcel assessor's parcel number (APN) | Simplifies linking permit data to existing assessor records |
| Date of first permit submission | Based on when developer submits first required form in the entitlement or building permit process |
| Estimated value | Either self-reported by developer or estimated by city planners; distinguishing the two possible sources would improve reporting |
| Project streamlined by state law | Includes columns for each possible policy that overrides local zoning, such as California's ministerial approval |
| Units to be demolished on-site | Important variable to derive net new units built by developer |
| **Variables for further analysis** | **Guidelines** |
| Number of affordable units | Calculating affordability implies some estimate of rent or home prices, either reported by developer or by planners |
| New parking spots planned | Record net new spots (i.e., change from any parking already on proposed site); this value can be negative |
| City ward or neighborhood | Especially useful for cities where ward officials can request review |
| Project uses housing tax credits | Track whether projects are financially viable as a result of federal low-income housing tax credit or state incentives |
| Floor plan characteristics | Total floor space, possibly divided into types of floor space use; possibly a verbal note that can be processed later by collaborators |
| Identity of developer | In practice, the holding corporations will be reported, but outside collaborators could analyze data |

In practice, it is assumed that external researchers will collaborate with practitioners on outputting results from these research designs. This section's goal is to provide guidance about when each design is useful and to highlight the data requirements. To better understand how to run the analysis, readers should consult two well-received textbooks: Huntington-Klein (2021) and Cunningham (2021). Both books emphasize that the statistical models behind these designs are primarily linear regression models. Footnotes throughout this section provide more detailed guidance on how to code the analysis and produce results.

**FIGURE 2.** Using data to conduct a boundary discontinuity design



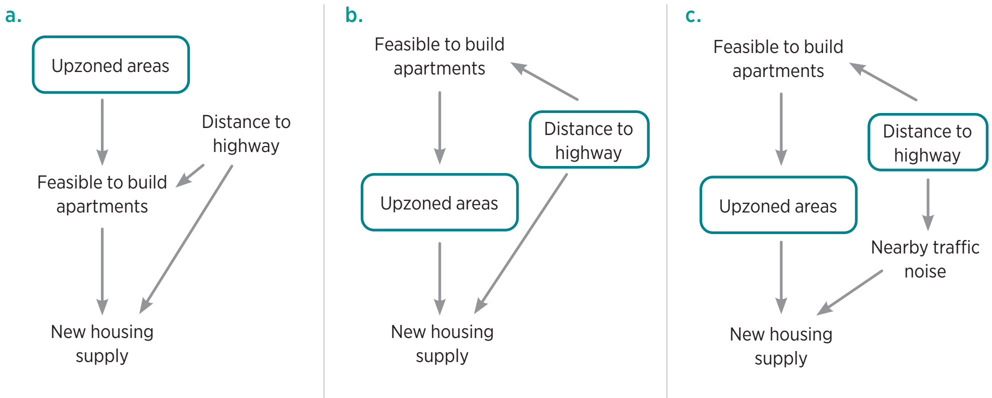

## 4.1. Research designs exploiting changes in space

Instead of changing the land use of individual parcels, suppose a local authority drew a straight line through a city's borders. All land on one side of the line is subject to new pro-housing reforms, and none of the land on the other side is subject to reforms. When properties within a city are compared, it is expected that they would receive similar public services, such as schools. By comparing only the properties close to this line, one can infer that there should be no difference in access to public services or nearby amenities before the reform. Therefore, the parcels close to the line where reforms do not apply appear to be good counterfactuals for the treated parcels.

The technical term for this kind of study is the *boundary discontinuity design*. If a pro-housing reform is passed for an area along new boundaries and researchers know the geographic coordinates for lots across the city, then they can measure how close the lots are from the boundary line on the basis of the shortest distance lines. Analyzing the data together shows the difference effect of the reform on outcomes by comparing the outcome value next to the boundary on one side to the value on the other side.

A feature of this research design is that the way to back out estimates can be directly visualized, as can be seen in figure 2. Panel a visualizes the first data cleaning step. Given the locations of observations in space and the reform boundaries, researchers can use GIS programs to calculate the nearest distance to the

**FIGURE 3.** Using causal diagrams to organize confounders




*Note*: Variables in blue boxes are those that must be collected in the analysis for the research design to estimate causal effects.

boundary.[10] Panel b visualizes the design itself, illustrating how the observation's outcomes change near the boundary.

The treatment effect estimated under the design is the difference in predicted values, estimated separately on both sides. Except in rare cases where most of the observations are very close to the boundary, the predicted values usually come from a curve-fitting exercise on both sides—the regression prediction using observations near and far from the boundary.[11]

## 4.2. The logic of causal inference designs

In the method described in subsection 4.1, it was critical that the straight line was arbitrary: the line was not meant to avoid or go through specific areas. No confounder, past or present, can be used to explain why the line was drawn the way it was, and that fact can be represented with a causal diagram (see figure 3). In the diagram, variables are connected with paths that represent what are assumed to be true relationships between them. Panel a of figure 3 represents an environment where upzoned areas were arbitrarily determined; for example, confounders likely do not affect which areas now allow tall apartments.

10. It might be useful to set several points on the boundary and calculate the distance matrix between the observation and those points, keeping only the smallest value for each observation.
11. More advanced discussions, including code and what to do if locations are not correctly observed, are in Keele and Titiunik (2015)the regression discontinuity design (RD.

If a city allows apartments next to a highway only and the reform is limited to those areas where apartments had been zoned, the result is a diagram like the one in panel b. It makes sense that buildings next to a highway command lower rents, which means developers would be unwilling to build next to them. It then becomes unclear whether a treated parcel remains undeveloped because the reform does not work or whether distance to the highway matters more for development. A clearer answer is possible if the researcher can control for this confounder, blocking off the causal path by collecting data for that variable and using statistical methods.

A twist on this model is shown in panel c, where it might appear that the traffic noise from a highway, not the sight of the highway, makes developers unwilling to build. But it is assumed that if a tenant lives farther from the highway, the tenant will hear less noise. If distance to highway is controlled for, that blocks off the confounder of more noise from upzoned areas so that the estimates are not confounded. Moreover, with this model, there is no need to record traffic noise.

If administrative data are cleaned and geocoded as described in section 3, it is much easier to merge the observed confounders into free spatial data. Appendix table A3 lists some datasets containing likely confounders, available at fine to neighborhood-level scales. Furthermore, confounders that matter for one outcome may not be relevant for another; for example, planners might conclude that developers are no less likely to build townhomes near a highway than to build anywhere else.

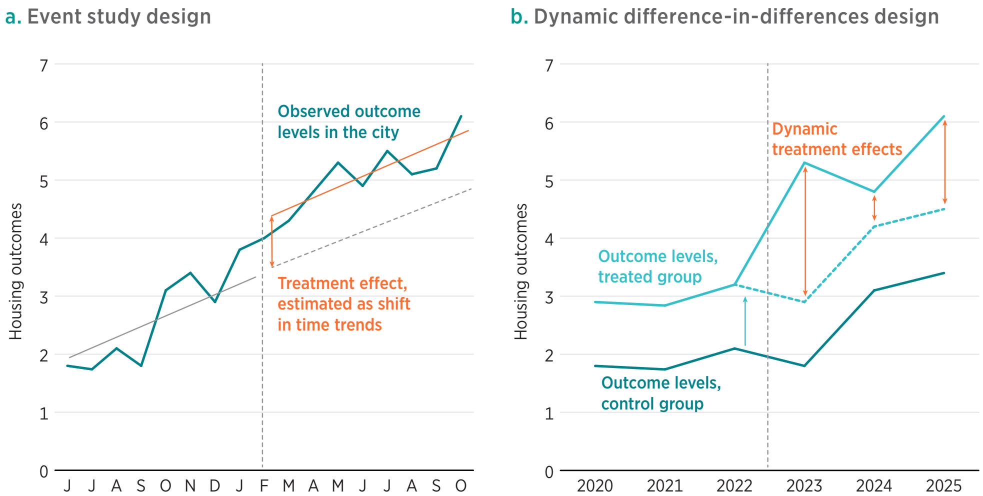
## 4.3. Research designs exploiting changes in time

If planners know about a confounder that determines why some areas rather than others had a pro-housing reform, the boundary discontinuity design is invalidated. The most common alternative design in this case is to compare changes in outcomes over time instead of at one point in time. It is then necessary to collect some data on outcomes from at least one period before the policy was in effect.

At a citywide level, a simple time-varying design is to compare levels and rates in an earlier year versus a later year. Or the design could use an average of values across different blocks and census tracts—specifically, the ones with potential to be affected by the reforms. Methods that adjust for time trends are called *event studies*.

An event study controls for many confounders only if the timing of a confounding factor was unpredictable—that is, not explained by market conditions and not announced in advance. This assumption is unrealistic because most

**FIGURE 4.** Identifying effects from changes in trends




*Note*: Panel a shows data at a monthly frequency (June, July, August, etc.), while panel b does so annually.

reforms are responses to growing unaffordability. Therefore, the more accepted method is to compare post-reform trends for treated areas with those of a valid control group. With two groups and at least two periods, one can run a difference-in-differences (DiD) design.

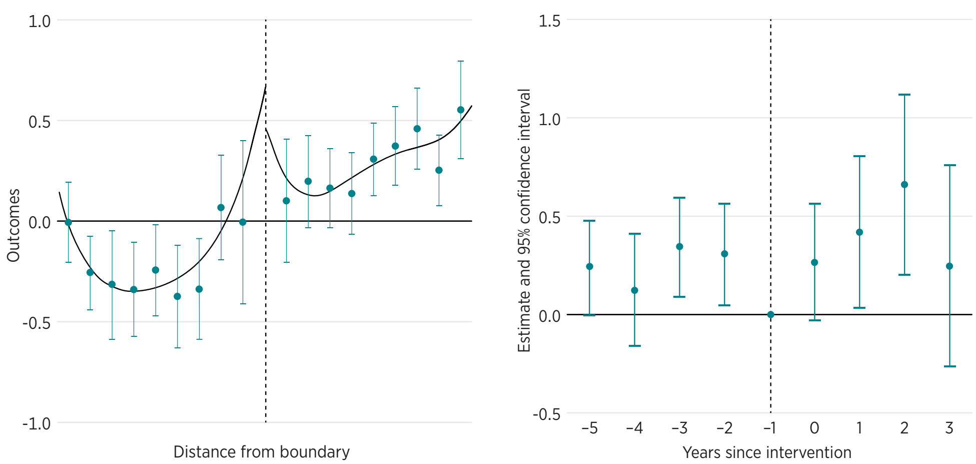
Figure 4 plots the ways that an event study and a DiD design can isolate treatment effects. Suppose some reforms legalized duplex construction everywhere, and those reforms were not anticipated by developers. The event study illustrated in panel a looks at how growth in a city's outcome changed before and after the month of adoption of a new policy. This research design works to compare what occurs after reforms to the correct counterfactual, assuming that, without the reform, one could accurately predict the market on average by projecting the past trends.

Panel b shows a dynamic DiD design, in which an outcome can be averaged over two groups for more than two periods. Suppose some reforms allowing duplex construction were adopted in response to increasing unaffordability—but only on one side of a main road going through the city. The hope is that, on the side where reforms do not apply, housing outcomes there reflect the same market confounders driving unaffordability as they would on the side seeing reforms. How researchers isolate the counterfactual at this design is more sophisticated. They would take the trend over the control group, which can be observed, and shift it to the treated group's baseline level before the reform took place.

**FIGURE 5.** Adding confidence intervals to visualize statistical error

a. Error in boundary discontinuities

b. Error in dynamic difference-in-differences




The boundary discontinuity design also requires individual geolocated units, but coding the distances is not necessary for a DiD design. The treated and control groups could be different parcels of land, or they could be average outcomes across neighborhoods or cities. Even if only one city is observed, such that averages cannot be determined, there are well-known models, such as the synthetic control model, that work in similar ways.[12]

## 4.4. Visualizing evolving findings

All the research designs described up to now have strong visual elements. Convincing laypeople that reforms have impact is easier when they see a change in trends right after a reform's introduction. Prioritizing a visual over a table also emphasizes data transparency; the data in hand can support whether a control group provides good counterfactuals. In academic research, the degree of confidence in a counterfactual can change once the uncertainty is plotted in the estimates.

Panel a in figure 5 presents what a boundary discontinuity design plot should look like on more realistic data. Trends around the boundary are visualized in two ways. First, the parcel observations are put into groups. What is plotted is not every observation in the sample, but sample averages among

12. For synthetic control models, McClelland and Mucciolo (2022) provide a helpful guide for an entire evaluation team.

observations close to each other. Statistical software also produces confidence intervals around each group of observations: the more observations in the group, the less uncertain the value could be due to statistical error. Second, data around the boundary are used to fit curves. These curves predict values right around the boundary, which gets at the treatment effect like in subsection 4.2.

Figure 2 illustrates how to use real data to estimate outcome values on both sides of the boundary. In the case of figure 5, however, things go wrong: The graph in panel a suggests little data close to the boundary. One could estimate values using trends in the data available, but here those estimates extrapolate to the border oddly. Wide confidence intervals and unusual extrapolations all point to the need for more data to estimate smaller effects—rather than altering the models to suggest more certainty than there is.

Panel b shows a standard dynamic DiD plot with confidence intervals for each period. Instead of plotting the separate trends for treatment and control groups, here the dynamic effects like those highlighted in figure 4 are plotted. The estimates should reflect two tradeoffs. First, the dynamic estimates do not appear to be the same across all periods after the intervention, but they are individually noisy. If one wants to assume the effects should be the same over multiple years, pooling the effects can increase statistical accuracy.

Second, in some years before the intervention, the outcomes in the treated group were significantly different from those in the control. This effect is called a *violation of parallel pre-trend*s: if the treated areas diverge in some way before the treatment started, there might be a confounder that continues to change treated areas even without a policy intervention.

The overarching idea is that the evaluation is not useless just because statistical noise is high or assumptions appear to be violated. The first step is to rethink how quickly data on outcomes can be collected or whether data on confounds can be collected and controlled for. If effects are noisy or very large, they should be considered with respect to a larger body of evidence. Subsection 4.5 discusses this point further.

## 4.5. Registering pre-analysis plans

Much of what has been discussed—generating hypotheses and outcomes, and describing the policy intervention, basic methodology, and possible confounders—is part of what social scientists must do to obtain findings on piloting experiments. To ensure peer reviewers expect the evaluation to be high quality, researchers can write a pre-analysis plan to specify how they decided to analyze the data and account for any shortfalls or risks.

One reason to consider writing the pre-analysis plan is to lay out the evaluation plan in a standard way, making it easy for outside organizations to understand and allocate grants to support the plan. Another reason is that with many areas piloting reforms, sharing pre-analysis plans can simplify communication between local officials on what outcomes each one is investigating.

A pre-analysis plan should not cover all potential analyses or downsides. Rather, it offers a clear overview of why a study will produce new findings that are credible (Duflo et al. 2020). A research team uses a checklist of evaluation aspects to investigate; appendix B offers an example list, based on Chuang and Wykstra (2015). Olken (2015) describes an economist's experience writing pre-analysis plans that is approachable for researchers and practitioners.

Two sections of the checklist are worth explaining in greater detail. Even in a perfect study, there are many ways in which chance alone in how the sample is chosen can result in strong effect estimates in either direction. The discussion in subsection 4.4 addresses visualizing confidence intervals or determining whether a null hypothesis of no policy effect can be rejected. The other angle is whether the research design has statistical power. If the outcome varies greatly in the population, but practitioners did not expect the reform to have large effects at some point in time, testing the hypotheses will require more data to reject a specific hypothesis with a degree of confidence.[13]

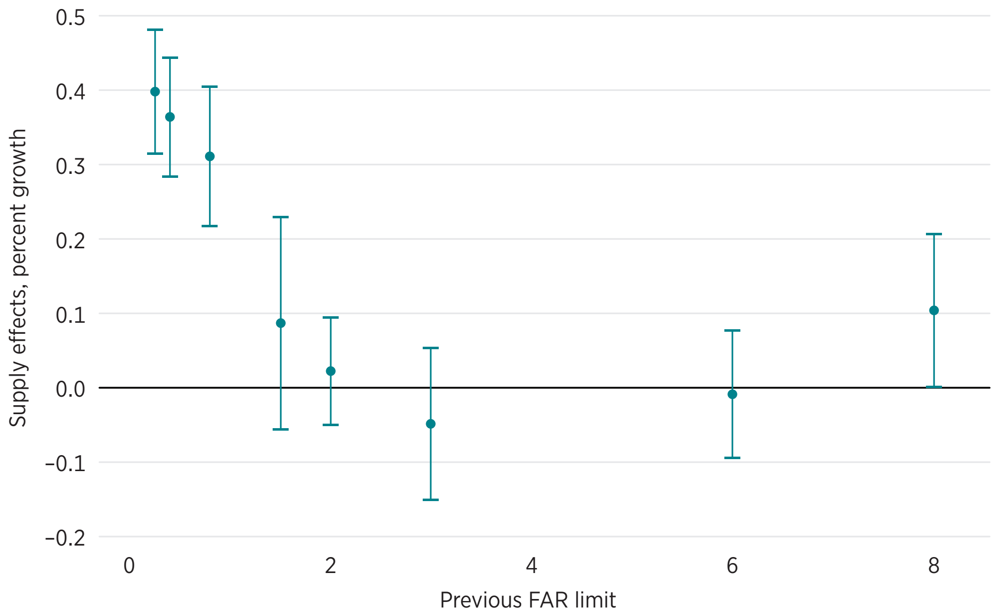
The other important factor for pro-housing reforms is the expected timeline. Active debate about the reforms occurs at all levels of government, so policies studied today could be repealed in the future or superseded by another law at the local or state levels. If that happens, the current evaluation should be thought to be terminated. In addition, reforms tend to be staggered—often first piloted in specific zones or neighborhoods before being scaled up.

Because home construction can take years to complete even without constraints, a timeline for seeing effects could be one to two years—or even longer. It is worth noting any risk of repeal or potential for further reforms in the short run. The same caution applies if the reforms were rolled out over multiple periods, an example of staggered adoption.[14]

---

13. One way to make these "power calculations" is to use an existing calculator, such as the PowerUp! tools for Microsoft Excel (Dong and Maynard 2013). In general, statistical power is higher if researchers plan for larger samples or if they think interventions have large effects relative to the overall noise in data outcomes.
14. There is a technical concern here: some statistical models do not deliver meaningful causal effects if policy rollout is staggered and planners also expect dynamic effects that vary according to place. McKenzie (2022) offers an advanced overview that is not overly complex.

**FIGURE 6.** Hypothetical reform effects to use in a meta-analysis




*Note*: This figure simulates how effects can be estimated for a standard reform (here, elimination of all parking-space minimums) but separately, depending on a built environment's characteristics (previous floor area ratio [FAR] limits). Each study has a certain confidence interval, but additional meta-analysis methods can pool studies to obtain more-precise relationships.

## 4.6. From analysis to meta-analysis

Even though reform usually involves many policies being adopted at once (as section 2 points out), this subsection discusses a treatment effect as if it flips on or off like a switch. That said, subsection 2.3 offers a template to code the intensity of policy changes. As an example, it becomes easier to learn from multiple cities eliminating their parking-space minimums if it's understood what the existing market conditions and the macroeconomic trends affecting housing investment were at the time.

The general idea for combining results from separate trials to obtain additional estimates is called *meta-analysis*. This is a common procedure for more precisely calculating the effects of drug treatments in healthcare, so the statistical method is well established. The logic of meta-analysis can also extend to combining separate pro-housing reforms in a time of policy experimentation. Practitioners need only to ensure that different results are comparable by converting the data to the same units.

Figure 6 illustrates results ready for meta-analysis. The data could, for example, represent high-quality estimates of the effects of the elimination of parking-space minimums on housing supply in separate cities. Each study has statistical uncertainty and different market conditions, such as the FARs established in areas undergoing policy change. The data could also come from a single citywide elimination of parking-space minimums, with effects estimated separately by different zones. Knowing these design differences matters because the lack of a control group not undergoing reform, as in the latter case, makes these estimates less credible.

The outcome is not expressed in the number of new units built but is a measure of growth in housing available—in this case, a percentage growth in housing stock. The outcome is then not in terms of units, which makes sense if separate results are calculated over larger or smaller areas. A meta-analysis would then add further assumptions to model new estimates—for example, whether planners can statistically reject the hypothesis that the parking reform levels off in impact when existing construction is at a FAR of 2.

A final difficulty with interpreting effects, which is a judgment call, is the potential for spillover effects. If a reform legalizes apartment construction citywide, the developer of a new apartment might have dropped a project to build single-family homes in certain neighborhoods. The treatment effect in this case, therefore, cannot account for the opportunity cost of leaving those single-family homes unbuilt.

Recent best practices in the academic literature describe how to use past trends in outcomes that may have received a spillover effect to calculate how much spillovers matter. External researchers interested in applying these methods should refer to Rambachan and Roth (2023), as well as to an application of the methods to zoning in Greenaway-McGrevy and Phillips (2023).

## Conclusion

As housing reforms pass and real data trickle in, this paper offers a road map to pro-housing policy evaluation. High-quality evaluations follow some common practices to produce estimated effects—additional economic approaches help planners apply estimates to determine which policies might relieve the housing crisis. The research agenda is broad: the amount of ongoing policy experimentation at the state and city levels offers many avenues to understanding which pro-housing reforms can affect the market. Evidence from these policies needs to be communicated in a common language, using standard outcomes and empirical methods described in this paper's road map.

As a starting point for the research agenda, this paper walks through the causal inference approach to policy evaluation. There are other credible ways to analyze open data. One is descriptive analyses to accurately identify market responses to reforms, and another is to feed open data into market simulators to model how likely certain projects are to be financially feasible (Metcalf 2024). All these analyses pursue a common goal: use the policy experiments occurring around the world to establish some basic facts so that practitioners can learn in real time about both the limits and the potential of policy reforms.

## Appendix A: Primer on Geocoding Data

Geocoding housing data for analysis is crucial: research designs or descriptive analysis cannot start without a structured way to format an observation's location. In an ideal world, each state has a digitized map of parcels that locate any parcel in space. When those data do not exist, any local agency with a building's address can still match it to geographic coordinates, either for free or at reasonable rates.

The free way to geocode for data within the United States is to use the Census Bureau's census geocoder. Through the bureau's website, anyone can upload a spreadsheet containing addresses and receive the same file back with longitude and latitude coordinates for that address. The geocoder FAQ explains how the address needs to be formatted and saved; Excel spreadsheets (.xlsx files) are acceptable. To get coordinates from the geocoder, choose the "Find Locations" option. An alternative is to choose "Find Geographies," which instead gives the census tract and block of the address. This alternative might be helpful for those running a DiD design over blocks, but less so if it is a boundary discontinuity requiring precise distances.

A reasonably priced option is Google's geocoding service. Specifically, if the observations with addresses are loaded into Google Sheets, one can use a plug-in to run the Google geocoder from the Sheets file. One example is the "Geocoding by SmartMonkey" plug-in.[15] As with the Census Bureau's geocoder, addresses must be input in a certain format. Each account gets 500 free requests before a paid plan is required ($5 per 1,000 rows).

For GIS users, another alternative is to upload addresses to batch geocoders that come with a GIS application. ArcGIS users have the ESRI geocoder, while QGIS users have the MMQGIS geocoder, which relies on a different address dataset, OpenStreetMap.

More step-by-step guides can be found in Dougherty and Ilyankou (2024).

15. The plug-in can be found at https://workspace.google.com/marketplace/app/geocoding_by_smartmonkey/1033231575312.

## Appendix B: Checklist for Pre-analysis Plans

Note: This appendix adapts a checklist provided by Chuang and Wysktra (2015) for randomized trials of policies. Details relevant for randomized trials are replaced with more description of the pro-housing reforms studied.

- ☐ General information
  - ☐ Title of the project
  - ☐ Researchers involved
  - ☐ Outside partner institutions (if applicable)
- ☐ Introduction
  - ☐ Project summary
  - ☐ Aims, rationale, and background
- ☐ Study design
  - ☐ Hypotheses
  - ☐ Treatment effects and measurement
    - ☐ Primary outcome, precisely defined
    - ☐ Secondary outcome, precisely defined
    - ☐ Data sources
  - ☐ Details of intervention
    - ☐ Table that categorizes all passed reforms
    - ☐ Intensity of changes or strings attached
    - ☐ Boundaries of areas with reforms (if applicable)
  - ☐ Details of research population
    - ☐ Demographic information on targeted areas
    - ☐ Housing market indicators ahead of reform
  - ☐ Inclusion or exclusion criteria
  - ☐ Research design or statistical model
    - ☐ Randomization procedure (if applicable)
    - ☐ Regression form for design
    - ☐ Potential confounders
    - ☐ Differences in confounders across groups
    - ☐ Power analysis (if applicable)
  - ☐ Expected timeline
    - ☐ Staggered adoption of reforms
    - ☐ Early termination possibility
    - ☐ Potential superseding laws

# Appendix Tables

**TABLE A1. Accounting for all reforms applied during Auckland upzoning**

| Policy | Reforms applied | Pre-reform | Post-reform |
|---|---|---|---|
| Reform use restrictions. | In mixed housing zones: additional uses allowed<br>Apartment zone: unrestricted residential | Over 400 sq. kilometers zoned for below 0.8 floor area ratio; 4 stories and, in downtown, up to 10 | All upzoned areas allow triplexes |
| Reform density restrictions. | Greater density in all new zones<br>Yard setbacks replaced with site coverage ratios<br>New minimum size for housing units | | Max. heights set to 5 stories for apartments<br>Uniform min. house size at 45 sq. meters |
| Remove parking-space minimums. | Progressive removal of parking-space minimums | | |
| Bypass discretionary review. | Relaxation of historical preservation, architectural review | | |
| Restore past multifamily standards. | No major changes | | |
| Review building code standards. | No major changes (New Zealand's multifamily building code is in line with global standards). | | |
| **Strings attached** | **Main conditions** | | |
| Location requirements | Mixed housing zones make up most of the urban core, leading to 3/4 of land being upzoned.<br>Apartment zones are areas near existing transit (transit-oriented development). | | |
| Affordable housing requirements | No major requirements; special housing areas, where they were applied, were rolled back. | | |

*Sources*: Greenaway-McGrevy and Jones 2023; Greenaway-McGrevy and Phillips 2023.

*Note*: Rows shaded in blue indicate that the upzoning removed possibly binding zoning regulations. Absence of shading means no prior regulation in the category was in effect.

**TABLE A2. Sample results on housing market outcomes for Minneapolis and Auckland**

| Minneapolis | | Auckland | |
|---|---|---|---|
| Outcome | Paper and Research Design | Outcome | Paper and Research Design |
| Units | None as of August 2023 | Units | Greenaway-McGrevy and Phillips (2023): difference-in-differences across zoning borders<br>Most conservative estimates imply that reform resulted in 4,400 more permitted units each year (0.8 percent of housing stock). |
| Rents | Flisrand (2022); Maltman (2023): descriptive analysis (no explicit research design)<br>Declines in rents over 5 years for 1 and 2 bedrooms, possibly 12 percent from control group | Rents | Greenaway-McGrevy (2023): synthetic control (trend differences between cities)<br>Declines (e.g., in 6 years, average 2 bedroom rent decreased 16 to 28 percent from control group) |
| House values | Kuhlmann (2021): difference-in-differences across city borders<br>Increases after 2 years for developed parcels; larger effects for lots with open space | House values | Greenaway-McGrevy, Pacheco, and Sorensen (2021): difference-in-differences across zoning borders within a city<br>Increases for all developed parcels, except ones that have denser apartments |

**TABLE A3.** Suggested datasets that include control variables

| Dataset | Fineness of geography | Variables included |
|---|---|---|
| Decennial census, 100 percent data | As small as a census block (corresponding to urban blocks) | Population, housing unit, and race and ethnicity breakdowns |
| American Community Survey | Census tract (neighborhoods of ~4,000 people), as small as a block group (600–3,000 people) | Demographics, plus average income, rent, and home value data |
| Global human settlement layer | Most are available at 100-meter raster resolutions, some at 10 meters | Measures of land use intensity and density of built structures, based on satellite data |
| EPA air quality system | Data come from monitors; most easily aggregated to county level | Historical air quality measures, such as exposure to different particulates |
| Stanford education data archive | School district–level data or county-level data | Demographics of school districts and measures of student achievement |
| OpenStreetMap files | Maps precisely marking objects in space | Can query data to get nearby transit routes, highways, or urban attractions |
| FBI Uniform Crime Reports | City police department level; check whether local police department has more refined data | Annual counts of homicides or other violent and property crimes; police departments vary in the accuracy of reporting |

## References

Asquith, Brian J., Evan Mast, and Davin Reed. 2023. "Local Effects of Large New Apartment Buildings in Low-Income Areas." *Review of Economics and Statistics* 105 (2): 359–75.

Been, Vicki, Ingrid Gould Ellen, and Katherine M. O'Regan. 2023. "Supply Skepticism Revisited." NYU Law and Economics Research Paper 24-12, New York University Furman Center, New York, November.

Been, Vicki, Alex Jonlin, and Noah Kazis. 2023. "Encouraging Transit-Oriented Development." *Critical Land Use and Housing Issues for New York State*, Issue 1, New York University Furman Center, New York.

Brey, Jared. 2023. "Zoning Changes Have Small Impact on Housing Supply." *Governing*, April 10, 2023.

Brooks, Leah, and Jenny Schuetz. 2023. "Does Housing Growth in Washington, DC, Reflect Land Use Policy Changes?" *Learning from Land Use Reforms: Housing Outcomes and Regulatory Change*, New York University Furman Center, New York.

California Department of Housing and Community Development. 2023. "San Francisco Housing Policy and Practice Review." Housing Policy Development Division, San Francisco, October.

Chuang, Erica, and Stephanie Wykstra. 2015. "A Guide to Pre-analysis Plans." Materials 2016 BSPS Workshop, Innovations for Poverty Action, Washington, DC, September 24, https://osf.io/sh4v8.

Cunningham, Scott. 2021. *Causal Inference: The Mixtape*. New Haven, CT: Yale University Press.

Dong, Nianbo, and Rebecca A. Maynard. 2013. "PowerUp!: A Tool for Calculating Minimum Detectable Effect Sizes and Minimum Required Sample Sizes for Experimental and Quasi-Experimental Design Studies." *Journal of Research on Educational Effectiveness* 6 (1): 24–67.

Dougherty, Jack, and Ilya Ilyankou. 2024. *Hands-On Data Visualization: Interactive Storytelling from Spreadsheets to Code*. Sebastopol, CA: O'Reilly Media.

Duflo, Esther, Abhijit Banerjee, Amy Finkelstein, Lawrence F. Katz, Benjamin A. Olken, and Anja Sautmann. 2020. "In Praise of Moderation: Suggestions for the Scope and Use of Pre-analysis Plans for RCTs in Economics." NBER Working Paper 26993, National Bureau of Economic Research, Cambridge, MA, April.

Eastland, Maggie. 2023. "The Suburb That Defied Nimby." *Wall Street Journal*, August 16.

Eliason, Michael. 2021. "Seattle's Lead on Single Stair Buildings." *The Urbanist*, December 20.

Elmendorf, Christopher S., Eric Biber, Paavo Monkkonen, and Moira O'Neill. 2021. "State Administrative Review of Local Constraints on Housing Development: Improving the California Model." *Arizona Law Review* 63 (1): 609–77.

Flisrand, Janne. 2022. "Minneapolis Rents Drop, Bucking National Trends." streets.mn, May 6. https://streets.mn/2022/05/06/minneapolis-rents-drop/.

Freemark, Yonah. 2023. "Zoning Change: Upzonings, Downzonings, and Their Impacts on Residential Construction, Housing Costs, and Neighborhood Demographics." *Journal of Planning Literature* 38 (4): 548–70.

Gardiner, Dustin, and Susie Neilson. 2022 "627 Days, Just for the Permit: This Data Shows the Staggering Timeline to Build Homes in S.F." *San Francisco Chronicle*, December 14.

Goggin, Brian. 2018. "Measuring the Housing Permitting Process in San Francisco." Research and Policy Professional Report, Terner Center for Housing Innovation, University of California, Berkeley, July 24.

Gray, M. Nolan. 2021. "How Californians Are Weaponizing Environmental Law." *Atlantic*, March 12.

Greenaway-McGrevy, Ryan. 2023. "Can Zoning Reform Reduce Housing Costs? Evidence from Rents in Auckland." Working Paper 016, Economic Policy Centre, University of Auckland Business School, Auckland, New Zealand.

Greenaway-McGrevy, Ryan, and James Allen Jones. 2023. "Can Zoning Reform Change Urban Development Patterns? Evidence from Auckland." Working Paper 012, Economic Policy Centre, University of Auckland Business School, Auckland, New Zealand, December.

Greenaway-McGrevy, Ryan, Gail Pacheco, and Kade Sorensen. 2021. "The Effect of Upzoning on House Prices and Redevelopment Premiums in Auckland, New Zealand." *Urban Studies* 58 (5): 959–76.

Greenaway-McGrevy, Ryan, and Peter C. B. Phillips. 2023. "The Impact of Upzoning on Housing Construction in Auckland." *Journal of Urban Economics* 136 (July): 103555.

Gyourko, Joseph, Jonathan Hartley, and Jacob Krimmel. 2021. "The Local Residential Land Use Regulatory Environment across U.S. Housing Markets: Evidence from a New Wharton Index." *Journal of Urban Economics* 124 (July): 103337.

Gyourko, Joseph, and Raven Molloy. 2015. "Regulation and Housing Supply." In *Handbook of Regional and Urban Economics* vol. 5B, edited by Gilles Duranton, J. Vernon Henderson, and William Strange, 1289–337. Amsterdam: North-Holland.

Hamilton, Emily. 2023a. "How DC Densified." *Works in Progress*, Issue 11, May 23, 2023.

Hamilton, Emily. 2023b. "Four Elements of a Successful Housing Task Force: Lessons from the Montana Miracle." Mercatus Policy Brief, Mercatus Center at George Mason University, Arlington, VA, August 23.

Horowitz, Alex, and Ryan Canavan. 2023. "More Flexible Zoning Helps Contain Rising Rents." Pew Charitable Trusts, Washington, DC, April 17.

Huntington-Klein, Nick. 2021. *The Effect: An Introduction to Research Design and Causality*. Boca Raton, FL: CRC Press.

Jordan, Tony. 2019. "What Is Parking Reform?" *Parking Reform Network*, December 19.

Justus, Andrew. 2023. "One Year Later: Pre-Approved Home Plans for South Bend Infill." Analysis/Commentary, Niskanen Center, Washington, DC, August 7.

Kahn, Eli, and Salim Furth. 2023. "Breaking Ground: An Examination of Effective State Housing Reforms in 2023." Mercatus Policy Brief, Mercatus Center at George Mason University, Arlington, VA, August 1.

Keele, Luke J., and Rocío Titiunik. 2015. "Geographic Boundaries as Regression Discontinuities." *Political Analysis* 23 (1): 127–55.

Krimmel, Jacob, and Betty Wang. 2023. "Upzoning with Strings Attached: Evidence from Seattle's Affordable Housing Mandate." *Learning from Land Use Reforms: Housing Outcomes and Regulatory Change*, New York University Furman Center, New York.

Kuhlmann, Daniel. 2021. "Upzoning and Single-Family Housing Prices." *Journal of the American Planning Association* 87 (3): 383–95.

Kulka, Amrita, Aradhya Sood, and Nicholas Chiumenti. 2023. "Under the (Neighbor)Hood: Understanding Interactions among Zoning Regulations."

Maltman, Matthew. 2023. "A Detailed Look at the Outcomes of Minneapolis' Housing Reforms." *One Final Effort* (blog), April 17.

Manji, Shazia, Truman Braslaw, Chae Kim, Elizabeth Kneebone, Carolina Reid, and Yonah Freemark. 2023. "Incentivizing Housing Production: State Laws from across the Country to Encourage or Require Municipal Action." Terner Center for Housing Innovation, University of California, Berkeley, and Urban Institute, Washington, DC.

Marantz, Nicholas J., Christopher S. Elmendorf, and Youjin B. Kim. 2023. "Evaluating California's Accessory Dwelling Unit Reforms: Preliminary Evidence and Lessons for State Governments." *Learning from Land Use Reforms: Housing Outcomes and Regulatory Change*, New York University Furman Center, New York.

McClelland, Robert, and Livia Mucciolo. 2022. "An Update on the Synthetic Control Method as a Tool to Understand State Policy." Urban Institute and Brookings Institution Tax Policy Center, Washington, DC.

McKenzie, David. 2022. "A New Synthesis and Key Lessons from the Recent Difference-in-Differences Literature." *World Bank Blogs*, January 10.

Metcalf, Ben. 2024. "A National Housing Policy Simulator: A Plan for Modeling Policy Changes to Spur New Housing Supply." Policy memo, Federation of American Scientists, Washington, DC.

Nagpal, Geetika, and Sahil Gandhi. 2024. "Scaling Heights: Affordability Implications for Zoning Deregulation in India." Job Market Paper, Population Studies and Training Center, Brown University, Providence, RI.

National League of Cities. 2014. “City Open Data Policies.” National League of Cities, Center for City Solutions and Applied Research, Washington, DC.

Neighbors for More Neighbors. 2021. “Final Update on Minneapolis Built Form Zoning Regulations.” Neighbors for More Neighbors, January 3.

Olken, Benjamin A. 2015. “Promises and Perils of Pre-analysis Plans.” *Journal of Economic Perspectives* 29 (3): 61–80.

Rambachan, Ashesh, and Jonathan Roth. 2023. “A More Credible Approach to Parallel Trends.” *Review of Economic Studies* 90 (5): 2555–91.

Schieferdecker, Alex. 2021. “Five Takeaways from Ten Years of Minneapolis Planning Commission Approvals.” *streets.mn*, February 23.

Seattle Information Technology. 2016. “Open Data Playbook, Version 1.0.” City of Seattle, Washington.

Smith, Stephen. 2023. “Why We Can’t Build Family-Sized Apartments in North America.” *Center for Building in North America Blog*, May 4.

Speckert, Conrad. n.d. “Jurisdictions.” *The Second Egress: Building a Code Change* (website), accessed August 16, 2023. https://secondegress.ca/Jurisdictions.

Stacy, Christina Plerhoples, Chris Davis, Yonah Freemark, Lydia Lo, Graham Macdonald, Vivian Zheng, and Rolf Pendall. 2023. “Land-Use Reforms and Housing Costs: Does Allowing for Increased Density Lead to Greater Affordability?” *Urban Studies* 60 (14): 2919–40.

## About the Author

Tianfang (Tom) Cui is a postdoctoral fellow at the Furman Center, whose research covers topics in public and urban economics. Tom's work combines new measures of regulatory restrictiveness and economic modeling to analyze which policies are able to shift housing supply and demand. As housing affordability becomes a policy priority, he also analyzes how competing local interests form political constraints on housing development. Tom received his PhD in applied economics at the Wharton School and his BA and MA in economics at Northwestern University.

## About the Mercatus Center at George Mason University

The Mercatus Center at George Mason University is the world's premier university source for market-oriented ideas—bridging the gap between academic ideas and real-world problems.

As a university-based research center, the Mercatus Center trains students, conducts research of consequence, and persuasively communicates economic ideas to solve society's most pressing problems and advance knowledge about how markets work to improve people's lives.

Our mission is to generate knowledge and understanding of the institutions that affect the freedom to prosper and to find sustainable solutions that overcome the barriers preventing individuals from living free, prosperous, and peaceful lives.

Since 1980, the Mercatus Center has been a part of George Mason University, located on the Arlington and Fairfax campuses.