

University of Pennsylvania Carey Law School

## Penn Law: Legal Scholarship Repository

---

Faculty Scholarship at Penn Law

---

2-15-2021

### Leases as Forms

David A. Hoffman

*University of Pennsylvania Law School*

Anton Strezhnev

*Harvard University*

Follow this and additional works at: [https://scholarship.law.upenn.edu/faculty\\_scholarship](https://scholarship.law.upenn.edu/faculty_scholarship)



Part of the [Civil Rights and Discrimination Commons](#), [Contracts Commons](#), [Housing Law Commons](#), [Human Geography Commons](#), [Inequality and Stratification Commons](#), [Law and Economics Commons](#), [Law and Race Commons](#), [Law and Society Commons](#), [Other Legal Studies Commons](#), [Place and Environment Commons](#), [Political Economy Commons](#), [Property Law and Real Estate Commons](#), [Race and Ethnicity Commons](#), and the [Real Estate Commons](#)

---

#### Repository Citation

Hoffman, David A. and Strezhnev, Anton, "Leases as Forms" (2021). *Faculty Scholarship at Penn Law*. 2259.

[https://scholarship.law.upenn.edu/faculty\\_scholarship/2259](https://scholarship.law.upenn.edu/faculty_scholarship/2259)

This Article is brought to you for free and open access by Penn Law: Legal Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship at Penn Law by an authorized administrator of Penn Law: Legal Scholarship Repository. For more information, please contact [PennlawIR@law.upenn.edu](mailto:PennlawIR@law.upenn.edu).

## LEASES AS FORMS

*David A. Hoffman*\*      *Anton Strezhnev*<sup>†</sup>

February 15, 2021

### ABSTRACT

We offer the first large scale descriptive study of residential leases, based on a dataset of ~170,000 residential leases filed in support of over ~200,000 Philadelphia eviction proceedings from 2005 through 2019. These leases are highly likely to contain unenforceable terms, and their pro-landlord tilt has increased sharply over time. Matching leases with individual tenant characteristics, we show that unlawful terms are surprisingly likely to be associated with more expensive leaseholds in richer, whiter parts of the city. This result is linked to landlords' growing adoption of shared forms, originally created by non-profit landlord associations, and more recently available online for a nominal fee. Generally, such shared form leases contain worse rules for tenants than the proprietary leases they replace. Over time, it has become easier and cheaper for landlords to adopt such common forms, meaning that access to justice for landlords strips tenants of rights. We observe few within landlord effects: rather, property owners specialize in particular areas in the city. This specialization leads black tenants to be more susceptible to eviction based on crime or drug use on the premises, an effect concentrated in whiter neighborhoods. Our results offer a significant advance in the empirical study of consumer contracting, building the field by examining individual differences in adherents, geography-effects, information costs and time trends.

---

\*Professor of Law, University of Pennsylvania Law School. We thank Andrew Baker, Meirav Furth, Jonah Gelbach, Mitu Gulati, Jon Klick, Rory Van Loo, Julian Nyarko, Manisha Padi, Natasha Sarin, Tess Wilkinson-Ryan and Eyal Zamir, as well as participants at the Fourth Annual Empirical Contracts Conference, a Penn Law faculty workshop, and lawyers from the Philadelphia Public Interest Law Center for comments. Jonathan Pyle of Philadelphia Legal Assistance scraped the underlying eviction dockets and developed the basic coding framework we employ. We are particularly grateful to Mike Crosson (Penn '18) whose preliminary work on these data helped inspire the project. Sean Bender, Madison Gray, Jordan Konnell, Colleen O'Connor, Megan Russo, Patrick Miller-Bartley, and Sam Whillans provided invaluable research assistance. Address comments to [dhoffman@law.upenn.edu](mailto:dhoffman@law.upenn.edu)

<sup>†</sup>CDS-Moore-Sloan Data Science Fellow, New York University.

## TABLE OF CONTENTS

INTRODUCTION	3
I LEASES AS OBJECTS OF STUDY	4
A <i>The Conventional Account: The Landlord's Dominion</i> . . . . .	4
B <i>Leases as Forms</i> . . . . .	6
C <i>Predictions</i> . . . . .	7
II THE LEASES DATASET	7
A <i>Data</i> . . . . .	7
B <i>Unenforceable or Oppressive Terms</i> . . . . .	9
C <i>The Selection Question</i> . . . . .	10
III RESULTS	14
A <i>Summary Statistics</i> . . . . .	14
B <i>Geographic Covariates</i> . . . . .	17
C <i>Individual Effects</i> . . . . .	21
D <i>Templating</i> . . . . .	23
E <i>Behavioral Control</i> . . . . .	33
IV IMPLICATIONS AND RESEARCH DIRECTIONS	35
A <i>Regulating Templates</i> . . . . .	36
B <i>The Role of Race</i> . . . . .	38
C <i>Limitations and Next Steps</i> . . . . .	39
CONCLUSION	39
A APPENDIX: PENNSYLVANIA LAW AND UNLAWFUL AND OPPRES- SIVE LEASE PROVISIONS	43

## INTRODUCTION

Residential leases are ubiquitous, economically significant, but understudied contracts. While more than a third of American households rent, the literature has systematically examined only around a hundred distinct leases. That deficit is particularly striking for contracts whose alleged breach results in evictions, as these agreements provide the basis for the exercise of the force of the state to (quite literally) turn poor families out on the street.

To understand these contracts better, we gather a novel dataset of ~170,000 leases filed in support of over ~200,000 eviction proceedings in Philadelphia from 2005 through 2019. For each lease, we collect information about the leasehold (where it is in space, the size of the property as well as local demographic and income characteristics), the lease (its terms, price and length), the tenant (individual level race, income, gender, and other measures of social capital), the landlord (name, attorney, repeat player status), as well as information about the eviction proceeding itself. We focus on several questions.

*How Bad Are Leases?* The conventional account holds that leases are uniformly averse to tenants. We show, in line with previous research, that unenforceable and oppressive terms in private leases are very common, but leases do differ significantly in how likely they are. Moreover, we find a time trend: the incidence of bad terms has sharply increased over the last 20 years.

*Why Are They Bad?* Legal scholars typically argue that oppressive legal terms result from landlords' bargaining power: powerless tenants accept illegal terms because they cannot do otherwise, though perhaps trading off aversive terms with cheaper rents. But our data reveal a striking, counter-intuitive, fact: unenforceable terms are more likely in the leases of richer, whiter, tenants. We explain this result by focusing on the role of standardized leases that are widely shared between landlords. Though we study over a hundred thousand private contracts, a handful of distinct standard form leases make-up the plurality the dataset. Such *shared leases* are more likely to be used in more expensive leaseholds, and also more likely to contain unenforceable terms. We show that the sharp increase in unlawful terms in the last twenty years can be thus attributed to increasing adoption by small landlords of shared leases, in turn resulting from the ability to access cheap internet forms, which replace the oral or *proprietary leases* of the past. In this way, we illustrate standardization's perverse effects.

*Are Eviction Leases Different?* One hazard in our approach concerns the selected nature of our dataset. We study those leases that landlords have attached to eviction filings, and not a random selection of Philadelphians' leases: for instance, they are disproportionately drawn from the leases of the poor. Even within the set of poor tenants, eviction isn't random: tenant characteristics, landlord strategies, and neighborhood characteristics may make it more likely. We investigate whether there is a connection between these *eviction selection* effects and *lease selection*. We are primarily concerned with knowing if the

sorts of leases attached to eviction filings are meaningfully different from those of poor Philadelphians generally. Attacking that problem from a variety of angles, we argue that they are not.

*What is the Effect of Local Geography?* In Philadelphia, the lease you sign has more to do with where you live than who you are. Because landlords specialize in places, and almost never vary leases between tenants, geography generally dominates tenant characteristics in explaining the makeup of leases. That tight link between lease, place and people significantly complicates the conventional account, which suggests that leases are priced in part based on their legal terms.

*Do Tenant Characteristics Matter?* In a world dominated by widely shared forms, it should be unsurprising that we find almost no within-landlord differences. Simply put, landlords do not tailor leases for tenants. However, we do observe that some landlords appear to specialize based on tenants' race, and, when doing so, use distinctive lease strategies. The result is that black tenants living in white neighborhoods are more likely (than white tenants) to be asked to sign leases that permit landlords to evict them if anyone uses drugs on the premises or commits a crime. As this request is not made in minority-majority tracts, it has the effect of enabling race-based private policing of tenants.

These observations—the dominance and growth of illegal terms, the role of forms, and the pernicious effects of geography and race — lead naturally to set of prescriptions, which we discuss in the last section of the paper.

We begin with a literature review.

## I. LEASES AS OBJECTS OF STUDY

As Suchman (2003) argued, contracts are “social artifacts,” things worthy of study as such. Much can be learned from micro and macro analysis of the production of these objects, just as scholars have benefits from the study of other technological innovations, from the wheel to the mousetrap. The study of lease contracts as artifacts has been, to date, limited. In this section we detail the results of that previous learning, and then describe how it relates to the burgeoning field of contractual innovation writ large.

### A. *The Conventional Account: The Landlord's Dominion*

There have been a handful of small-scale empirical projects studying leases. Curtis Berger's pioneering work on leases focused on “representative forms” from 16 cities' Real Estate Boards in 1972, out of 50 cities surveyed. Berger (1974) found that of the 16 model leases, around half purported to waive the warranties of habitability, while most created additional landlord remedies and tenant obligations. The result was a set of documents that Berger argued treated the “residential tenant as a latter-day serf. One sees a near-pathological concern with tenant duties and landlord remedies, occupying from 50 to 80

percent of every form. Much of the remaining text seeks to immunize the landlord against the claims of his tenant.”

This theme—that leases permit landlords to dominate tenants—continued in later works. Most argue that form leases are an faceless evil, as “it makes little difference which form one purchases, however, for their contents are strikingly uniform.” (Bentley 1974) Scholars have asked why forms persistently contain enforceable provisions, as such terms are often denied enforcement by courts. (Sullivan 2009; Berger 1974; Furth-Matzkin 2017) The general consensus is that such clauses are imposed as a part of a strategy by landlords, who seek to extract surplus from tenants given tenants’ low readings rates and inability to fully police such clauses in court. (Furth-Matzkin 2018; Mueller 1970)

The work closest to ours is Furth-Matzkin’s work on Boston leases. (Furth-Matzkin 2017). Employing a snowball sample, she studied 70 residential agreements, of which a third were student leases. Of those leases, a little more than half had individual (non-corporate) landlords. Half of the studied leases were standardized forms. Furth-Matzkin identified a relatively high number of unenforceable and misleading terms. When controlling for leasehold size, rental amount, tenant type, and duration, she concluded that corporate landlords, using commercial forms, and higher value leaseholds, all used forms with fewer unenforceable terms. Furth-Matzkin attributes her results to commercial firms and drafters being “typically more careful than individual landlords in the drafting of their leases.”<sup>1</sup>

In general terms, these studies extend a consumer law tradition which focuses on the purportedly negative effects of disparate bargaining power. Whiter and richer tenants, and

---

<sup>1</sup>In an unpublished student paper, Nora Crawford studied 20 Philadelphia leases, 17 from students and 3 from landlords. She found that 6 of the 17 student leases used language drafted by a single law firm in Philadelphia, which listed its copyright on the leases themselves. (Crawford N.D.) Crawford interviewed the lawyer-drafter of the form residential leases in her sample. He explained why he thought terms in leases were particularly sticky: landlords “look at [the lease] like the Bible, and they will not change the terms of the Bible.” The drafting lawyer continued that he had departed from the model Realtor Association lease because it had drawn attention to tenant remedies in ways not required by law: “My clients don’t want me to tell the other side how to sue us and when.” He suggested that because Pennsylvania law on landlord tenant issues is “sort of confusing,” landlords seeking to comply face difficulties. For “mom and pop” operators, the drafter believed that they originally “buy a form” but then do not update it for later developments in the law. By contrast, larger landlords were more likely to “revise [the lease] periodically, but not regularly, to reflect changes in the law.” Finally, the drafter discussed how his leases were disseminated in the housing market. His firm’s standard residential lease is “very inexpensive” and gives the purchaser an unlimited license at their properties. Customized leases are more expensive, and most landlords take the standard form. He described that form as a loss-leader—the “low-end” of the practice—intended to build relationships with clients who might later be interested in higher end business. Unfortunately, sometimes property managers “take” the lease from building to building as they move, or the client takes the lease to new properties without his notice. The result is that his leases, originally drafted to build a book of business for the law firm, have diffused through the residential market, particularly in buildings associated with richer landlords.

men, are thought to be signatories of “better” leases, or at least leases with fewer unenforceable terms. Moreover, the literature seems to largely assume that landlords calibrate price based on the characteristics of the applicable legal regime and the tenant. Scholars thus posit that landlords will charge tenants for “better” lease contracts (Ackerman 1971).

### *B. Leases as Forms*

The above account suggests that leases are tailored for tenants. But the literature on contracts-as-forms is to the contrary. Multiple studies have found that even for material contracts, terms are not well priced, nor do they change after radical shifts in the governing legal regime (Gulati and Scott 2012; Choi and Gulati 2004). How can this be? The answer seems to be a very strong form of inertia and aversion to change of the template (Nyarko 2020).

This in turn lends credence to drafting cost explanations for boilerplate terms. Professor John Coates, describing weak takeover defenses, argued that information costs played an important role (Coates 2001). Where terms were less material, lawyers would be likely to adopt those nearest to hand, meaning that contact boilerplate terms were “likely to exhibit geographic correlations.”

Law firms that are closer to each other in physical proximity are more likely to share information, either formally (by sharing documents or experiences), or by lateral hiring, or by conscious borrowing from large, locally prominent law firms with high IPO market shares (Silicon Valley law firms are likely to look to Wilson Sonsini; smaller New York law firms are likely to look to Skadden Arps or Sullivan & Cromwell) or via common counter-parties (particularly accountants, investment bankers, or VCs). In any event, more geographically proximate law firms are more likely to think of one another as salient sources of public company boilerplate.

Coates continued that once established, economic concentrations on the client side would be nurtured by parallel, close-knit, networks of lawyers, who would have an interest in preserving similar forms and practices (at the expense of outsiders). The result, in Coates’ sample, was a “silicon valley law firm effect.” Other scholars have found similar differences in economically significant contract practices based on law-firm fixed effects (e.g. Choi, Gulati, and Posner 2012; Jennejohn 2018).

The consumer literature is less well developed, largely because there are fewer large-*n* datasets of consumer facing contracts. Florencia Marotta-Wurgler has shown that firms adjust the terms in their consumer contracts as they learn about their consumers (Marotta-Wurgler and Taylor 2013). But, though terms may change over the long run, there is very little evidence that consumers react to those changes as they happen. Reading rates

for mass contracts are too low to generate informed minority price responses to terms. Outside of a few very salient terms (warranty, interest rate and term in financial contracts), we are aware of no studies that find price effects for particular ordinary contract terms. (Zamir 2014)<sup>2</sup>

### C. Predictions

Putting this work together leads to cross-cutting intuitions. The extant evidence on leases focuses on unenforceable terms, and finds that landlords use “worse” leases for the least powerful tenants. It also holds that there are few, if any variations between forms (rarely, in fact, mentioning the possibility of competition between lawyers). Finally, the evidence is blind to the effects of local geography. But these conclusions rely on very few observations.

Work on contracts-as-forms, by contrast, suggests that we should expect variation between forms to be driven in part by drafting costs, and that such costs might result in geographically based differences in form types. It also suggests that there will be little change in forms over time, and that unenforceable terms may persist, but the explanation may turn more on the costs of change rather than an intent to dominate. This strand of literature, in short, does not necessarily predict that weaker tenants will get worse leases.

## II. THE LEASES DATASET

### A. Data

We obtained our dataset of leases with the help of the non-profit Philadelphia Legal Assistance (PLA), an organization that provides free civil legal services to low-income Philadelphians, as a part of a broader project analyzing eviction proceedings that scraped the online docket of the Philadelphia Landlord-Tenant Court. Each landlord-tenant case contains a standardized complaint document along with a number of attached exhibits—in PDF format—containing any supplementary documents. From available dockets, PLA downloaded all such exhibits with titles that contained any mention of the term “lease.” This yielded a total of 225,409 potential leases that justified evictions from 2005 to 2019.

These PDFs, being images of scanned paper documents, are neither machine-readable nor searchable in their original form. With difficulty, we converted them to text.<sup>3</sup>After

---

<sup>2</sup>Studies of mortgage contracts (and credit-card agreements) do observe behavioral changes in response to changes in salient material terms (Bucks and Pence 2008; Padi 2018).

<sup>3</sup>To analyze the data in any reasonable timeframe, it was first necessary to convert the images into a machine-readable text format. We used the open-source Tesseract 4.0 Optical Character Recognition (OCR) engine to process the PDF images into text files. Often leases are scanned at poor resolutions, at difficult-to-read angles and with hand-written text and therefore the image conversions typically contain errors. Even in



converting, we constructed a search via a number of fuzzy regular expressions—search patterns with some tolerance for mistakes or errors—to detect whether a document actually contained a lease and on which page the lease began. Of the 225,409 docket entries, we identified 171,915 as residential leases for which data on the plaintiff and defendant were present and where we were able to detect the presence of a lease via OCR.<sup>4</sup> We removed 609 leases which had missing data on the ongoing rent. Overall, we have captured the vast majority of written leases underlying Philadelphia landlord-tenant disputes from 2005 to 2019.

In this sample of 171,306 residential leases, we searched for whether the dispute related to a public housing or subsidized unit. While leases involving public housing were directly identifiable by the presence of the Philadelphia Housing Authority (PHA) as the plaintiff, a substantial share of lease contracts with private landlords involved subsidies to the tenant administered through the HUD housing choice voucher program. Public housing and subsidized leases are subject to significant regulation in what contract terms are permitted, with both exculpatory clauses and waivers of notice of legal proceedings explicitly forbidden.<sup>5</sup> We search each of the leases for any mentions of “housing assistance” or “assistance payment,” indicating that the lease is connected to a housing assistance payment contract between the tenant and the public housing authority. Of our 171,306 leases, we identified 18,030 as public housing leases where the PHA acts as the plaintiff and 16,214 subsidized leases, leaving 137,062 unsubsidized lease contracts.

From the docket entries themselves, PLA extracted a number of covariates related to the dispute and the property. Most important for us is the address of the property, which PLA obtained from each Landlord-Tenant complaint and attempted to geocode. 164,856 leases were able to be geocoded to a unique latitude/longitude and located within a Philadelphia census tract. Of these, 132,515 are unsubsidized leases.

Beyond property-level covariates, we obtain information on the name of the plaintiff (landlord), the outcome of the case (win, loss, or withdrawal/judgment by agreement) and in the case of a judgment, whether it is a default judgment, the amount of ongoing rent demanded (which is almost always equal to the monthly rent specified in the lease), and whether the tenant had legal representation.

We also collect, from a private firm, information about race, gender and voting history along with consumer data where available for 35,735 individual tenants in cases from 2014

---

high-quality images, the OCR process does not necessarily detect paragraphs and segmentation in a reliable fashion, making it difficult to further segment the text into distinct provisions. Additionally, leases are often included with other documents such as any notices sent to the tenant or the landlord’s rental license.

<sup>4</sup>Not all scraped documents actually contained a lease. This is likely because the PLA search was somewhat overly-broad and included entries such as “Notices of Lease Termination.” Other failures to detect may be actual errors due to poor-quality OCR. Given the share of leases actually captured, we do not think this is a substantial issue.

<sup>5</sup>See 24 CFR § 886.327

to 2019. We use this information to help to validate a name-matching technique to estimate the tenant's race in the leases.

### *B. Unenforceable or Oppressive Terms*

We begin by focusing on the unenforceable or otherwise oppressive terms that they contain.<sup>6</sup> By talking to advocates in Philadelphia, we identified a number of potentially troublesome terms that had sufficiently clear precedent to draw conclusions from, and consistently enough written to code.<sup>7</sup> We think that three are probably or certainly unenforceable under applicable Pennsylvania law. One is enforceable but signals the landlord's power and perhaps its views on the likelihood of court enforcement. Appendix A describes the relevant caselaw in detail.

1. *Exculpatory clauses*, which disclaim liability for negligence. Such clauses, as we describe in the Appendix, are nearly always unenforceable in residential leases.
2. *As-Is clauses*, which purport to waive the implied warranty of habitability. These too are usually unenforceable.
3. *Holdover tenant penalty clauses*, which state that tenants who do not leave after the expiration of the lease term owe an increased sum, typically a multiple of the rent. These clauses are probably unenforceable in most of the leases we observe.
4. *Waivers of notice*, which permit the landlord to begin an eviction proceeding before the 15-30 day notice period required by the Landlord-Tenant Act. These clauses are likely permitted by law, but they are oppressive

Our strategy for identifying terms in leases relies on a series of searches using approximate (or “fuzzy”) regular expressions. These are search patterns which we crafted to identify the key phrases that reflect the presence of a particular provision.<sup>8</sup> We use a number of different search expressions for each clause to reflect the slight differences in language and phrasing found. While we generally find relatively good performance in identifying the specific provisions we focus on, this classification method is imperfect and

---

<sup>6</sup>Our first intuition was to compare lease similarity using existing supervised learning or other forms of matching tools. Because of the difficulty of parsing individual paragraphs in the leases and the prevalence of typographical errors from the transfer from PDF to text, that approach proved impracticable.

<sup>7</sup>In deciding which terms to focus on, we were confronted with a regional limitation in legal certainty. Though the field is nominally jointly governed by federal statutes and Pennsylvania's Landlord and Tenant Act of 1951, neither provides clear guidance about most provisions in modern leases. Caselaw about residential leases is sparse, leaving many provisions in a limbo, where analogies must be drawn from from decisions about commercial tenants, or more general principles of contract law.

<sup>8</sup>Specifically, we use the implementation found in the `regex` package in Python 3.

may not correctly classify some leases (such as cases where the language differs from what is typically used or the lease scan was incredibly poor quality).<sup>9</sup>

### C. *The Selection Question*

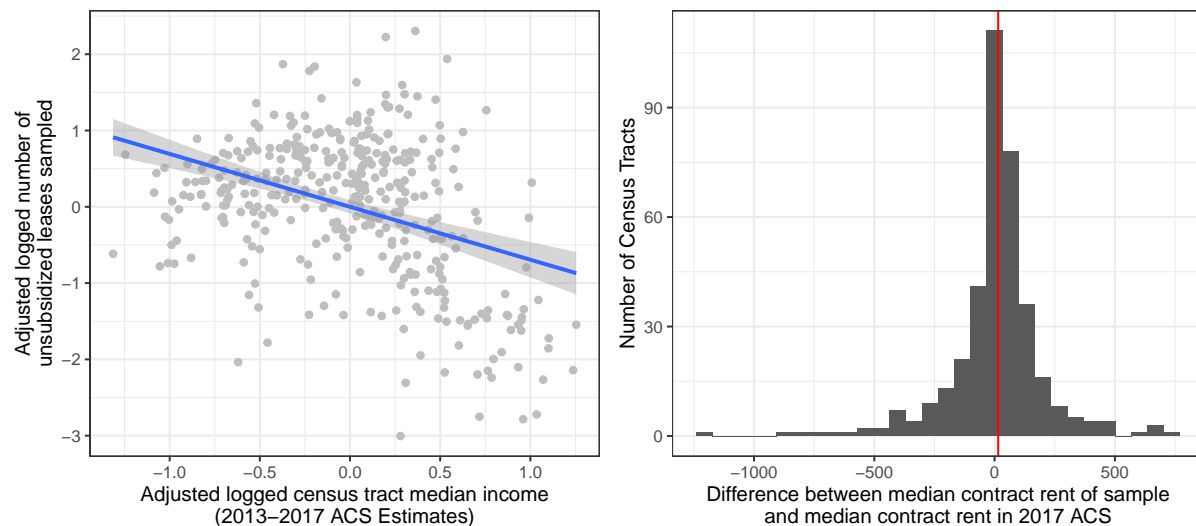
Leases are filed as exhibits to eviction proceedings and are not a random sample of all leases in the Philadelphia, let alone all national leases. While we do not claim to speak to the sorts of patterns that exist among Philadelphia leases in general, we do consider our sample to be a reasonable starting point for examining rental contracts among the poorer residents of Philadelphia. Nevertheless, we consider two possible avenues through which focusing on leases attached to evictions could generate a form of “selection bias.” We examine whether such mechanisms are plausible given additional evidence and, if they are, how our analyses adjust for that bias. First, we consider the question of whether selection threatens “external validity” – that is, whether our sample can reasonably generalize to the population of low-income resident leases or whether there are certain landlords or types of leases that never appear or are grossly underrepresented in our sample simply because those landlords are able to encourage tenants to leave their properties without recourse to formal rules. Second, we consider the risk of spurious correlations induced by unobserved factors associated with lease provisions and the onset of eviction proceedings – the “collider bias” problem (Cole et al. 2010).

Both external validity and collider bias concerns arise because we are only able to observe those leases for which an eviction proceeding was initiated and cannot observe those landlords who, for whatever reason, never choose to use the landlord-tenant court. Recent literature on the hazards that appear to drive eviction point to a number of factors that could also potentially be associated with covariates of interest. In recent work, Desmond and his co-authors randomly surveyed Milwaukee residents and find that (net of other factors) having children increased the likelihood of eviction, as does increase in crime in the neighborhood. Desmond has also found that being Latino in a predominantly white neighborhood makes one particularly at risk for eviction (from among the set of individuals who are behind on their rent) (Desmond and Gershenson 2017). At the same time, tenants with weaker social networks — close relationships who had recently experienced drug addiction, incarceration, eviction, etc. — were more likely to be evicted (Humphries et al. 2019). If the people in landlord tenant court are distinctive, it might stand to reason that their leases are too.

We first focus on the question of representativeness. Do we have reason to believe that the sample of leases we collected systematically differs from the true population of

---

<sup>9</sup>We do not believe that these errors are systematically associated with any fundamental characteristics of the underlying lease and eviction proceeding, but note that our approach will tend to be more likely to fail to classify a lease with a particular provision (false negative) than yield a classification where there is no such provision (false positive).



*Notes:* N = 375 census tracts. Grey bands denote 95% HC2 robust confidence intervals. Left panel presents an added-value plot from a regression of log number of leases sampled on log median income adjusting for log number of renter-occupied units. Right panel plots a histogram of the distribution across census tracts of the differences in median contract rent from sampled leases and median contract rent from ACS. Vertical red line denotes the median of the distribution. Median household income reported in 2017 inflation-adjusted dollars (ACS table S1901). Number of renter-occupied housing units based on 2013-2017 estimates (ACS table DP04). Median contract rent calculated from renter-occupied units paying cash rent (ACS table B25058).

Figure 1: Representativeness of sampled leases compared to ACS benchmarks

lower-income leases? We do find that, after conditioning on the number of renter-occupied units in each census tract (using data from the American Community Survey), we are more likely to sample leases from lower- and middle-income census tracts than we are leases from higher-income tracts. The left panel of Figure 1 shows the added-variable plot from a regression of the (logged) number of leases sampled in a census tract on the tract's (logged) median income, adjusting also for the (logged) number of units in a tract. This is consistent with the general pattern that the poor tend to be the targets of eviction proceedings and therefore appear in our sample (Collinson and Reed 2018). Since this paper specifically focuses on the contracts of poor residents, this is entirely expected – areas of Philadelphia are highly segregated by income and it would be surprising if a sample of the leases of lower-income residents uniformly represented all areas of the city. We would be somewhat more concerned about representativeness if we found that *within* each census tract we were consistently sampling leases that looked very different from the typical rental agreement in that tract. We do not find evidence that this is the case.

We examine the gap between the median monthly rent among sampled leases within each census tract and the median rent reported by the ACS. As shown in the right panel of Figure 1, the gap, on average, is very close to zero. Apart from a few extreme outliers – likely driven by single apartment complexes in otherwise low-eviction areas that are heavy initiators of eviction proceedings – we see little systematic difference between the median rent statistics from the ACS and median rent in our data.<sup>10</sup> This suggests the absence of *systematic* biases in the within-tract sampling process at least insofar as they would be correlated with rents. Since rent is an obviously highly relevant covariate when looking at leases, it is difficult to come up with a selection mechanism that is not, in some way, associated with rents. We are relatively confident that the main factors driving selection into our sample are geographic-level factors related to resident income which are expected and can be observed to a large extent.

Further, the structure of the lease ecosystem strongly suggests that it is unlikely that there is a type of lease that is entirely absent from the eviction sample. As we explore below, the dataset is dominated by shared leases – we can link around half of all leases in the dataset to shared, commercially available forms. These shared leases contain the worst clauses for tenants, and are increasing in relative prevalence over time. Within landlord differences are almost entirely absent. Thus, landlords (who evict) use the same forms for all of their tenants, and they increasingly tend to use forms also adopted by other landlords. One would have to posit that such landlords use an entirely different set of forms for that portion of their tenant population that they rationally expect *not to later evict* to worry that the selected leases and forms are unrepresentative *of those landlords*.

Moreover, while it is possible that the leases of landlords who don't file as many evictions are different, we discount that possibility for several reasons. First, Furth-Matzkin's work observes roughly equivalent numbers of unenforceable terms in her sample, composed largely of student leases, as we do. Furth-Matzkin (2017) This provides some comfort that the selection process is not biased toward finding "worse" leases, even though these are very documents that are submitted to court. In that light, it would be quite surprising if the set of landlords who don't evict had *better* terms for their tenants than those who do. Second, the prevalence of the notice waiver provisions suggest that selection out of litigation is somewhat limited: landlords can file evictions with relatively little fuss, and often do so supported by attorneys whose practice is highly routine. In fact, a recent lawsuit filed accuses those very attorneys of running, in effect, eviction-mills where insufficient attention was paid to the facts of each case.<sup>11</sup> Overall, as that eviction appears to result largely from post-rental change in tenant characteristics (such as having additional children while in the leasehold) we think that pre-eviction lease selection is unlikely.

---

<sup>10</sup>While there is quite a bit of variability, this is expected given the nature of the sampling process.

<sup>11</sup>See Lawsuit Documents, available at <https://www.pubintlaw.org/cases-and-projects/ending-exploitative-eviction-practices-by-collection-lawyers/>

In short, while it is obvious that examining leases associated with eviction will not yield a representative sample of all leases in the city, it does give us a good window into the the sorts of leases held by tenants on the lower end of the socio-economic spectrum that are more at-risk for eviction.

A second and more troubling concern about selection comes from the risk of “collider bias” or spurious correlation between variables induced by the fact that we are only examining leases associated with evictions. Many of the individual-level and demographic-level covariates that we study are known to have a strong association with the probability of eviction proceedings. For example, as recent work by the Reinvestment Fund suggests, Black tenants in Philadelphia are significantly over-represented in evictions relative to their share of the overall renter population (Reinvestment Fund 2021). Moreover, they are more over-represented specifically in those census tracts that are gentrifying and undergoing significant growth in rents over time. If there are other unobserved factors associated with both eviction and lease provisions, that may explain some of the associations we discuss below.

This would make an interesting explanation in its own right for some of the descriptive patterns in lease provisions that we observe and we make no claim of an explicitly causal association for the relationship between, for example, tenant race and lease provisions. One hypothesis is that this relationship may be driven by a third, unobserved factor, related to differences in property type between Black and white tenants within census tracts, but an alternative explanation may be due to gentrification patterns that are not being captured by our covariates. If landlords in rapidly gentrifying areas are more likely to adopt shared form leases and, on average, more likely to evict Black tenants, we would observe Black tenants *within the subsample of eviction-related leases* as being more likely to have provisions associated with shared forms. Such a pattern would nevertheless have important substantive consequences for our understanding of the eviction process.

However, we have some reason to doubt that this is necessarily the only story. First, we adjust for census tract-level covariates. If it is the case that Black tenants are disproportionately facing eviction in gentrifying areas of the city, then comparing Black residents to white residents in comparable regions would rule out gentrification as an explanation for the observed difference in our outcomes of interest. While our primary tract-level covariates are time-invariant, we include interactions between them and the filing date of the case in the regression analyses. Although we do not have an ideal *time-varying* measure of gentrification that encompasses our entire timeframe, this approach provides an initial way to capturing recent changes in the relationship between geography and eviction. Second, we find no evidence of heterogeneity across time in the association between whether a tenant is Black and the presence of unenforceable lease provisions like the notice waiver or exculpatory clause. We would expect that if the correlation were driven by a surge in evictions of Black tenants from gentrifying areas in recent years that this relationship would

grow stronger over time. Ultimately, while we are certainly cautious to not draw extremely strong conclusions about all leases based exclusively on the subset of leases from eviction proceedings, we have taken care to address the most likely sources of selection-induced correlation using what data can be gathered on census tract characteristics.

### III. RESULTS

#### A. Summary Statistics

Our geocoded sample consists of 164,856 total lease documents with 132,515 of these leases identified as unsubsidized and private. Based on the names of plaintiffs extracted from the landlord-tenant complaints, we find 42,683 unique plaintiffs-landlords, 39,701 of which have at least one unsubsidized lease. The modal number of times a unique plaintiff appears in the data is 1, so the share of repeat landlords is quite small. However, we do find 1,603 unique landlords with at least 10 landlord-tenant cases involving unsubsidized leases, allowing us to partially evaluate the extent to which there exists any meaningful within-landlord variation in lease provisions.

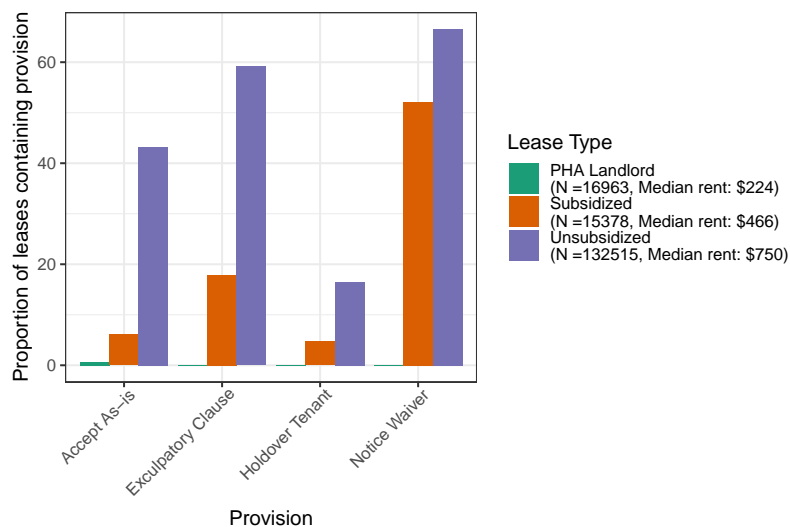
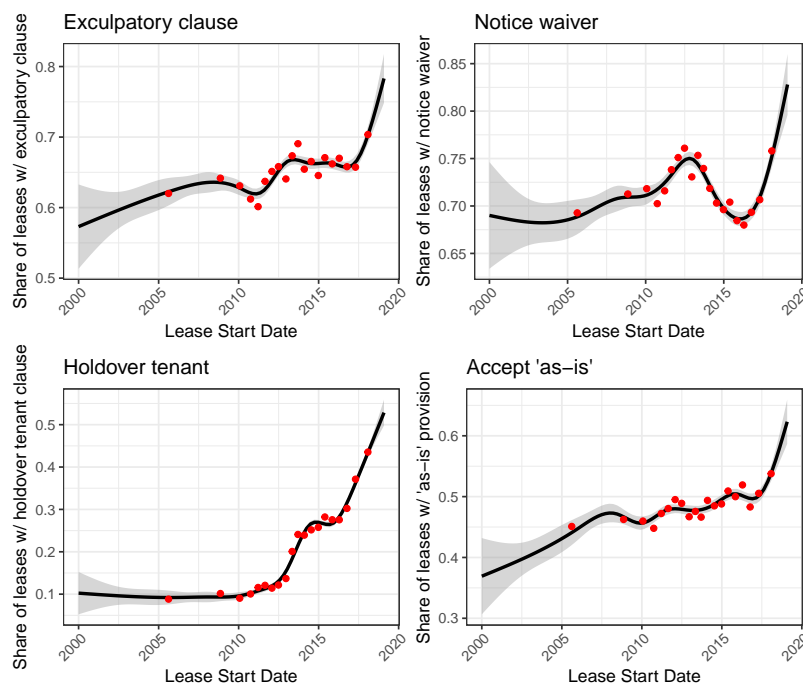


Figure 2: Summary of lease term prevalence in sample (all geocoded leases)

Figure 2 summarizes the overall frequency of each of the four provisions within the three classes of leases we identified. It also indicates the total number and median ongoing rent for each category as listed in the complaint.<sup>12</sup> Among the unsubsidized leases,

<sup>12</sup>Since rent is often handwritten into a lease template, it is not possible for us to systematically extract monthly rent from the machine-readable text as it is rarely OCR'd without error. As such, the information

the prevalence of exculpatory clauses and notice waivers is quite high: about 60 percent. A slightly larger share (67%) contain an explicit waiver of notice or a reduction in notice terms. Holdover tenant clauses are comparatively rarer and likely a newer phenomenon, with only 12% of unsubsidized leases containing this provision. “As-is” clauses are less common than either exculpatory clauses or waivers of notice, but more frequent than holdover tenant provisions. As expected, the Philadelphia Housing Authority leases do not contain any unenforceable terms.<sup>13</sup> These provisions also tend to co-occur. Of the unsubsidized leases in our sample, 18% have one of the provisions, 20% have two, 31% have three and 10% have all four.



*Notes:* N = 88,921 leases with observed start date post-2000. Estimates denote predicted probabilities from a cubic smoothing spline. Red points denote binned averages.

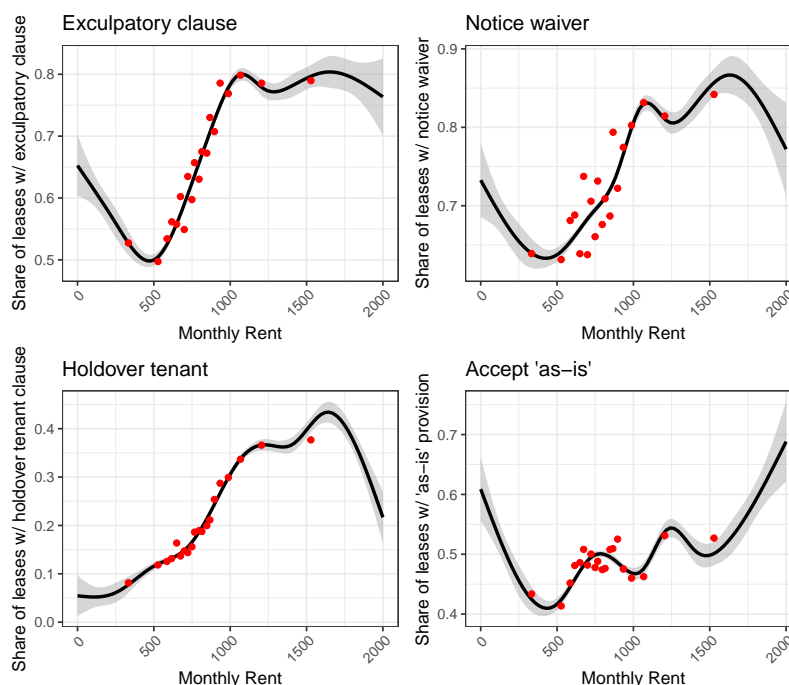
Figure 3: Prevalence of lease provisions by lease start date

from the landlord-tenant complaint on the amount of “ongoing rent” demanded by the plaintiff provides us with the closest approximation. Landlords have incentives to inflate this figure.

<sup>13</sup>We find a tiny fraction of matches to our search queries among the PHA leases which are likely false positives due to chance error in the OCR transfer. While we find a larger number of exculpatory clauses among the subsidized leases, this is not necessarily an indicator of a high false positive rate as it is not uncommon for a HUD-approved subsidized lease agreement to be accompanied by a standard lease template that may contain unenforceable provisions – both of which are signed by the landlord and tenant.



Leases have become much more unfriendly to tenants over time. Figure 3 plots the estimated proportion of leases with each of the four provisions over time. We use the lease start date listed in the docket and focus on leases starting from 2000 onwards as the sample size of leases that are listed as beginning prior to this year becomes very small.<sup>14</sup> For some clauses, the growth is stark: holdover tenant clauses were essentially unknown in 2000, but after a dramatic shift in 2010, they now are in most leases.



Notes: N = 86,880 leases with observed start date post-2000. Estimates denote predicted probabilities from a cubic smoothing spline.

Figure 4: Prevalence of lease provisions by monthly rent

Next, we examine the correlation between the monthly rent listed in the lease and the presence of these provisions among the unsubsidized leases.<sup>15</sup> Figure 4 plots the relationship between monthly rent and the prevalence of each type of provision. Focusing

<sup>14</sup>This data is missing for many docket entries and due to typographical errors, some leases are listed with nonsensical start dates. We are able to prune nearly all of these incorrect start dates by excluding leases that list a start date after the date of the filing of the dispute. This leaves us with a total of 92,111 unsubsidized leases post-2000 for which we know the start date and 88,921 which we are also able to geolocate.

<sup>15</sup>We focus on those leases with monthly rents listed as greater than 0 and less than or equal to 2000 dollars per month, which accounts for over 98% of leases in the dataset. Some of the very extreme values (>\$10,000 per month) are likely typographical errors, while others may be mis-classified commercial prop-

specifically on the range of monthly rents between \$500 and \$1500, which covers about 91 percent of the observations in the data, there is a clear, positive linear trend in the prevalence of each of these four provisions.<sup>16</sup> Notably these relationships persist even when adjusting for lease start year fixed effects to account for the fact that rents are generally rising over time.

These are striking patterns. Contrary to previous research, unenforceable and oppressive terms appear to be *positively* correlated with wealth, at least as measured by ongoing rent obligations in eviction leases. But is this relationship primarily driven by differences in the location of the properties or is there something about the individual characteristics of the tenants that predicts whether a lease will contain one of these terms? The next section examines the spatial distribution and the extent to which neighborhood characteristics predict term prevalence.

### B. Geographic Covariates

Given that lease terms are rarely the subject of direct negotiation between the landlord and tenant, and landlords typically use standardized form contracts, we expect that a significant amount of the variation in term prevalence will be driven by neighborhood-level characteristics. Over the 42,683 landlords we identify, around 30% (12,709) have two or more than two unsubsidized leases: 75% hold only a single leasehold.<sup>17</sup> Only 6,963 have properties across two or more census tracts and a mere 911 have properties across five or more tracts. That means that landlords, in effect, specialize by tenant type, as tenants cluster by wealth, class, race and ethnicity.

Philadelphia exhibits significant variation across neighborhood in wealth and income level and is likewise highly segregated across racial lines. We obtain from the American Community Survey 5-year data estimates of median income and racial composition in 2013-2017 at the level of the census tract. This period covers the signing dates of most of the leases that we have in our dataset. Both of these are highly correlated with one another, with poorer tracts tending to also have a larger proportion of Black residents. In general, the wealthiest and whitest areas of the city are located near the downtown area, and the northwest. In contrast, poorer tracts are found primarily in the north and west of the city.

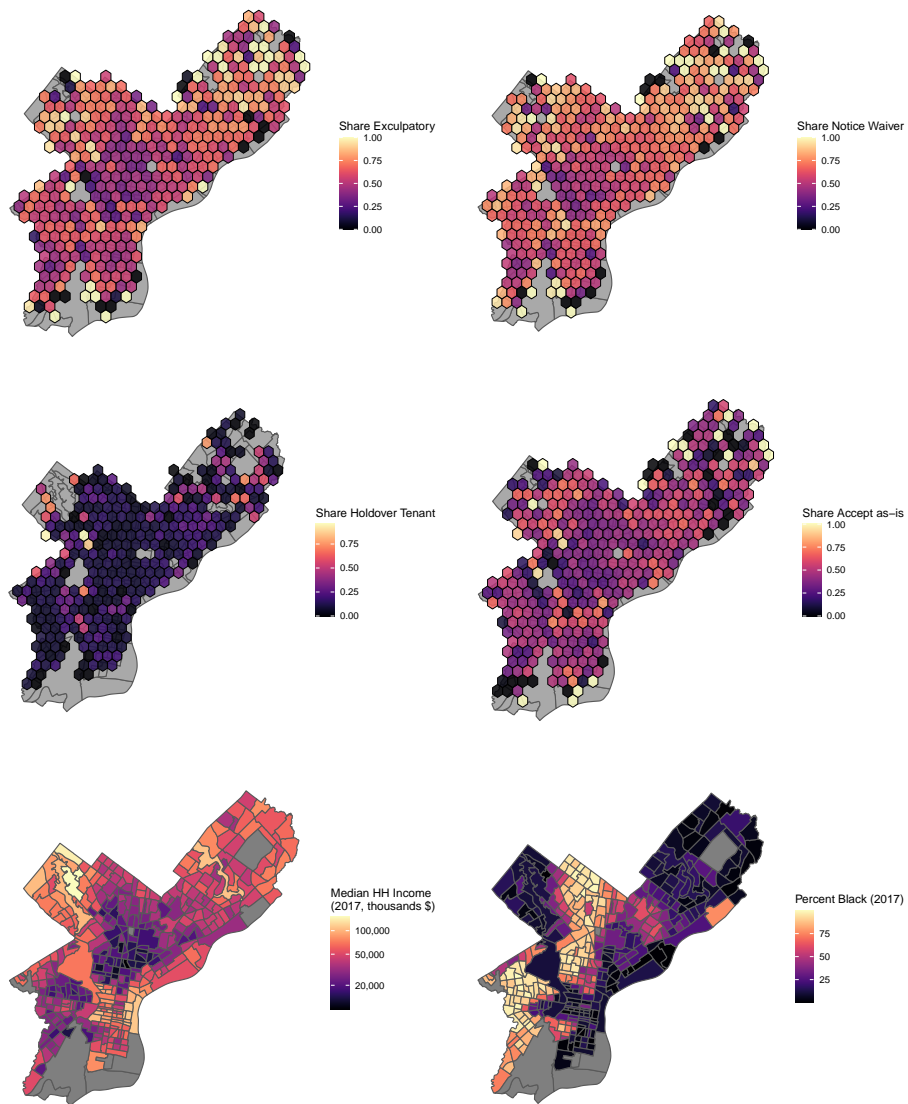
Simply plotting the prevalence of each of the provisions among our sample of unsubsidized leases strongly suggests regional effects. Figure 5 overlays on top of a map of

---

erties. Overall, pruning the data to these units this leaves us with 129,256 unsubsidized, geolocated leases. 86,880 have a known start date after 2000.

<sup>16</sup>We fit a cubic spline regression to estimate the share of leases containing each of the four provisions conditional on monthly rent. Closer to the tails of the distribution, there is a bit more instability in the estimated CEF due to the sparsity of observations near \$0 and greater than \$1500 and the flexibility of the model chosen.

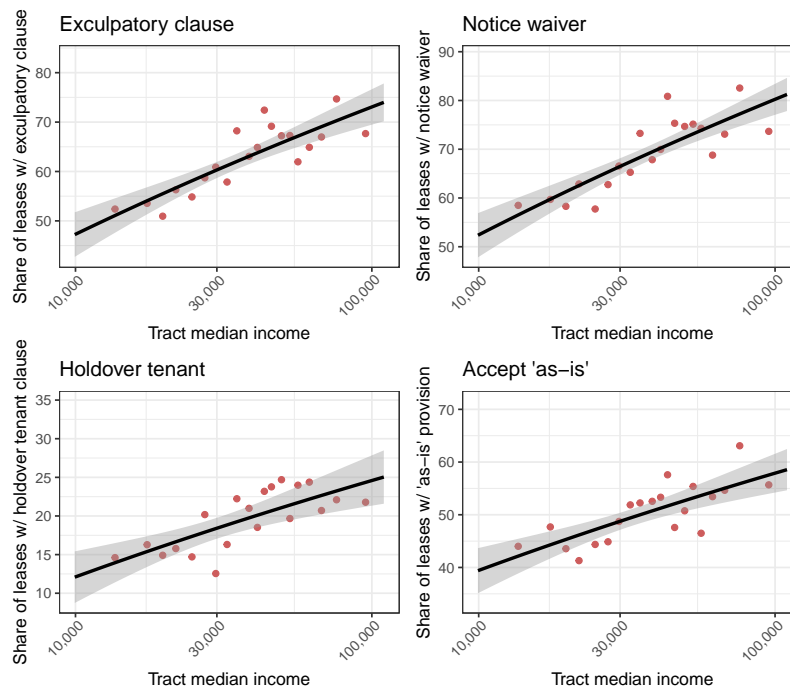
<sup>17</sup>We do not directly observe management companies.



*Notes:* Unsubsidized, geocoded leases with non-missing start date after 2000 (N = 88,921). Hexes are approximately 1 square kilometer in area. The bottom two figures are (left) median household income reported in 2017 inflation-adjusted dollars (ACS 5-year table S1901) and (right) the share of residents reporting their race as one race and Black or African American (ACS 5-year table DP05). Tracts with missing data in grey.

Figure 5: Geographic distribution of lease provisions in Philadelphia, Median Income (2017), and Race (2017)

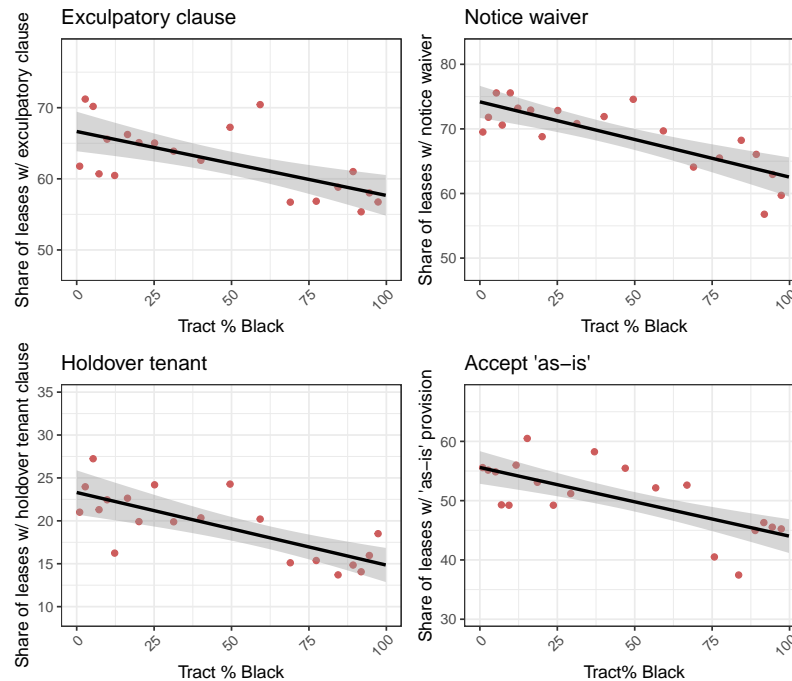
Philadelphia, a series of one-square-kilometer sized hexes. Within each hex, we calculate the proportion of geocoded leases that contain each of our four unenforceable lease provisions. Exculpatory clauses, notice waivers, holdover tenant clauses and accept as-is terms tending to be more common in the northeast, northwest and downtown regions and generally less common in north and west Philadelphia.



*Notes:* N = 363 census tracts with 10 or more sampled leases. Grey bands denote 95% HC2 robust confidence intervals. Median household income reported in 2017 inflation-adjusted dollars (ACS 5-year table S1901).

Figure 6: Regression of share of lease with each provision in a census tract on logged median income (2017)

To better visualize the relationship between local conditions and unenforceable provisions, we regress the proportion of each lease provision among leases geolocated to a census tract on 2017 median household income and on percent Black residents. Figure 6 shows that the relationship between census tract log median income is roughly linear and positive. In regions of the city with greater income levels, the proportion of leases in our sample that contain one of the unenforceable provisions is notably higher than in those regions with comparatively poorer residents. Figure 7 recovers an analogous nega-



*Notes:* N = 363 census tracts with 10 or more sampled leases. Grey bands denote 95% HC2 robust confidence intervals. Percent Black is the share of residents reporting their race as one race and Black or African American (2017 ACS 5-year table DP05). Points denote binned averages (20 equally-spaced bins). All regression slope coefficients statistically significant at  $p < .05$ .

Figure 7: Regression of share of lease with each provision in a census tract on % Black residents

tive association with the proportion of Black residents in the census tract.<sup>18</sup> While there is a substantial amount of noise in how prevalent a given contract provision is in a region, all of the provisions we consider are more prevalent in leases in wealthier and demographically whiter areas of Philadelphia. Given the problem of ecological inference, this result should not be interpreted as an effect of individual tenant race or income. However, these patterns do suggest likely clustering of lease templates favorable to landlords among those landlords in particular regions of the city associated with more well-resourced landlords.

<sup>18</sup>To the extent that the two covariates are highly correlated with one another, it makes it difficult if not impossible to meaningfully disentangle whether it is race or income that is *causally* responsible. Moreover, the question of whether one or the other is the “true” driving factor is poorly defined as the two are highly interrelated within a geographically segregated city like Philadelphia.

### C. *Individual Effects*

One way to address the ecological inference problem is to examine the impact of individual-level demographic covariates conditional on regional characteristics. That is, after accounting for general neighborhood characteristics, do wealthier or whiter *residents* still have leases with more of the unenforceable provisions? Detailed information about the demographic and socioeconomic status of particular tenants is difficult to obtain. However, using addresses and names, we can extract some information on the background of the tenant—in particular we are able to generate predictions regarding the race and gender of the tenant. We also match and validate name and address pairs using a comprehensive voter file dataset maintained by the private political research firm L2.<sup>19</sup>

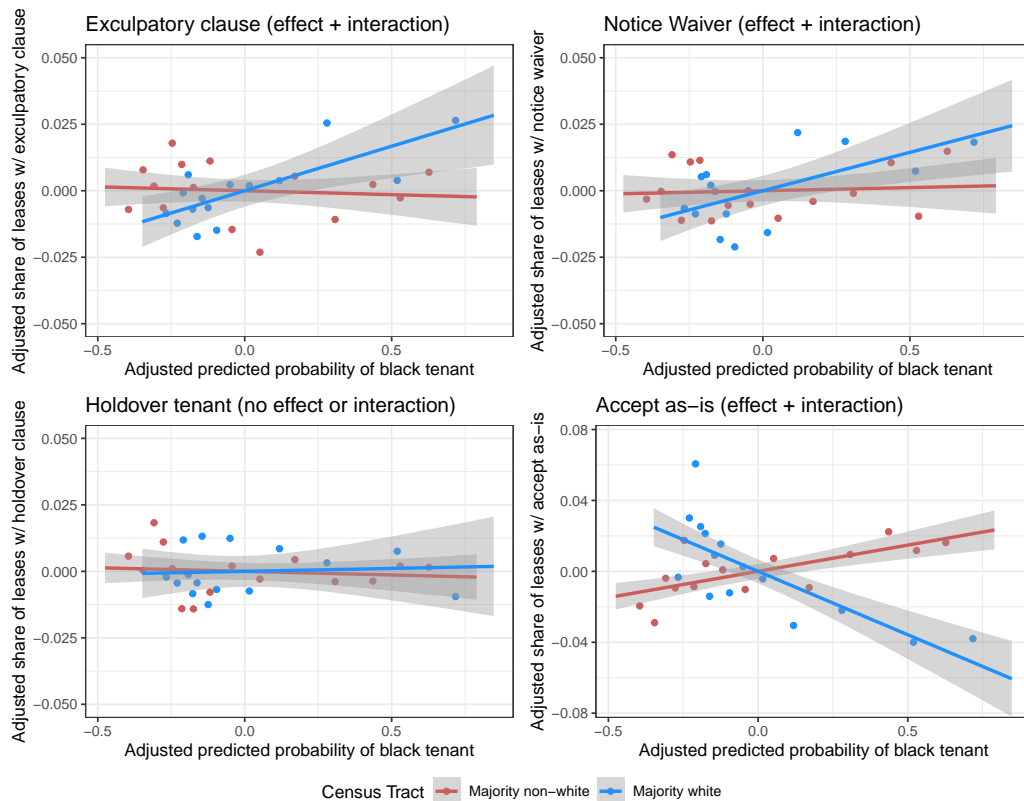
We predict tenant race using the surname analysis method frequently used in the social sciences (Fiscella and Fremont 2006; Imai and Khanna 2016) to impute individual-level race using known aggregate distributions from administrative data.<sup>20</sup> For our analyses below, our primary independent variables of interest are the predicted probability that an individual is non-Hispanic Black/African American and the predicted probability that an individual is male.<sup>21</sup>

---

<sup>19</sup>We extracted unique name-address pairings from our landlord-tenant dataset for cases filed between 2014 and 2019. We then requested that L2 match these name-address pairings to each of its annual historical voter file datasets from 2016 (as far back as was available) to the present day. Of the 79,861 unique tenant-address pairings from this time period, we were able to link 14,633 to an L2 voter file record. From these records, we obtain information on tenant gender, ethnic description, and estimated income where available from matched commercial data records.

<sup>20</sup>The approach leverages the fact that some surnames are more commonly among members of particular racial and ethnic groups. Given a known distribution of a particular surname using administrative data (such as census or voter registration data), the method uses Bayes' rule to obtain a predicted probability that an individual belongs to a particular ethnic or racial group conditional on their surname. We specifically implement the method developed by Sood and Laohaprapanon (2018) which also incorporates information on first names to improve the quality of the prediction. We use the full name model trained on the Florida voting registration data used by Sood and Laohaprapanon (2018) and implemented in the Python library `ethnicolr`, available at <https://github.com/appeler/ethnicolr>. Our method for imputing gender follows a similar methodology, using first names only using the approach described in Blevins and Mullen (2015). Our administrative dataset for first name-gender distributions is the Social Security Administration (SSA) dataset implemented in the `gender` R package, available at <https://cran.r-project.org/web/packages/gender/index.html>. We use the subset of individuals in the SSA dataset born between 1990 and 2012, ensuring that our name distribution reflects generally modern patterns of name-gender association even though we do not know the precise birth years of our tenants.

<sup>21</sup>We verified that our measure has good predictive power by comparing it to the race/gender classifications from those tenants who we were able to match to a commercial dataset obtained from political research firm L2. Using voter file data from 2016 onward, we matched 14,633 of 79,861 unique name-address pairings and obtained both registered voter data and commercial data. For gender, the estimated and actual shares line up quite neatly. However, race/ethnic classification in the L2 dataset is also imputed using similar methods as Pennsylvania voter registration does not collect data on race. Therefore, the relationship



*Notes:* N = 72,893 tenants, 48,039 addresses. Leases from 1999 to 2019. Multivariate linear regressions adjust for monthly rent (lease-level), percent White, and logged median income (census tract level, interacted with filing year) and lease start year fixed effects. Added variable plots visualize the bivariate relationship between tenant race and probability of lease provision adjusting for the other covariates. X-axis is the residual from a regression of tenant race on all other covariates. Y-axis is the residual from a regression of whether a lease contains the provision on all other covariates. Models fully interacted with an indicator for whether the lease census tract is majority white. All interactions between tenant race and majority white census tract statistically significant except for the holdover tenant outcome. Points denote binned averages (15 points per regression). Shaded intervals denote 95% cluster-robust confidence intervals clustered on address.

Figure 8: Regressions of lease provisions on predicted probability of black tenant - added variable plots

Figure 8 visualizes the results of our regressions of each of the four unenforceable provisions on individual tenant race, controlling for lease monthly rent, census tract per-

between our estimates and the proportion is positive but somewhat curvilinear due to imperfections in how the commercial dataset chose to coarsen the ethnic classification estimates.

cent white residents, census tract median income and year fixed effects. We estimate the relationship in census tracts where more than 50% of residents are white (156 tracts) and tracts where the percentage of white residents is less than or equal to 50% (216 tracts), fitting models fully interacted with an indicator for whether the address of the property comes from a tract that is majority white. We cluster standard errors at the level of the lease address, accounting for within-property correlations as multiple tenants can reside at the same property and the same property can appear in multiple eviction proceedings.

We find statistically detectable relationships between individual tenant characteristics and provision prevalence for three of the four provisions. For both exculpatory clauses and notice waivers, we find that in majority white sections of the city black tenants are more likely to receive leases with these types of provisions while no such relationship holds in majority non-white regions. On average, among properties in majority white areas, the probability of having a lease with an exculpatory clause increases by about 3.3 percentage points ( $p < .05$ ) when going from a tenant with an estimated 0% probability of being Black to a 100% estimated probability of being Black. Likewise, the probability of having a lease with a notice waiver increases by about 2.9 percentage points ( $p < .05$ ) for majority white areas. We find no statistically significant relationship between these two provisions and tenant race in majority non-white regions of the city and no relationship between tenant race and holdover tenant clauses in either of the two types of census tracts. Interestingly, the relationship between “accept as-is” clauses and whether a tenant is predicted to be black is notably *negative* in majority white areas and is slightly positive in *majority non-white* tracts (both statistically significant  $p < .05$ ).

These individual-level results may reflect a mix of differentiation by landlords and more fine-grained selection of tenants into certain types of properties even at the level of the census tract. We next included fixed-effects for each unique plaintiff-landlord.<sup>22</sup> When including landlord-level fixed effects, all of the relationships from the analysis above are no longer statistically distinguishable from zero. This suggests that the relationship between tenant race and lease provisions has less to do with unit-level choices by landlords and more to do with the types of properties being rented by Black tenants versus white tenants. The next section develops this account in greater detail by examining the growing prevalence of common *forms* that are shared among landlords.

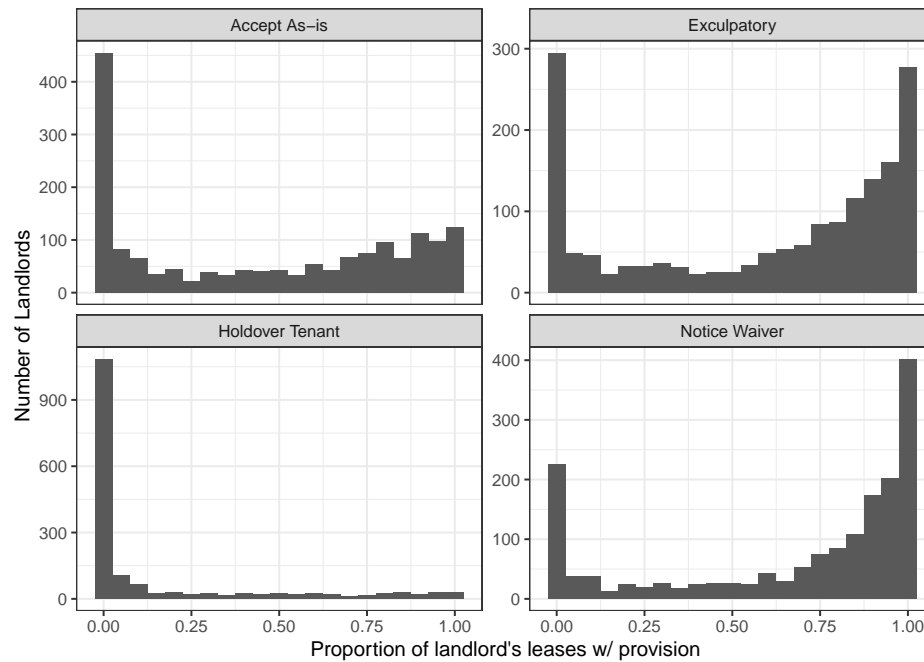
#### D. Templating

Do landlords vary whether they include certain clauses for different tenants? Examining the share of leases that contain each of the four provisions among frequent filing landlords suggests that they rarely do. Figure 9 plots histograms of the proportion of

---

<sup>22</sup>However, because some landlords maintain multiple LLCs and file their landlord-tenant cases under different names, this is a conservative approach to studying within-landlord variation.





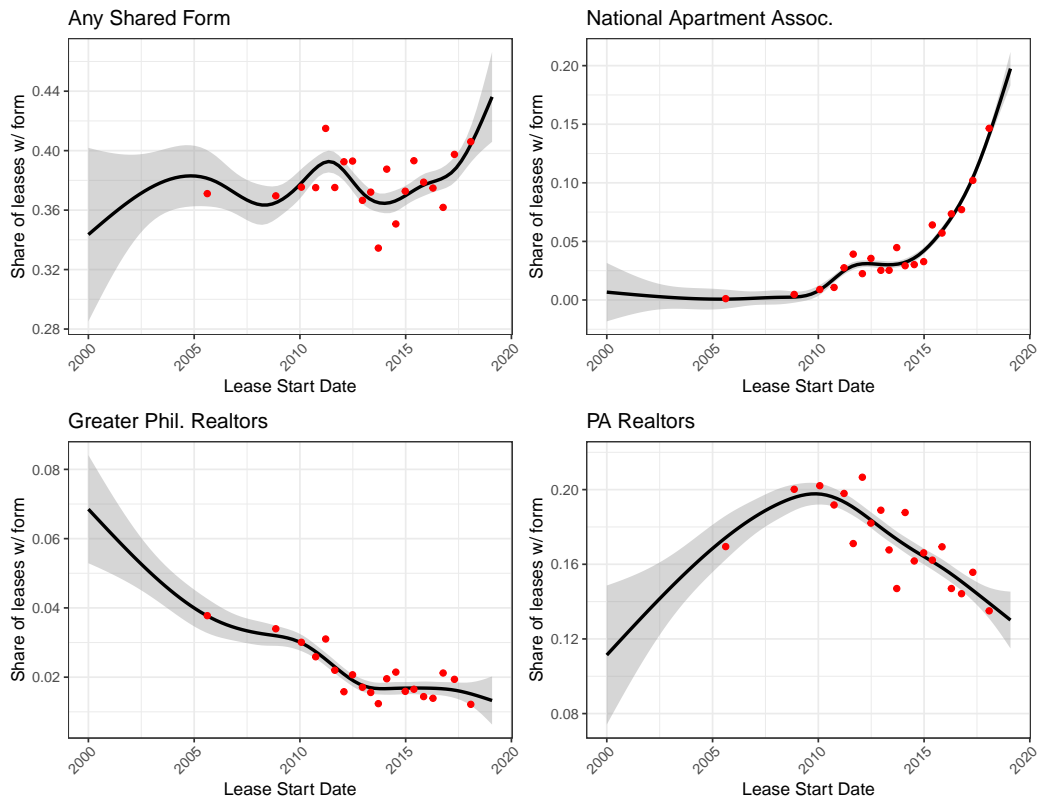
Notes: N = 1,673 landlords, 73,416 leases.

Figure 9: Prevalence of lease provisions within landlords (plaintiffs w/ 10+ leases)

leases within each landlord/plaintiff that contain one of the terms. We see clear bi-modal distributions with spikes at 0 and 1, especially for exculpatory clauses and notice waivers (with some noise expected due to our imperfect searches). For most of Philadelphia's landlords, the inclusion or exclusion of a particular provision is an all-or-nothing question and we find little evidence that there is any specific tailoring of leases to different tenants with respect to these clauses. Therefore the interesting variation in our data is not *within* landlords, but between them.

We hand-coded a random selection of 1000 leases, and found that many different landlords used the same widespread *shared lease*. Many of these shared leases are made available to landlords via online services that allow easy access to forms tailored to states. Using that hand-coding as the basis for text searches of the larger dataset, we found several dominant shared leases: the Philadelphia Association of Realtors Lease (approximately 16% of all leases in our dataset), the National Apartment Association Lease, the Landlord Association of Pennsylvania Lease, a "plain language" lease approved by the Pennsylvania Office of the Attorney General, the Greater Philadelphia Association of Realtors Lease, the Pennsylvania Apartment Association Lease, and a handful of internet leases

that were sourced from a number of websites such as ezLandlordForms.com, nolo.com or ZipLogix.<sup>23</sup>



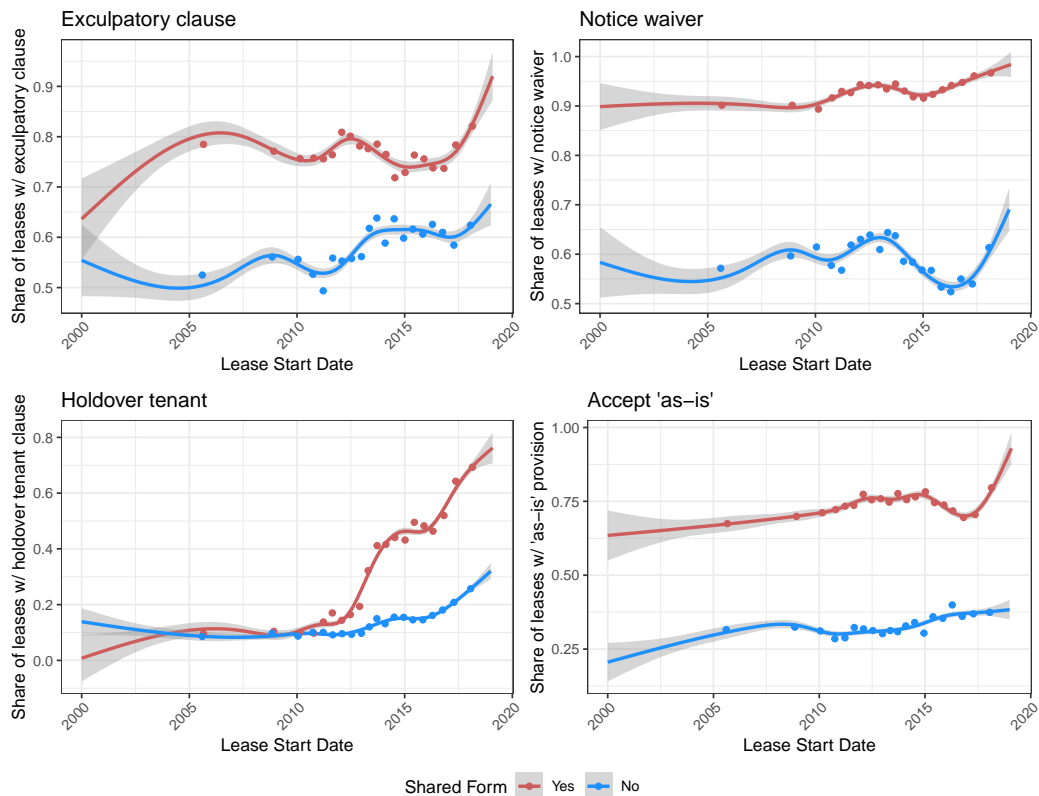
Notes: N = 88,921 leases with observed start date and geolocation post-2000. Estimates denote predicted probabilities from a cubic smoothing spline. Points denote binned averages (20 intervals).

Figure 10: Prevalence of shared leases over time

We coded each contract as being a *shared leases* if they matched one of these organization-sponsored leases or if they are listed as coming from a website that provides landlord forms. About 35% of leases in our dataset use one of these shared leases. This is likely an under-estimate of the total number of shared leases as some landlords may use identical (or largely identical) text from an existing lease while omitting relevant identifiers. Additionally, property management companies develop their own in-house leases which often have the same level of legal sophistication as the more popular shared leases. What we are

<sup>23</sup>Many of the forms obtained via online websites are also templates from one of these. For example, ZipLogix is often used to generate the Philadelphia Association of Realtors template. The role of non-profits and associations in generating forms is explored in (Davis 2007)

coding as *shared leases* are simply those that are common across many landlords. While essentially all leases are patterned off of models, inspecting a sample of the non-form leases in our dataset suggests that these are rather idiosyncratic – a combination of very simple leases and very long custom templates specific to a particular property manager. Going forward, we'll call contracts that aren't widely shared across landlords *proprietary leases*.



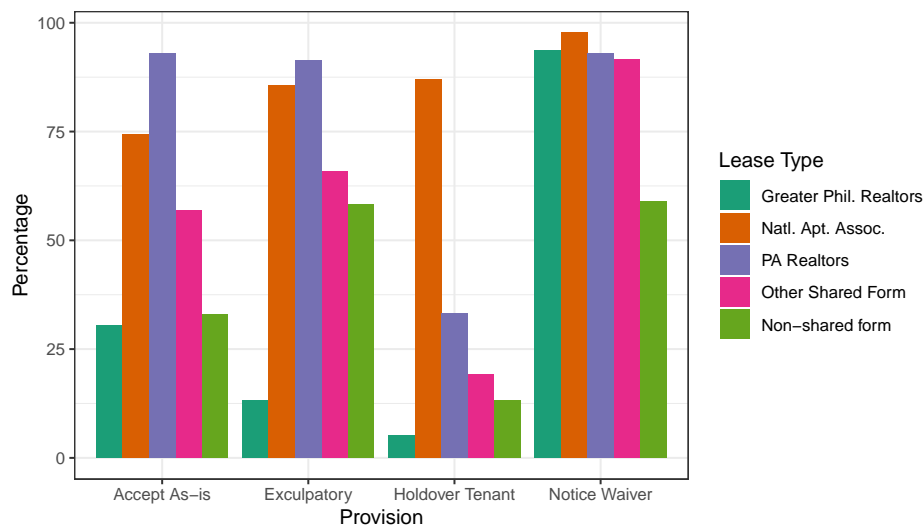
Notes: N = 88,921 leases with observed start date post-2000. Estimates denote predicted probabilities from a cubic smoothing spline. Points denote binned averages.

Figure 11: Prevalence of lease provisions by lease start date – Shared Leases vs. Proprietary Leases

The prevalence of shared leases increased with time. Figure 10 plots the estimated share of shared leases from 2000 to 2019. Their prevalence rises from about 34% in 2000 to over 43% in 2019. Additionally, the changes in *specific* shared leases have been more dramatic. For example, the National Apartment Association Lease is non-existent pre-2010 but rises to nearly 20% of the sample in 2019. Conversely, the Greater Philadelphia Realtors Lease has declined in popularity – dropping from about 6% of the sample in

2000 to essentially 0 in 2019. The PA Realtors Lease – one of the most popular templates, accounting for about 45% of our shared leases leases – rises in relative popularity from 2000 to 2010 but declines afterwards.

The spread of shared leases has important consequences for unenforceable provisions. Figure 11 plots the trend in the prevalence of each of our four provisions and compares shared to proprietary leases. First note the clear intercept shift — shared leases are much more likely to contain each of the four terms. Second, this gap appears relatively stable over time.<sup>24</sup> Third, we see some growth over time in the unenforceable provisions even among the proprietary leases, suggesting that there may be some convergence over time. This relationship is clearest when looking at the holdover tenant clause, which experiences a sharp increase in prevalence after 2010 that is driven almost by changes within shared leases: it was included in the 2013 edition of the Pennsylvania Association of Realtors Residential Lease. However, we see that the prevalence of this term grows even among the proprietary leases after 2010 as some landlords begin using proprietary leases containing this clause.

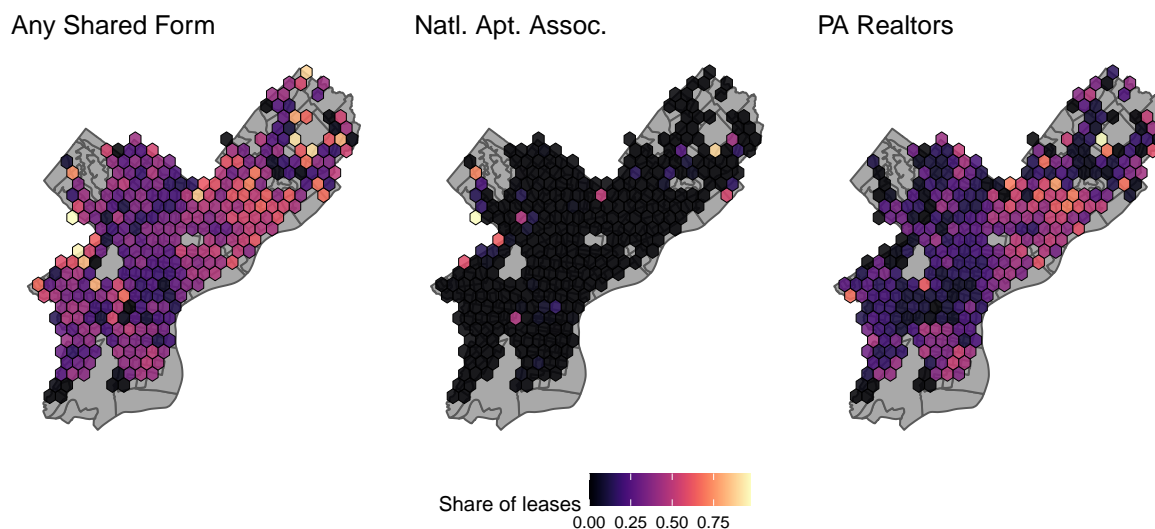


Notes: N = 88,921 leases with observed start date post-2000.

Figure 12: Share of leases containing each provision among those classified with particular templates

Furthermore, we find that the lease forms that are *growing* in popularity over time have worse provisions than those that are declining – that is, the composition of popular forms

<sup>24</sup>With one notable exception in the holdover tenant clause, which is almost entirely absent from all forms prior to 2010.



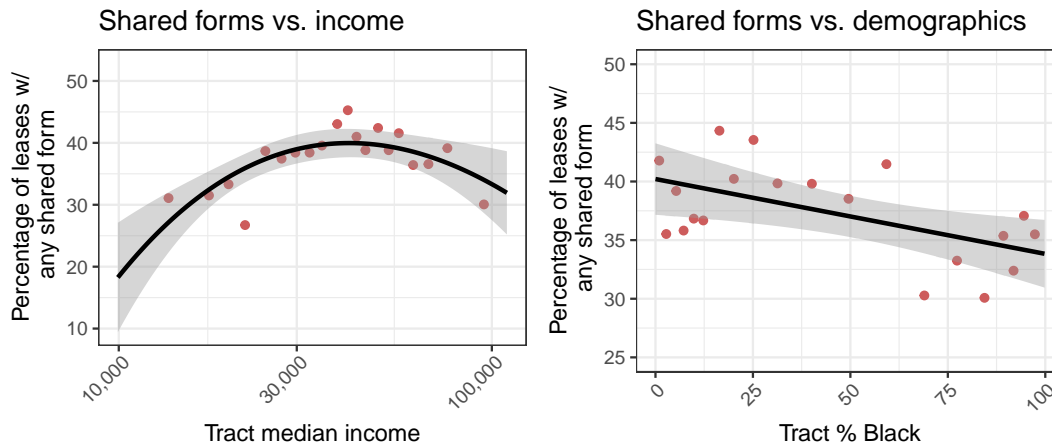
*Notes:* Unsubsidized, geocoded leases with non-missing start date after 2000 (N = 88,921). Hexes are approximately 1 square kilometer in area.

Figure 13: Geographic distribution of shared leases in Philadelphia

is changing in a way that is increasing the prevalence of our unenforceable terms. Figure 12 plots the share of leases with each of our four provisions among leases matching the Pennsylvania Association of Realtors (PAR) template, the National Apartment Association (NAA) template, and the Greater Philadelphia Association of Realtors (GPAR) template – three templates with different trajectories over time. In general, the forms that are growing their share of leases over time are worse for tenants than those in decline, which may explain some of the undulation in the time trends we see in Figure 11 as different templates wax and wane in popularity. It also may explain why the time trends in our four provisions are much more pronounced compared to the trend in templating. Not only are shared leases replacing proprietary leases, which tend to be shorter and more informal, but more pro-landlord templates are replacing less pro-landlord ones.

With shared lease networks explaining a significant chunk of why some leases contain these terms and others do not, do they exhibit similar clustering in space? Figure 13 plots the prevalence of shared leases in different parts of the city, along with specific sub-plots for two prominent types: the Pennsylvania Association of Realtors template and the National Apartment Association template. In general, we find heavier prevalence of shared leases in the downtown area and in the north-east. We find the National Apartment Association template concentrated in a few notable points in the city and nearly entirely absent elsewhere, likely reflecting use by large apartment complexes. The PA Association of Realtors template is much more diffuse as it is more widely available to smaller multi-

family and single-family housing but we still see clear areas of the city where it is much more common than others.

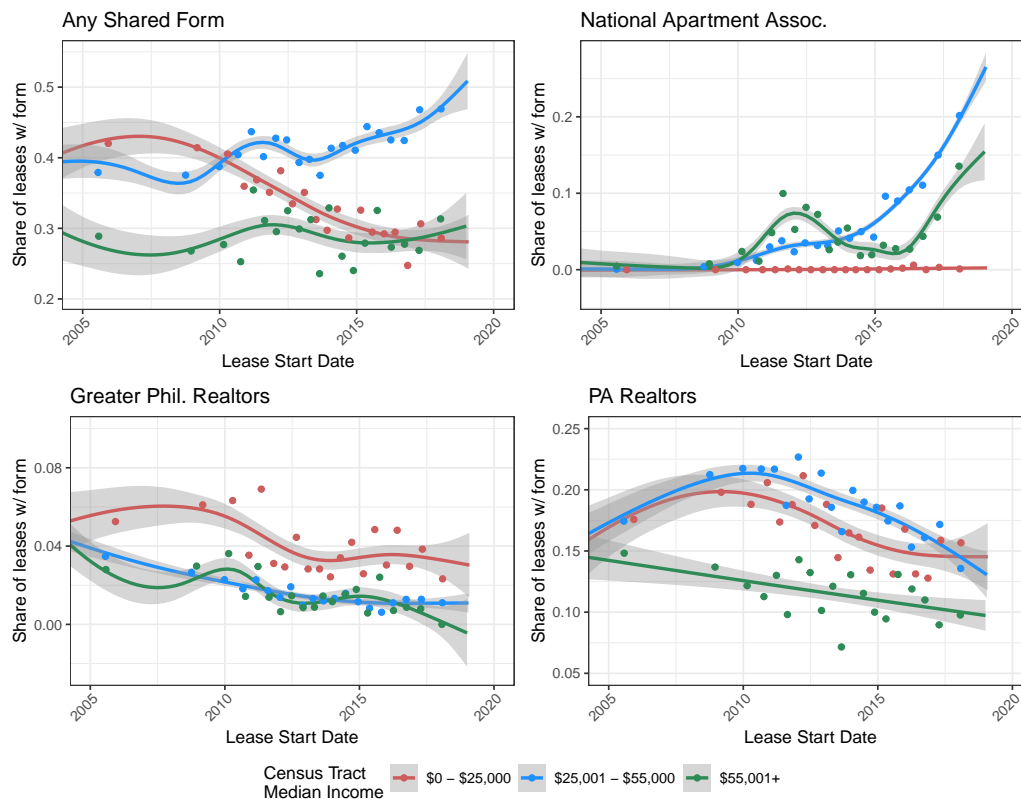


*Notes:* N = 363 census tracts with 10 or more sampled leases. Grey bands denote 95% HC2 robust confidence intervals. Median household income reported in 2017 inflation-adjusted dollars (ACS 5-year table S1901). Percent Black is the share of residents reporting their race as one race and Black or African American (2017 ACS 5-year table DP05). Points denote binned averages (20 equally-spaced bins). All regression slope coefficients statistically significant at  $p < .05$ .

Figure 14: Regression of share of leases containing shared forms in a census tract on logged median income and race (2017)

As expected, this clustering is associated with regional income and demographic characteristics. Figure 14 plots the regressions of the proportion of shared leases in each census tract on logged median income and % Black residents. We find one notable deviation from the patterns seen for the prevalence of unenforceable terms – tract income predicts more forms only up to a certain point. A quadratic polynomial fits the observed data much better than a linear model. Shared leases are most common in tracts where median incomes are in the \$50,000 range but their usage drops off in the most wealthy census tracts (which are comparable to the poorest in our dataset). We hypothesize that the use of custom templates among property management companies accounts may account for this pattern – these leases are often extremely dense and are developed by lawyers hired by the company but may not match any of the commonly used form templates. Unfortunately because such leases lack any common identifiers, evaluating this systematically is a difficult task.

What we do find, however, is that the growth of shared leases seems predominantly concentrated in census tracts with median incomes in roughly the 25th-75th percentile of all tracts. Figure 15 plots the time trend in the incidence of shared leases conditional on the census tract income bracket. Leases from census tracts in the middle of the distribution

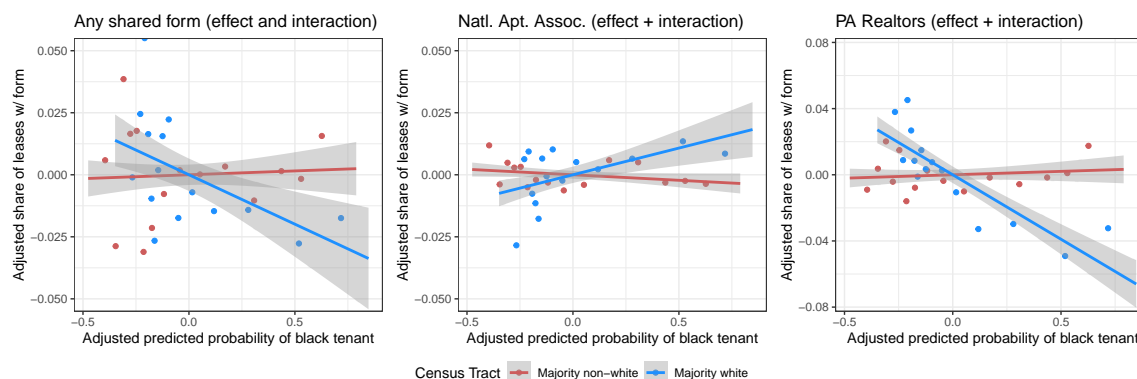


Notes: N = 88,698 leases with observed start date and geolocation post-2000 in census tracts with observed median income. Estimates denote predicted probabilities from a cubic smoothing spline.

Figure 15: Prevalence of shared leases over time by census tract median income

(56,185 leases from 191 tracts with median incomes between \$25,000 and \$55,000) are the ones with a clear growth over time in the proportion of shared leases being used. One thing that we find in the richer census tracts is relative stability in the overall share of shared leases but clear trends in *specific* types of forms – for example, a decline in the PA Realtors template coupled with a sharp growth in the National Apartment Association template. This to us is suggestive of *substitution* in form type towards templates more favorable to landlords among those landlords using forms and may further explain the high prevalence of the unenforceable provisions in richer vs. poorer census tracts despite comparable levels of form usage.

Turning to the results at the individual level, we find an interesting contrast to our original findings on lease provisions and race. Figure 16 shows that, on average, we find that Black tenants are less likely to have shared leases among those tenants in majority white census tracts. This is puzzling as we also find Black tenants are *more* likely to



*Notes:* N = 72,893 tenants, 48,039 addresses. Leases from 1999 to 2019. Multivariate linear regressions adjust for monthly rent (lease-level), percent White, and logged median income (census tract level, interacted with filing year) and lease start year fixed. Added variable plots visualize the bivariate relationship between tenant race and probability of lease provision adjusting for the other covariates. X-axis is the residual from a regression of tenant race on all other covariates. Y-axis is the residual from a regression of whether a lease contains the provision on all other covariates. Models fully interacted with an indicator for whether the lease census tract is majority white. Points denote binned averages (15 points per regression). Shaded intervals denote 95% cluster-robust confidence intervals clustered on address.

Figure 16: Regression of lease provisions on predicted probability of black tenant - added variable plots

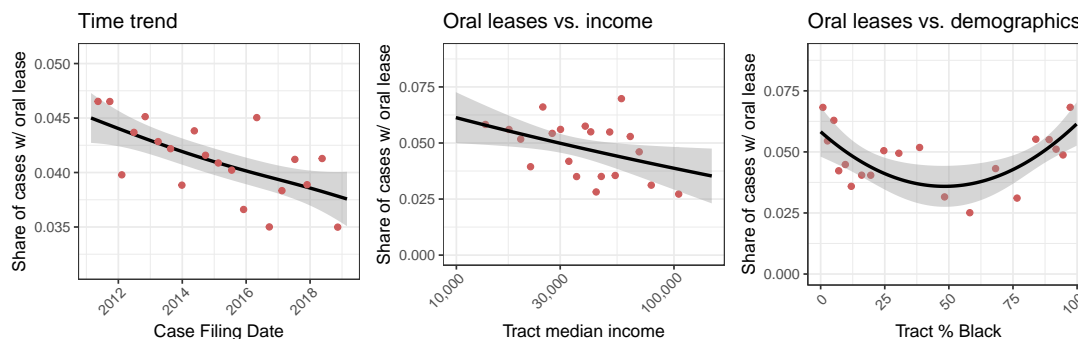
have leases containing the unenforceable provisions and that form leases tend to contain more unenforceable provisions. That is, *despite* being less likely to get a shared lease, Black tenants are more likely to get worse lease provisions. We consider two possible explanations for this.

The first is that among the universe of shared leases, black tenants are more likely to get those forms that are distinctly worse than others. We find some evidence for this — despite being less likely to get shared leases on average, Black tenants are on average more likely to have leases with the National Apartment Association template in particular, possibly reflecting a greater likelihood of renting from large multi-unit buildings that are NAA members. (This finding could follow from unobserved steering by brokers.)

Another explanation may be in the types of proprietary leases that Black tenants are likely to receive. Such leases either be short, idiosyncratic contracts or they can be one of the highly customized and dense leases used by larger property managers and apartment complexes. We hypothesize that Black tenants may be more likely to receive the latter. Notably, when we condition on the presence of a shared lease, the individual-level correlations between tenant race and lease provisions are no longer present and statistically significant, suggesting that the individual level variation in lease provisions between Black and non-Black tenants is primarily driven by differences among which of the proprietary



leases they receive.<sup>25</sup> In some sense, the pattern we see of greater formalization and spread of shared lease forms has been a convergence between white and Black tenants to a more pro-landlord set of provisions which were previously more common among Black tenants.



*Notes:* N = 180,082 landlord-tenant cases from 2011-2019. N = 376 census tracts with 10 or more sampled cases. Grey bands denote 95% HC2 robust confidence intervals. Median household income reported in 2017 inflation-adjusted dollars (ACS 5-year table S1901). Percent Black is the share of residents reporting their race as one race and Black or African American (2017 ACS 5-year table DP05). Points denote binned averages (20 equally-spaced bins). Time trend results are predicted probabilities from a cubic smoothing spline. Income and race regressions are linear/quadratic models estimated via ordinary least squares. All regression slope coefficients statistically significant at  $p < .05$ .

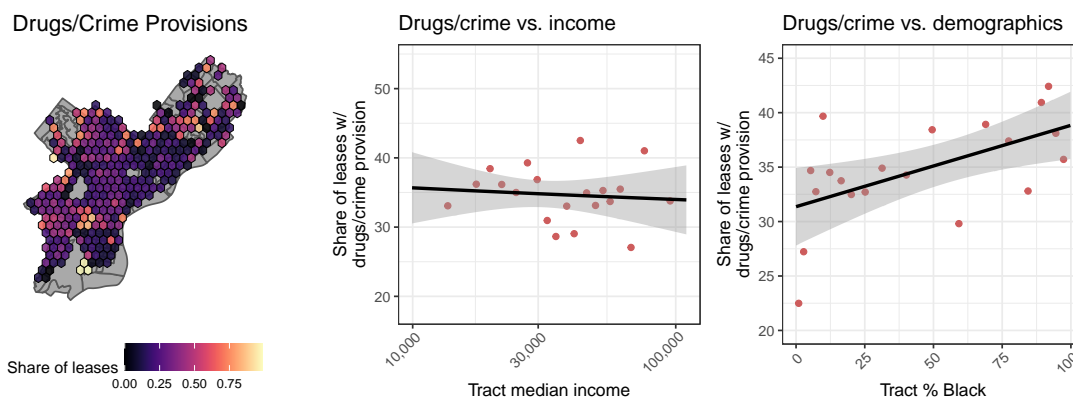
Figure 17: Time trends and census-tract level regressions for prevalence of oral leases in landlord-tenant cases (2011-2019)

Finally, we see more evidence supporting the hypothesis that legal formality is ascending. For the cases filed between 2011 and 2019, we have information recorded in the case filing on whether the lease agreement was oral or written. The vast majority of leases are written, but there are a few oral agreements (about 4 percent). Figure 17 shows that the use of oral leases has been declining over time (even within our relatively narrow timeframe) and is differentially distributed throughout the city. Census tracts with higher incomes have, as we would expect, fewer oral leases. Interestingly, the relationship with census tract demographics is curvilinear – tracts that have both with the tracts around 50% Black having the smallest share of oral leases among landlord-tenant cases brought.

<sup>25</sup>It is important to note that these results are meant somewhat tricky to interpret causally, even if we assume no unobserved confounders as whether a tenant receives a shared lease is arguably a “post-treatment” quantity and conditioning on post-treatment variables can induce artificial selection biases (see e.g. Montgomery, Nyhan, and Torres 2018). In this case, because Black tenants are less likely to get shared leases, the pool of non-Black tenants without shared leases and the pool of Black tenants taking proprietary leases may further differ on additional unobservable characteristics (e.g. property characteristics). While we do our best to control for what we can, it is important to approach this explanation with some caution even though we do find it highly plausible in this circumstance.

### E. Behavioral Control

Another set of lease provisions that are indicative of greater landlord power are clauses that permit eviction on the basis of tenant behavior. One very common set of provisions in many leases are “crime-free/drug-free” clauses, often included as addenda to the main lease, which permit termination of the lease and eviction in a variety of circumstances related to criminal activity or drug manufacture on the premises. These clauses (described in Appendix A) vary in their breadth, but often go far beyond just restricting the behavior of the tenants alone and permit eviction for criminal activity committed by third-parties. Such provisions originated in public housing with the development of the “one strike” policy by HUD during the 1990s which developed guidance for permitting public housing authorities to evict tenants if drug-related criminal activity occurs on the premises (Hannaford 2003). However, their application has grown into private-sector market leases through the enactment of “crime-free housing” ordinances (CHO) by many municipalities. These laws permit landlords to evict tenants for suspected criminal activity and often require that leases include provisions via a “crime-free lease addendum” that clarify the conditions under which eviction may occur for criminal activity (Ramsey 2018). Even where these provisions are not required by a CHO, many landlords nevertheless include such clauses either as part of the lease or as a supplementary addendum.



*Notes:* Unsubsidized, geocoded leases with non-missing start date after 2000 (N = 88,921). Hexes are approximately 1 square kilometer in area. N = 376 census tracts with 10 or more sampled cases. Grey bands denote 95% HC2 robust confidence intervals. Median household income reported in 2017 inflation-adjusted dollars (ACS 5-year table S1901). Percent Black is the share of residents reporting their race as one race and Black or African American (2017 ACS 5-year table DP05). Points denote binned averages (20 equally-spaced bins).

Figure 18: Geographic distribution of drugs/crime clauses in leases

We observe significant variation in which leases contain these provisions. Since there is no standardized language for such provisions, and because they often vary in precisely

what activities are covered, we conducted a broad search for any mention of “drugs” or “crime” in the lease texts.<sup>26</sup> Of the unsubsidized, private-sector leases in our dataset, 34.4 percent contain some mention of drugs or crime. This is in contrast to 99.7 percent of the leases with the Philadelphia Housing Association and 78.8 percent of leases that we identified as subsidized housing. Unlike the four unenforceable provisions, drug and crime provisions are only slightly more common in shared leases compared to proprietary leases (about 44.7 percent of form leases versus 37.4 percent of non-form). Some shared leases, like the National Apartment Association Lease, contain explicit mention of drug use and criminal activity while others like the PA Realtors do not. However, as these provisions are often included via addenda we expect that a landlord that wants to incorporate this additional control on tenant behavior can do so without modifying the content of the main lease contract.

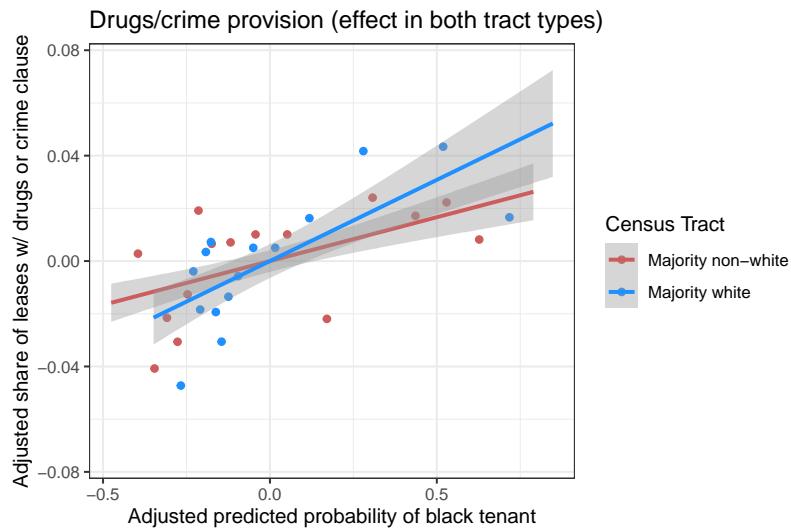
In contrast to our four unenforceable provisions, we find a different geographic distribution for the drug/crime clauses. Figure 18 plots the share of leases containing these clauses in different sectors of Philadelphia along with the census-tract level correlation with median income and percent Black. We find essentially no association between census-tract median income and the prevalence of drug/crime clauses but a strong positive relationship between the % of residents who are Black and the share of leases that contain these provisions. On average, the percentage of leases with a drug/crime provision increases by about 7.4 when comparing tracts with 0% Black residents to those that are 100% Black.

This relationship holds not only at the aggregate level but also at the individual level. Figure 19 plots the estimated relationship between the probability of a lease containing a drug or crime provision and tenant predicted race (probability Black) conditional on year fixed effects, census-tract characteristics and monthly rent. We find a strong, positive and statistically significant relationship between whether a tenant is black and the probability that they receive a lease with such a clause. Unlike the unenforceable terms, we see this relationship in both majority white and majority non-white census tracts. Moreover, unlike what we saw with the unenforceable terms where the variation seemed primarily driven by the non-form leases, this pattern remains regardless of whether we condition on either non-form or form leases. Black tenants who have form leases appear to be more likely to get those templates that contain a drug/crime clause or have drug/crime addenda.

However, as with the four unenforceable terms, when including landlord-level fixed effects (based on plaintiff name) these individual-level results no longer hold. That is, we do not observe landlords adopting different leases in distinct parts of the city, or for tenants of different races, genders or incomes. Rather, landlords appear to specialize in

---

<sup>26</sup>We do note that a number of standardized crime-free addendum templates do exist such as the template distributed by the International Crime Free Association (ICFA). See Ramsey (2018) pp. 1162-1163



*Notes:* N = 72,893 tenants, 48,039 addresses. Leases from 1999 to 2019. Multivariate linear regressions adjust for monthly rent (lease-level), percent White, and logged median income (census tract level, interacted with filing year) and lease start year fixed. Added variable plots visualize the bivariate relationship between tenant race and probability of lease provision adjusting for the other covariates. X-axis is the residual from a regression of tenant race on all other covariates. Y-axis is the residual from a regression of whether a lease contains the provision on all other covariates. Models fully interacted with an indicator for whether the lease census tract is majority white. Points denote binned averages (15 points per regression). Shaded intervals denote 95% cluster-robust confidence intervals clustered on address (to account for correlation in lease provisions between residents in the same unit).

Figure 19: Regression of drugs/crime lease provisions on predicted probability of black tenant - added variable plot

places, kinds of buildings, and therefore tenant-types. The race effects result in differences between landlords, not within them.

#### IV. IMPLICATIONS AND RESEARCH DIRECTIONS

We have focused on a selected dataset: contracts submitted to a court to justify an eviction. Nonetheless, we find that most such contracts are marked by multiple unenforceable terms. The primary exception are those contracts associated with public housing agencies, which have created a regulatory inspection process which scrubs leases of illegal language. Overall, we illustrate that *ex post* term policing doctrines appear to have low efficacy.

Landlords appear to adopt one of two strategies in choosing leases. Some adopt proprietary leases, which are often short, contain few unenforceable terms, and are associated with cheaper leaseholds in poorer census tracts with non-repeat play landlords. That is,

“mom-and-pop” leases. More often, and more recently, landlords adopt shared leases. These are typically drafted by non-profits or commercial providers and disseminated at low cost. They are longer and more pro-landlord than their proprietary counterparts. Over the last twenty years, the latter strategy has increasingly dominated the former, which we attribute to the decreasing cost of accessing partially-customized internet forms. Regional and individual-differences patterns are still apparent in these data. But whether they adopt proprietary or shared leases, or whether they are large or small, one thing we do not see is significant numbers of landlords tailoring leases for individual tenants.

This Part explores the implications of these findings, considers further research questions and describes certain limitations.

### A. *Regulating Templates*

One clear take-home message is that *ex ante* regulation is associated with significantly fewer pro-landlord terms than when such policing is left to *ex post* court decisions. That’s so even though all the studied leases were submitted in support of evictions, meaning that most landlords left themselves open to ready-made defenses if tenants could secure competent representation. Moreover, unenforceable terms are becoming more common over time, suggesting that although internet-based reputational markets are increasingly available, they do not work particularly well to regulate lease terms. By contrast, the Federal Housing Authority has been quite effective in cleansing leases of unenforceable terms. We do not find meaningful backsliding from federal mandates in subsidized leases, though the enforcement regime is under-resourced. These findings—the prevalence of illegal terms, and the weakness of *ex post* court policing, reinforce the small existing literature, but the comparative story is novel.

These data also illuminate the role of *forms*. We show a story of substitution over time: from oral to written leases, and from proprietary leases to widely-shared lease templates. As we show, formality and standardization is generally associated with significantly more landlord-friendly leases. This is an important addition to the literature on consumer contracting, on two levels.

*First*, it illustrates that, although these are rather expensive consumer contracts, of great importance to their adherents, there is essentially no within-landlord variation in terms. Leases are not tailored to tenants. The sort of strategic drafting behavior that the literature posits is a luxury available to those with the money to pay for counsel. Until recently, such landlords did not even have access to legal forms at all, let alone the capacity to calibrate forms to tenant- or building-specific risks. That is illustrated by Figure 17 which shows that oral leases are concentrated in poorer regions. Moreover, because the landlord-tenant court rarely focuses on the details of the lease in ordering evictions, particularly those that result from defaults, landlords are free to under-invest in legal protections.

It is perverse to find that access to justice, at least for landlords, may have hurt tenants. Proprietary, shorter, and oral leases, though they accompany decidedly lower-resourced leaseholds, contain fewer terms that take away tenants' rights. This finding runs counter to the relatively small extant studies on this topic. We show that as the costs of formalization and standardization decreased (due to the internet), and tenants increasing faced legally unfriendly leases. Thus, we illustrate one mechanism by which formality and access to legal services may hurt the poor.

Our results lay a path toward a justify a more muscular regulatory strategy. A common argument against pro-tenant interventions has been that they will have perverse effects, increasing the price for all tenants while merely benefiting those before the eviction court. But, ironically, because landlords appear to select between leases with little regard to tenant characteristics, and drafting costs seem to drive market outcomes, these data suggest that pro-tenant regulatory interventions are unlikely to increase rents. More vigorous enforcement of private law prohibitions on particular lease terms, or novel prohibitions on particular lease terms, would help tenants in court proceedings, and potentially reduce the incidence of evictions.

But though regulatory and *ex post* policing may do little harm, our results suggest a more direct strategy to improve leases. Advocates should focus their reform efforts on the purveyors of shared leases. Many such providers are non-profit associations that may be susceptible in the near term to public pressure in ways that courts and state legislatures are not. At the same time, a very small number of for-profit firms are providing legal forms over the internet that contains terms which we think to be at least arguably illegal. Reformers might consider legal strategies to change these firms' practices.

*Second*, our results have implications for the study of contracting writ large. Legal scholars have struggled to find examples of exogenous shocks that will produce widespread changes to consumer contracts. (Marotta-Wurgler and Taylor 2013) Here, we observe dramatic changes in practice—the rise of holdover tenant clauses, for example—in the absence of any corresponding shifts either in the legal regime or the rental market. Rather, non-profit associations' forms appear to drive the changes we observe in contract terms. This fits well with Davis (2007)'s theory of the role of such entities in contractual innovation, but is its first first empirical demonstration.

This does lead to a puzzle for further research, which is provide a more general account for the relative rise and fall of particular types of shared leases. Does the relative success of particular lease templates illustrated in Figure 10 arise from competitive forces acting on the lease level, or is it merely downstream from struggles between different landlord groups. To put it differently, we observe what looks like a competition between forms, with some emerging as more "fit" than others. But this may be an illusion, and what is really happening is the nationalization of the rental market, resulting in the de-localization of the forms they use.

### *B. The Role of Race*

Even in a world of shared leases, we uncover a potentially troubling racial finding. Recall Figure 19, which shows that overall black tenants are more likely to sign leases permitting eviction based on criminal conduct or drug use, and that this effect is stronger in white census tracts. This immediately suggests that landlords are (implicitly) choosing leases for such tenants based on a stereotyped projection about their behavior. That inference would follow from previous work on the persistence of racially restrictive covenants and other forms of racial discrimination in consumer markets. (Sood, Speagle, and Ehrman-Solberg 2019; Ayres and Siegelman 1995). This would not necessarily mean that different tenants in the same building sign different documents but rather that landlords use different leases for different properties based on their views of the mix of tenants in those places.

But because within landlord controls eliminate the effect, a different channel for the race-effect we observe is more likely.<sup>27</sup> Tenants choose where they live in a world of constraints. Spatial discrimination—i.e., black workers are effectively prohibited from moving freely around a city due to factors like access to mass transit or a lack of capital—may affect employment. (Zenou and Boccoard 2000) This is a reminder that race is intertwined with other characteristics—like income, social capital, employment and geographic mobility—which leads tenants to rent at particular properties. We observe both a geography effect—tenants in black precincts are more likely to see such clauses—and an individual one—black tenants in white precincts are more likely than similarly situated white tenants to get these provisions. These dual findings imply that landlords specialize both in places and in kinds of buildings/tenants, perhaps on racialized lines, and that they impose on black tenants living in white neighborhoods rules that they've imported from other property types.

Regardless of the precise etiology of why black tenants in white neighborhoods are more likely to face clauses permitting their eviction for crime or drugs, we do observe that they unequally confront that hazard. Such clauses provide landlords with an option to exercise at their discretion — simply put, black and white tenants in Philadelphia are not equally likely to face consequences for crime or drug use in their leaseholds, and the difference can not be explained by hyper-local crime trends. In future work, we intend to investigate how that discretion has been exercised. For now, regulators may consider whether to investigate large landlords' disparate use of these clauses as a leading indicator of other forms of discriminatory activity.

---

<sup>27</sup>As we discuss, a different mechanism is collider bias. That the effect persists when controlling for gentrification trends is reassuring, but we continue to seek more precise measures.

### *C. Limitations and Next Steps*

We have earlier discussed selection concerns. But even putting them aside, Philadelphia is a distinct city. It is poor (for a large American city), geographically diffuse, and its rental market appears to be dominated by small landlords. Whether the relationship between templating and enforceability would recur in other cities is up in the air, and, consequently, the generalizability of our findings to the tenancies across the country is unknown.

Important questions about this contracting ecosystem remain. We are interested in learning about how landlords react to changes in applicable legal rules over time, as a recent literature suggests the importance of such learning in how consumer contracts change. (Marotta-Wurgler and Dari-Mattiacci 2018). Recent changes to the rules in Philadelphia's Landlord-Tenant court may permit us to observe how legal rules spur changes in leases. Second, were it possible to learn more about management firms, we could determine how such intermediaries drive lease selection, and whether learning occurs at the network level or at the landlord level. Finally, we would be interested (and are exploring) in seeing how matches in the characteristics of landlords and tenants (such as race and gender, to the extent landlords are individuals) affect lease terms and price.

Our more immediate concern, however, is to study how lease terms affect the hazards of eviction itself. That is, are tenants forced to sign a lease depriving them of notice of eviction more likely to lose their tenancy? Does agreeing to be evicted for criminal conduct mean that eviction is more likely, controlling for the likelihood of criminal activity? Our next paper using these data will turn to those questions.

### CONCLUSION

This novel dataset consists of hundreds of thousands of Philadelphia leases drafted over the last twenty years. Shared lease templates are widespread and growing more prevalent over time. As such common contracts increasingly substitute for oral leases and proprietary forms, tenants face a more hostile legal terrain. We also uncover that tenants' race matters to what kind of leases they get, an effect we largely identify with segregated housing even at the hyper-local level. Residential leases thus reveal, and reinforce, existing disparities how formal legal rules affect citizens. With a clarified understanding of the etiology of those disparities in mind, better regulatory strategies, from changes in public laws to pressure campaigns against landlord associations, may come into focus.



## REFERENCES

- Ackerman, Bruce. 1971. "Regulating Slum Housing Markets on Behalf of the Poor: Of Housing Codes, Housing Subsidies and Income Redistribution Policy". *The Yale Law Journal* 80 (6): 1093–1197. ISSN: 00440094. <http://www.jstor.org/stable/795276>.
- Ayres, Ian, and Peter Siegelman. 1995. "Race and Gender Discrimination in Bargaining for a New Car". *American Economic Review* 85 (3): 304–21. <https://EconPapers.repec.org/RePEc:aea:aecrev:v:85:y:1995:i:3:p:304-21>.
- Bentley, Allen R. 1974. "An Alternative Residential Lease". *Columbia Law Review* 74 (5): 836–883. ISSN: 00101958. <http://www.jstor.org/stable/1121521>.
- Berger, Curtis J. 1974. "Hard Leases Make Bad Law". *Columbia Law Review* 74 (5): 791–835. ISSN: 00101958. <http://www.jstor.org/stable/1121520>.
- Blevins, Cameron, and Lincoln Mullen. 2015. "Jane, John... Leslie? A Historical Method for Algorithmic Gender Prediction." *DHQ: Digital Humanities Quarterly* 9 (3).
- Bucks, Brian, and Karen Pence. 2008. "Do borrowers know their mortgage terms?" *Journal of Urban Economics* 64 (2): 218–233. <https://EconPapers.repec.org/RePEc:eee:juecon:v:64:y:2008:i:2:p:218-233>.
- Choi, S., G. Gulati, and E. Posner. 2012. "The Dynamics of Contract Evolution". *Law & Society: Private Law eJournal*.
- Choi, Stephen J., and Mitu Gulati. 2004. "Innovation in Boilerplate Contracts: An Empirical Examination of Sovereign Bonds". *Emory Law Journal* 53:929.
- Coates, John C. 2001. "Explaining Variation in Takeover Defenses: Blame the Lawyers". *California Law Review* 89 (5): 1301–1421. ISSN: 00081221. <http://www.jstor.org/stable/3481160>.
- Cole, Stephen R, et al. 2010. "Illustrating bias due to conditioning on a collider". *International journal of epidemiology* 39 (2): 417–420.
- Collinson, Robert, and Davin Reed. 2018. "The effects of evictions on low-income households". [https://economics.nd.edu/assets/303258/jmp\\_rcollinson\\_1\\_.pdf](https://economics.nd.edu/assets/303258/jmp_rcollinson_1_.pdf).
- Davis, Kevin. 2007. "The Role of Nonprofits in the Production of Boilerplate". *New York University Law and Economics Working Papers* 104 (). doi:10.1017/CBO9780511611179.012.
- Desmond, Matthew, and Carl Gershenson. 2017. "Who gets evicted? Assessing individual, neighborhood, and network factors". *Social Science Research* 62:362–377.
- Fiscella, Kevin, and Allen M Fremont. 2006. "Use of geocoding and surname analysis to estimate race and ethnicity". *Health services research* 41 (4p1): 1482–1500.

- Furth-Matzkin, Meirav. 2017. "On the Unexpected Use of Unenforceable Contract Terms: Evidence from the Residential Rental Market". *Journal of Legal Analysis* 9 (1): 1–49.
- . 2018. "The Harmful Effects of Unenforceable Contract Terms: Experimental Evidence". *Ala. L. Rev.* 70:1031.
- Gulati, Mitu, and Robert E. Scott. 2012. *The Three Minute Transaction: Boilerplate and the Limit of Contract Design*.
- Hannaford, Rachel. 2003. "Trading Due Process Rights for Shelter: Rucker and Unconstitutional Conditions in Public Housing Leases". *U. Pa. J. Const. L.* 6:139.
- Humphries, John Eric, et al. 2019. *Does Eviction Cause Poverty? Quasi-Experimental Evidence from Cook County, IL*. Working Paper, Working Paper Series 26139. National Bureau of Economic Research. doi:10.3386/w26139.
- Imai, Kosuke, and Kabir Khanna. 2016. "Improving ecological inference by predicting individual ethnicity from voter registration records". *Political Analysis*: 263–272.
- Jennejohn, Matthew. 2018. "The Architecture of Contract Innovation". *Boston College Law Review* 59:71.
- Marotta-Wurgler, Florencia, and Giuseppe Dari-Mattiacci. 2018. "Learning in Standard Form Contracts: Theory and Evidence". *NYU Law and Economics Research Paper No. 18-11*.
- Marotta-Wurgler, Florencia, and Robert Taylor. 2013. "Set in Stone? Change and Innovation in Consumer Standard Form Contracts". *New York University Law Review* 88:240.
- Montgomery, Jacob M, Brendan Nyhan, and Michelle Torres. 2018. "How conditioning on posttreatment variables can ruin your experiment and what to do about it". *American Journal of Political Science* 62 (3): 760–775.
- Mueller, Warren. 1970. "Residential Tenants and Their Leases: An Empirical Study". *Michigan Law Review* 69:247.
- Nyarko, Julian. 2020. "Stickiness and Incomplete Contracts". *Chicago Law Review* 87.
- Padi, Manisha. 2018. "Consumer Protection Laws and the Mortgage Market: Evidence From Ohio".
- Ramsey, Kathryn V. 2018. "One-Strike 2.0: How Local Governments Are Distorting a Flawed Federal Eviction Law". *UCLA L. Rev.* 65:1146.
- Reinvestment Fund. 2021. "Evictions in Philadelphia: Race (and Place) Matters". *Reinvestment Fund Policy Brief*. [https://www.reinvestment.com/wp-content/uploads/2021/02/ReinvestmentFund\\_PHL-Evictions-Race-and-Place-Matters.pdf](https://www.reinvestment.com/wp-content/uploads/2021/02/ReinvestmentFund_PHL-Evictions-Race-and-Place-Matters.pdf).

- Sood, Aradhya, William Speagle, and Kevin Ehrman-Solberg. 2019. “Long Shadow of Racial Discrimination: Evidence from Housing Covenants of Minneapolis”. doi:10.2139/ssrn.3468520.
- Sood, Gaurav, and Suriyan Laohaprapanon. 2018. “Predicting race and ethnicity from the sequence of characters in a name”. *arXiv preprint arXiv:1805.02109*.
- Suchman, Mark. 2003. “The Contract as Social Artifact”. *Law & Society Review* 37 (): 91–142. doi:10.1111/1540-5893.3701003.
- Sullivan, Charles. 2009. “The Puzzling Persistence of Unenforceable Contract Terms”. *SSRN Electronic Journal* (). doi:10.2139/ssrn.1373841.
- Zamir, Eyal. 2014. “Contract Law and Theory: Three Views of the Cathedral”. *University of Chicago Law Review* 81.
- Zenou, Yves, and Nicolas Boccoard. 2000. “Racial Discrimination and Redlining in Cities”. *Journal of Urban Economics* 48 (2): 260–285.

A. APPENDIX: PENNSYLVANIA LAW AND UNLAWFUL AND OPPRESSIVE LEASE PROVISIONS

*Exculpatory Clauses*

Exculpatory clauses (disclaiming liability for negligence) are the best known unenforceable term in the existing scholarship about leases. (Mueller 1970) Exculpatory clauses in *commercial* leases are enforceable in Pennsylvania so long as they are very particular—that is, not merely exculpating in general terms.<sup>28</sup> They are clearly void if they seek to vitiate the landlord’s liability for statutorily-required duties (like providing fire exits.)<sup>29</sup> In between those extremes, Pennsylvania Courts focus on bargaining power, and disapprove of exculpatory clauses where “each party [is not] a free bargaining agent, which, it is evident, a prospective tenant for an apartment being unable to bargain away an exculpatory clause, is not.”<sup>30</sup> Following that rule, Pennsylvania courts have relatively routinely invalidated exculpatory clauses in residential leases in the absence of strong evidence of equal bargaining power or particular specific forms of negligence.<sup>31</sup>

Figure 20 shows an extract from one of the lease documents in our sample that we found used the 2017 PAR template. Sub-section B contains the exculpatory clause which disclaims the Landlord’s responsibility for any injuries occurring on the property.

185 20. INSURANCE AND RELEASE  
 186 (A) Tenant understands that Landlord's insurance does not cover Tenant, Tenant's personal property, or Tenant's guests. Tenant is advised to  
 187 obtain personal property and liability insurance to protect Tenant, Tenant's personal property, and Tenant's guests who may be injured  
 188 while on the Property.  
 189  IF CHECKED, Tenant must have insurance policies providing at least \$ \_\_\_\_\_ personal property insurance  
 190 and \$ \_\_\_\_\_ liability insurance to protect Tenant, Tenant's personal property and Tenant's guests who may  
 191 be injured while on the Property. Tenant must maintain this insurance through the entire Term and any Renewal Term. Tenant  
 192 will provide proof of insurance upon request. Tenant will notify Landlord within 10 days of changes to or cancellation of these  
 193 policies.  
 194 (B) Landlord is not legally responsible for any injury or damage to Tenant, Tenant's family, or Tenant's guests that occurs on the Property.  
 195 (C) Tenant is responsible for any loss to Landlord caused by Tenant, Tenant's family or Tenant's guests, including reasonable attorney's  
 196 fees associated with that loss, if awarded by a court.

Figure 20: Sample exculpatory clause - Pennsylvania Association of Realtors 2017 Standard Residential Lease

*As-Is Clauses*

Pennsylvania courts enforce a non-waivable implied warranty of habitability.<sup>32</sup> “As is” provisions, which violate that prohibition, are *per se* unenforceable.

<sup>28</sup>Topp Copy Products, Inc. v. Singletary, 626 A.2d 98, 99 (Pa. 1993)

<sup>29</sup>Boyd v. Smith, 94 A2d 44, 46 (Pa. 1953)

<sup>30</sup>Galligan v. Arovitch, 219 A.2d 463, 465 (Pa. 1966)

<sup>31</sup>Santiago v. Truitt, 23 Pa. D. C. 3rd 1982)

<sup>32</sup>Fair v. Negley, 390 A.2d 240, 245 (Pa. Super. Ct. 1978) (“We can only conclude that an attempted waiver of the implied warranty of habitability in residential leases is unconscionable and must be held to be ineffective.”)

The 2017 PAR lease likewise contains an “as is” provision as shown in Figure 21. Of course, there may be some variation in the disclaimer, and that variation could (under the caselaw) rescue the clause from a finding that it (in the words of the governing authority) made the leasehold unfit “for its intended purpose to provide premises fit for habitation by its dwellers.”<sup>33</sup>

110 14. CONDITION OF PROPERTY AT MOVE IN  
 111 Tenant has inspected the Property and agrees to accept the Property “as-is,” except for the following: \_\_\_\_\_  
 112 \_\_\_\_\_  
 113 \_\_\_\_\_

Figure 21: Sample as-is clause - Pennsylvania Association of Realtors 2017 Standard Residential Lease

### *Holdover Tenant Penalty Clauses*

A holdover tenant clause states that a tenant who stays over past the end of her tenancy owes some sum, typically a multiple of the rent, together with associated expenses. Below is an example from the 2013 PAR standard lease

22. **HOLDOVER TENANTS**  
 If Tenant occupies the Property after the Ending Date or end of any Renewal Term, Tenant will be considered a holdover tenant and will be causing the Landlord damages. These damages will be equal to 3 times the monthly Rent plus any lodging expenses of the new occupant, eviction costs and attorney fees, paid on a daily basis without demand.

Figure 22: Sample holdover tenant clause - Pennsylvania Association of Realtors 2013 Standard Residential Lease

Are such terms enforceable? The issue is familiar for contract students. Is the holdover tenant stipulated damage payment a reasonable estimate of the landlord’s liquidated damages, or is it a penalty clause, which attempts to punish breach? Pennsylvania courts have not ruled on such a clause in the residential lease context.<sup>34</sup> Other jurisdictions, however, appear to lean against enforcement of stipulated damage clauses against residential holdover tenants.<sup>35</sup>

<sup>33</sup>Pugh v. Holmes, 405 A. 2d 897, 906 (Pa. 1979). The point is that the “as is” clause is itself legally unenforceable.

<sup>34</sup>Rittenhouse 1603, LLC v. Barbera Eyeglasses, 2019 WL 1787475 (Pa. Sup. 2019). In commercial leases, Pennsylvania courts have accepted double rent. Pennsylvania Warehouse Beverage Stores, Inc. v. Brookhaven MZL, LP, 2015 WL 7258463 (Pa. Super. 2015).

<sup>35</sup>Mark S. Dennison, *Landlord’s Recovery of Damages for Tenant’s Wrongful Holding Over of Leased Premises*, 68 AM. JUR. POF 3D 1 (2019) (“The question of whether a provision for the payment of a stipulated amount, should the tenant withhold possession upon the expiration of a lease, is a provision for liquidated damages or for a penalty is determined by the facts of the particular case. If the court determines that the clause fixing damages is merely to secure performance of the agreement, it will be treated as a penalty and only actual damages proved can be recovered. In doubtful cases, the courts are inclined to construe the stipulated sum as a penalty.”)

In general, Pennsylvania follows a traditional, skeptical, view toward stipulated damages:

“[A]greements to pay a fixed sum without any reasonable relation to probable damages for breach tends to negative any notion that the parties really meant to provide a measure of compensation. Where a stipulated damages clause is intended as a form of punishment with the purpose, in terrorem, to secure compliance, the principles of compensation are subordinated and the provision must fail as an unenforceable penalty. A penalty is said to be fixed not as a pre-estimate of probable actual damages, but as a punishment, the threat of which is designed to prevent the breach.”<sup>36</sup>

In sum, though holdover tenant penalties have not been tested explicitly in Pennsylvania, we are fairly sure that most of them would be deemed unenforceable.

#### *Waiver of Notice*

The fourth provision that we consider is a waiver of the 15-30 day notice period required by the Philadelphia Landlord-Tenant Act of 1951 in the case of an initiation of eviction proceedings for a breach of the lease. This type of provision is enforceable as the Landlord-Tenant Act stipulates that a tenant may accept a shorter time or waive notice entirely in the lease.<sup>37</sup> However, we consider it an indicator of greater landlord power over the tenant in the case of eviction as it reduces the time available to a tenant to mount a defense. Since the vast majority of the judgments actually issued by the Philadelphia Landlord-Tenant court are default judgments in favor of the plaintiff, (approximately 89% in our sample), simply showing up and contesting an eviction attempt reduces the landlord’s likelihood of getting an eviction. Additionally waivers of notice, like exculpatory clauses, are explicitly banned in HUD subsidized housing leases.

An example of a waiver of notice, also found in the Pennsylvania Association of Realtors template is provided in Figure 23.

232 (B) IF TENANT BREACHES THIS LEASE FOR ANY REASON, TENANT UNDERSTANDS AND AGREES THAT TENANT  
 234 HAS WAIVED OR GIVEN UP TENANT'S RIGHT TO A NOTICE TO MOVE OUT UNLESS A DIFFERENT PERIOD  
 235 FOR PROVIDING NOTICE IS REQUIRED BY LOCAL ORDINANCE OR IS STATED HERE: \_\_\_\_\_  
 236

Figure 23: Sample waiver of notice - Pennsylvania Association of Realtors 2017 Standard Residential Lease

#### *Clauses About Criminal Conducts*

<sup>36</sup>Holt’s Cigar Co. v. 222 Liberty Associates, 404 Pa. Super 578, 587 (1991) (cleaned up)

<sup>37</sup>68 P.S. § 250.501 (e)

In addition to examining lease provisions that are unenforceable, we also consider the presence of “crime-free” or “drug-free” lease provisions that can be found in a sizeable number of leases but that we also suspect might exhibit systematic variation across geography and demographics. Many landlords include such provisions, either as an addendum or in the lease itself, which consider the presence of criminal activity – often specifically “drug-related criminal activity” – on the premises to be a violation of the lease. Since the language used by these provisions varies from lease to lease and some leases explicitly specify drug use or manufacture to be in breach, we constructed a broad search for any mention of “illegal drugs” or “controlled substances” or any mention of “criminal” or “illegal” activity.