Essays in Environmental and Healthcare Market Design

by

Anna Russo

B.S. Applied Mathematics and Economics, Yale University, 2017

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2024

Authored by:	Anna Russo Department of Economics May 15, 2024
Certified by:	Amy Finkelstein John & Jennie S. MacDonald Professor of Economics, Thesis Supervisor
Certified by:	Nikhil Agarwal Professor of Economics, Thesis Supervisor
Accepted by:	Isaiah Andrews Professor of Economics Chairman, Departmental Committee on Graduate Studies

Essays in Environmental and Healthcare Market Design

by

Anna Russo

Submitted to the Department of Economics
on May 15, 2024 in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

ABSTRACT

This thesis studies the design of government intervention in environmental and healthcare markets.

The first chapter, joint with Karl M. Aspelund, studies how asymmetric information influences the performance and design of government-established markets for conservation. Market mechanisms aim to deliver environmental services at low cost. However, this objective is undermined by participants whose conservation actions are not marginal to the incentive — or "additional" — as the lowest cost providers of environmental services may not be the highest social value. We investigate this potential market failure in the world's largest auction mechanism for ecosystem services, the Conservation Reserve Program, with a dataset linking bids in the program's scoring auction to satellite-derived land use. We use a regression discontinuity design to show that three of four marginal winners of the auction are not additional. Moreover, we find that the heterogeneity in counterfactual land use introduces adverse selection in the market. We then develop and estimate a joint model of multi-dimensional bidding and land use to quantify the implications of this market failure for the performance of environmental procurement mechanisms and competitive offset markets. We design alternative auctions with scoring rules that incorporate the expected impact of the auction on bidders' land use. These auctions increase efficiency by using bids and observed characteristics to select participants based on both costs and expected additionality.

The second chapter explores the observation that healthcare is often allocated without prices, sacrificing efficiency in the interest of equity. Wait times then typically serve as a substitute rationing mechanism, creating their own distinct efficiency and distributional consequences. I study these issues in the context of the Veterans Health Administration (VA) healthcare system, which provides healthcare that is largely free but congested, and the Choice Act, a large-scale policy intervention that subsidized access to non-VA providers to reduce this congestion. Using variation in Choice Act eligibility in both patient-level and clinic-level difference-in-differences designs, I find that the price reduction for eligible veterans led to substitution away from the VA, an increase in overall healthcare utilization and spending, and reduced wait times at VA clinics in equilibrium. I then use the policy-induced price and wait time variation to estimate the joint distribution of patients' willingness-to-pay and willingness-to-wait. I find that rationing via wait times redistributes access to healthcare

to lower socioeconomic status veterans, but at a large efficiency cost (-24%). This equity-efficiency trade-off is steep: rationing by wait times is an inefficient form of redistribution across a range of equity objectives. By contrast, I find that a coarsely targeted, modest increase in copayments increases consumer surplus by more than the Choice Act, at lower cost to the VA, while disproportionately benefitting low-income veterans.

The third chapter, joint with Abigail Ostriker, investigates the effects of regulations designed to correct a wedge between privately- and socially-optimal construction in areas at risk of flooding in Florida. Using a spatial regression discontinuity around regulatory boundaries and an event study around the policy's introduction, we document that floodplain regulation reduces new construction in high-risk areas and mitigates damages at homes constructed under flood-safe building standards. Embedding these effects in a model of the housing market, we find the policy reduces damages to the socially-efficient level, but incurs higher costs than a first-best corrective tax. Improved targeting of the existing policy achieves 94% of first-best welfare gains, or $7,567 per newly-constructed house.

JEL Codes: D4, D6, D8, H2, H5, I1, I3, Q2, Q5

Thesis supervisor: Amy Finkelstein
Title: John & Jennie S. MacDonald Professor of Economics

Thesis supervisor: Nikhil Agarwal
Title: Professor of Economics

Acknowledgments

To Amy, Nikhil, Ben, and Parag: Thank you, I could not have imagined a better advising team. You taught me to ask ambitious questions, tackle them with care and discipline, and communicate my ideas clearly and effectively. You are dedicated, thoughtful, kind advisors who challenged me to produce work that far exceeded any of my expectations. Most importantly, you taught me that economics research can and should be *so much fun*.

To the Public Finance and Industrial Organization faculty and students at MIT ('the PFs and IOs'): Thank you for being warm and welcoming communities for the past six years. You made it a joy to work across fields. To Mike and Tobias especially: Thank you for your many thoughtful comments and for being sources of kindness and encouragement throughout the PhD.

To my co-authors of this thesis, Abby and Karl: Thank you, I have learned so much from you. Working together makes research infinitely more fun, and I can't wait to keep at it for many years to come.

I am grateful for the financial support that allowed me to dedicate my time to this thesis. I acknowledge support from the National Science Foundation Graduate Research Fellowship Program, the Martin Family Society of Fellows for Sustainability, and the NBER Pre-Doctoral Fellowship in Aging and Health Research.

Finally, to my family: Thank you for your unwavering support. And to Zach: Thank you for being the best teammate, in this endeavor and in life.

Contents

List of Figures

List of Tables

Chapter 1

Additionality and Asymmetric Information in Environmental Markets: Evidence from Conservation Auctions

1.1 Introduction

Land-use change contributes 13% of global greenhouse gas emissions (Friedlingstein et al., 2022) and leads to biodiversity loss, water pollution, and erosion (Dirzo et al., 2014; Vörösmarty et al., 2010; Borrelli et al., 2017). While environmental markets can, in theory, reduce environmental degradation at low cost (Samuelson, 1954a; Anderson and Libecap, 2014; Teytelboym, 2019), many believe that existing mechanisms have failed to meet this potential (Anderson, 2012; Filewod, 2017; Maron et al., 2016; West et al., 2020; Jones and Lewis, 2023). A leading explanation for this failure is the possibility of inframarginality: some participants may have engaged in the incentivized action even absent an incentive. The notion of "additionality," defined as the likelihood that an action is marginal to an incentive, is a central challenge to the design and success of many environmental markets (Engel et al., 2008; West et al., 2023).

Does the challenge of additionality drive markets to failure, undermining environmental incentive policies and offset markets?[1] Or can markets be designed to achieve low-cost climate change

[1] In offset markets, private buyers purchase contracts that "offset" environmental degradation acre-for-acre or ton-for-ton. Offset markets exist in a range of settings, due to direct implementation from regulators (wetlands and air pollution), to allow for gains from trade between regulated and unregulated industries (e.g. compliance offsets in California's cap-and-trade program), between countries to provide flexibility in meeting international emissions commitments (the Clean Development Mechanism and REDD+), and due to the large volume of voluntary net-zero commitments among firms (McKinsey Sustainability, 2021, 2022). See Salzman et al. (2018) for an overview of Payments for Ecosystems Services, specifically.

mitigation? We explore these questions by analyzing the challenge of additionality as a market failure due to asymmetric information. Social welfare in markets for environmental conservation depends on both a landowner's unobserved additionality and her private cost of complying with the market requirements. Market mechanisms, however, screen only on the latter. If asymmetric information prevents incentives from reflecting heterogeneity in landowner additionality, market mechanisms may not achieve allocative efficiency and in the extreme, may fail (Akerlof, 1970a). In this paper, we use this perspective to analyze, test, and quantify this potential failure and to examine remedies in alternative market designs.

We conduct our analysis in the context of the United States Department of Agriculture's (USDA) Conservation Reserve Program (CRP), one of the oldest and largest Payments for Ecosystem Services (PES) mechanisms in the world.[2] The CRP incentivizes agricultural land retirement and conservation actions via procurement auctions of conservation contracts. CRP contracts pay landowners $1.6-$1.8 billion per year to take cropland out of production and to plant grass mixes, plant or maintain trees, or establish habitats for a duration of ten years. Combining administrative data and high-resolution satellite imagery, we construct a dataset that links landowners' multi-dimensional bids in the CRP scoring auction to their land use, which we use to measure additionality. The CRP auction provides a rich empirical setting for each step of our analysis: assessing the extent of additionality, testing for asymmetric information, and quantifying their implications for social welfare under current and alternative market designs. Moreover, the insights from this setting are broadly applicable: CRP contracts are structured similarly to other PES programs, to contracts traded in global offset markets, and to private competitive agricultural offset markets in the US.[3]

We first analyze the market failure introduced by additionality with a stylized framework that builds on the graphical analysis in Einav et al. (2010a). Landowners differ in two dimensions. The first is their cost of contracting, which includes the forgone option value of cropping and the hassle costs of complying with program requirements. The second is their conservation behavior without the contract, which determines their additionality. The social value of contracting depends on a landowner's cost and additionality, but her choices depend only on her cost and the market incentive. This difference can lead to allocative inefficiency. When a landowner's cost of contracting is positively correlated with her additionality — for example, landowners who expect to conserve regardless of the program have lower opportunity costs of contracting — there will be adverse selection in the market. In procurement, adverse selection can limit the implementability of efficient allocations and undermine the performance of standard mechanisms (Manelli and Vincent, 1995). In competitive offset markets, adverse selection can limit trade because buyers consider the

[2] Over its history, the CRP is the largest PES mechanism in the world. Within a given year, the CRP is second to China's Sloping Land Conversion Program.

[3] See Kinzig et al. (2011), Engel et al. (2008), and Stubbs et al. (2021), respectively.

expected additionality of all market participants, not only those contracting at the margin. These challenges can be remedied if markets are designed to close the gap between socially-optimal choices and the choices made in the market.

The stylized framework provides guidance for empirical analysis. Social welfare under current and counterfactual market designs depends on the distribution of landowner contracting costs and the population expectation of additionality at each value of costs. Contracting costs and additionality may be correlated due to landowners' expectations of low payoffs from cropping land. However, landowners may have only limited information about future payoffs to cropping over the contract's duration and incur hassle costs that may be arbitrarily correlated with additionality (Jack and Jayachandran, 2019). The extent of additionality, the existence of adverse selection in the market, and together, their quantitative implications for the performance and design of markets for environmental services are empirical questions.

We begin by examining the extent of additionality in our setting. Credible estimates of additionality, particularly in large-scale mature markets, are scarce as they require knowledge of an unobserved counterfactual. We use the discontinuity in contracting around the winning bid in the procurement auction to evaluate additionality at the margin of contract awards. We find that, as incentivized by the CRP, landowners substitute away from agriculture to natural vegetation and grasslands upon contracting. However, only one quarter of marginal winners are additional, which we calculate by comparing the regression discontinuity treatment effect to the magnitude of land contracting at the margin. In other words, three quarters of landowners at the margin of contract awards would have conserved without a CRP contract. However, the status quo auction implicitly assumes all landowners are additional in the design of its scoring rule (Claassen et al., 2018).

To test for adverse selection in the market, we correlate heterogeneity in additionality with heterogeneity in the costs of contracting reflected in landowner bids. We make two assumptions — perfect compliance and no spillovers, both of which we test and validate — to obtain a landowner-specific measure of additionality for all rejected bidders (82% in the most restrictive auction). We examine the relationship between landowner-specific additionality and bids following classic tests for asymmetric information in insurance markets (Chiappori and Salanie, 2000) and auctions (Hendricks and Porter, 1988). We document substantial heterogeneity in additionality and a positive correlation between additionality and bids, indicating the presence of adverse selection in the market. The positive correlation persists even conditional on a rich set of observed characteristics. This analysis also highlights opportunities for improvements to market design: heterogeneity in additionality is predicted by landowners' choice of contract in the mechanism and by the observed characteristic of soil productivity.

These facts demonstrate that both additionality and asymmetric information are relevant to the

function of this market; to quantify their welfare implications and evaluate the performance of counterfactual market designs, we develop and estimate a joint model of bidding and additionality. First, we infer costs of contracting from revealed preferences in optimal bidding. Then, we estimate landowner additionality, including how it varies with costs, by matching the moments of land use presented in the first half of the paper.

In the CRP auction, landowners submit multi-dimensional bids on a menu of heterogeneous contracts, which are ranked by a scoring rule. This provides a rich environment for market design, as scores across the menu of contracts and observed asymmetry terms are tools to increase social welfare. In the first part of the model, we extend the multi-dimensional bidding models of Asker and Cantillon (2008b) and Che (1993) to a setting with a discrete contract choice and a non-linear scoring rule. In the second part, we model additionality with a conditional expectation function that relates land use to both observed characteristics and unobserved landowner costs. This conditional expectation function is the component of the model that captures the possibility of inefficient or adverse selection.

We estimate the model in three steps. The first two steps adapt standard procedures for the empirical analysis of auctions (Guerre et al., 2000a; Hortaçsu, 2000; Hortaçsu and McAdams, 2010; Agarwal et al., 2023). First, we estimate bidder beliefs via simulation. Second, we estimate bidder costs via revealed preferences in optimal bidding. Because of the discrete choice in the bidding problem, we rely on variation in the scoring rule for identification. In the final step, we estimate additionality, including how it varies with unobserved landowner costs, by matching the levels of additionality and the correlation between additionality, landowner characteristics, and optimal bids observed in our linked land use and bid data. We use our estimates of additionality to calculate the social benefits of contracting based on valuations of environmental services from the CRP literature and the USDA's revealed preferences across landowners and contracts implied by the scoring rule.

Using these estimates, we first examine whether the existence of a market for conservation increases social welfare. When some landowners in the market are not additional, this is theoretically ambiguous; we investigate it empirically in a simple uniform market for the base contract. We find substantial social welfare gains under the socially-optimal uniform price ($14.66 per acre-year) and in a stylized competitive offset market ($14.11 per acre-year). The difference between these two market structures (-4%) reflects the trade-limiting effects (-15%) of adverse selection in competitive markets. Despite landowners who are not additional and the adverse selection this introduces, we find that the market does not fail.

We then evaluate the performance of the status quo auction mechanism. We estimate that the status quo mechanism leads to social welfare gains of $126 million per auction, relative to no

market. However, it implements only 15% of the social welfare gains of an efficient allocation. This allocation determines contract awards based on both landowners' costs and expected social benefits, which depend on additionality.

Implementing the efficient allocation with an incentive compatible auction may not be possible (Myerson, 1981a). Because they are less additional, lower cost landowners are not always higher social value. Our estimates imply that the allocation rule for this efficient benchmark need not be monotone in bidder cost.

We instead propose and evaluate alternative auctions with scoring rules that trade off bidders' costs against both their conservation actions' heterogeneous social benefits *and* their expected additionality. This differs from the status quo scoring rule, which does not consider the latter. Alternative scoring rules adjust asymmetry across bidder observables and scores across the menu of contracts based on predictions of additionality. Instead of restricting participation in the market through eligibility requirements, our approach re-designs the auction to *impact conservation,* acknowledging that some landowners in the market may not be additional.

Simple modifications to the auction's scoring rule close the gap between the status quo and efficient allocation by 41%, increasing social welfare by $284 million per auction. All gains are due to changes that incorporate landowner additionality. A large share are from setting the socially-optimal uniform adjustment to the scoring rule. Further gains accrue from using the rule to differentiate among heterogeneously additional landowners. By contrast, switching from the status quo (inefficient) auction to an (if all landowners were additional, efficient) Vickrey-Clarke-Groves mechanism that remains naive to additionality reduces social welfare.

We conclude with the implications of supply-side adverse selection for competitive offset market design.[4] Competitive markets introduce distinct considerations: a differentiated market may or may not be more efficient than a uniform one. Differentiation based on available covariates would increase social welfare in a stylized competitive offset market for the base contract by 15%, reducing both inefficient selection and inefficiently-limited trade due to adverse selection. Next, we consider which contracts could be traded. Markets for tree planting and maintenance unravel, while social welfare losses from adverse selection in other markets, including grasses planting and habitat creation, are limited to at most 3%.

Together, our results highlight that although additionality and the adverse selection that it introduces are relevant in practice, and in theory, can cause markets to fail, voluntary environmental markets can deliver on their promise of low-cost conservation. However, successful market design must consider not only the heterogeneity in private costs that determine choices, but also the

[4]The Growing Climate Solutions Act of 2021 includes provisions for the USDA to serve as a regulator of agricultural offset markets.

implications of these choices for additionality and social welfare in the market.

Related Literature Our primary contribution is to develop an empirical framework to evaluate social welfare under current and counterfactual market designs in the presence of the additionality market failure. We also provide credible estimates of the extent of additionality and evidence of adverse selection in a large-scale, mature market for ecosystem services. Our regression discontinuity estimates of additionality contribute to a literature estimating treatment effects of payments for ecosystem services (Jack, 2013; Alix-Garcia et al., 2015; Jayachandran et al., 2017a; West et al., 2020; Rosenberg et al., 2022) and inframarginality in offset markets (Calel et al., 2021). Our framework builds on theoretical analyses of asymmetric information (van Benthem and Kerr, 2013; Mason and Plantinga, 2013; Li et al., 2022; Haupt et al., 2023) and empirical tests for selection (Montero, 1999; Jack, 2013; Jack and Jayachandran, 2019) in environmental incentive programs and offset markets.

Though our context is additionality in conservation incentives, our framework relates broadly to the design of environmental incentive programs and other voluntary regulation (Allcott and Greenstone, 2012; Borenstein, 2012; Allcott and Greenstone, 2017; Einav et al., 2022; Ito et al., 2023) and complements work studying other sources of inefficiency in markets for environmental conservation (Harstad, 2016; Harstad and Mideksa, 2017; Aronoff and Rafey, 2023). Beyond environmental markets, our approach to auction design relates to a literature evaluating market designs based on treatment effects, not only revealed preferences (e.g. Agarwal et al. (2020)).

We contribute to a literature studying quality concerns in procurement auctions (Manelli and Vincent, 1995; Decarolis, 2014a; Carril et al., 2022; Lopomo et al., 2023), where we provide an empirical framework to evaluate alternative auction designs in the presence of adverse selection on bidder quality (additionality in our setting). This empirical framework draws on a large literature studying selection in insurance markets (Akerlof, 1970a; Chiappori and Salanie, 2000; Einav et al., 2010a; Bundorf et al., 2012a; Marone and Sabety, 2022).

Methodologically, our model and estimation strategy use techniques from a rich literature advancing the empirical analysis of auctions (Guerre et al., 2000a; Hortaçsu, 2000; Hortaçsu and McAdams, 2010; Jofre-Bonet and Pesendorfer, 2003; Agarwal et al., 2023). We draw on and extend existing work on scoring and other multi-dimensional auctions (Che, 1993; Asker and Cantillon, 2008b, 2010; Lewis and Bajari, 2011; Sant'Anna, 2017; Hanazono et al., 2020; Kong et al., 2022; Allen et al., 2023; Bolotnyy and Vasserman, 2023b) to incorporate discrete bidding, a non-linear scoring rule, and a correlation between additionality and bidder costs.

1.2 Theoretical Framework

We present a framework to analyze additionality in markets for environmental services.

1.2.1 Model

There exists a continuum of landowners, indexed by i, each making a decision to contract, $x_i \in \{0,1\}$, to obtain a transfer, p. In Section 1.5, we adapt this framework to a finite number of landowners bidding for contracts in a procurement auction with a quantity constraint.

The contract involves a promise to provide an environmental service $(a_i = 1)$ versus not $(a_i = 0)$. In our setting, $a_i = 1$ denotes agricultural land retirement (conservation) and $a_i = 0$ denotes cropping. The action $a_i = 1$ generates social benefits from positive environmental externalities. The buyer of the contract — either a regulator or a private buyer in an offset market — values the social benefits from $a_i = 1$ at $B > 0$.

Define a_{i1} as landowner i's action when $x_i = 1$ and a_{i0} as her action when $x_i = 0$. We assume perfect compliance, so $a_{i1} = 1$. Because B is generated whenever i chooses $a_i = 1$, regardless of contract choice x_i, the benefit of contracting with i is only the *incremental value* $B \cdot (1 - a_{i0})$. a_{i0} is unobserved whenever $x_i = 1$ and is therefore non-contractible.

A landowner's decision to contract x_i is the only available instrument to affect the provision of the environmental service. A Pigouvian subsidy (B) on all conservation would correct the externality, but is unavailable due to considerable practical constraints.[5] We thus focus on the performance and design of existing markets for environmental services. In practice, the instrument x_i is a binding long-term contract, and compliance may involve hassle costs, as documented across social programs, including in Payments for Ecosystem Services programs (Jack and Jayachandran, 2019).

Landowner Types Each landowner i is characterized by a type $\theta_i = (c_i, a_{i0})$ distributed according to the cumulative distribution function $F(\theta)$. c_i is a landowner's cost of contracting, defined as the minimum transfer p required for a landowner to accept the contract $x_i = 1$. a_{i0} is a landowner's expected action absent the contract.[6] We do not restrict the joint distribution of c_i

[5] The $36 billion Payments for Ecosystem Services industry (Salzman et al., 2018) has emerged as a second-best substitute to a first-best Pigouvian corrective instrument. This is motivated by limits to the feasibility of a Pigouvian instrument in practice. A Pigouvian subsidy on all conserved agricultural land in the United States would cost over forty times the current budget of the Conservation Reserve Program. A tax on $a_i = 0$, while raising revenue, may face political constraints. Both may be costly to monitor and administer. Moreover, binding long-term contracts allow for an extended period of conservation, which facilitates ecosystem development.

[6] We define landowners by the expected action a_{i0}, given each landowner's information about the payoffs to cropping versus conserving, at the time of contracting. Results that apply to a hidden information model also apply to a hidden action model (Milgrom, 1987).

and a_{i0}. Landowners may have a low c_i because they have unprofitable land that they do not plan to crop $(a_{i0} = 1)$. But landowners may face uncertainty over the stochastic payoffs to cropping that determine their forgone option value of cropping over ten-year contracts, and contracting in realistic settings involves activities beyond choosing not to crop that impose hassle costs that enter c_i and may be arbitrarily correlated with a_{i0}.[7] It is therefore ambiguous whether and how c_i and a_{i0} are related.

It will be useful to define the conditional expectation function:

$$\tau(c) = \mathbb{E}\left[1 - a_{i0} \,|\, c = c_i\right]. \tag{1.1}$$

This function describes the expected additionality, or the expected impact of contracting on a_i, among all landowners with the same cost of contracting.

Social versus Landowner Incentives The social surplus of contracting with landowner i is:

$$SS_i = B \cdot (1 - a_{i0}) - c_i. \tag{1.2}$$

Gains from trade occur when the incremental value of environmental services due to contracting is higher than a landowner's cost of contracting.

Landowners choose $x_i = 1$ if $p \geq c_i$. Let

$$x_i^*(p) = \mathbb{1}\{p - c_i \geq 0\} \tag{1.3}$$

be landowner i's choice to contract at price p. Equations (1.2) and (1.3) show that landowner i transacts based only on p and her contracting cost c_i, but social surplus depends also on $1 - a_{i0}$, or her additionality. p will therefore not necessarily incentivize the highest social surplus landowners to contract.

Efficient Prices and Allocations The socially-optimal uniform price solves:[8]

$$\max_p \int (B \cdot \tau(c) - c) x^*(p;c) f_C(c) \, dc, \tag{1.4}$$

[7] These include complying with mandates to purchase specific seed mixes whose costs to obtain differ across regions, effort costs to comply with specific configurations of grass planting, tree planting, or habitat establishment required in the contract, paperwork burdens to process payments, audits to manage compliance, and any taste or distaste for participating in an environmental market.

[8] A uniform price is motivated by the absence of observables. This could be because they have already been conditioned on, where equation (1.4) defines the pricing problem in a sub-population.

where the density f_C is the marginal of $F(\theta)$ on contracting costs, c_i, and $x^*(p;c) = \mathbb{1}\{p - c \geq 0\}$. The solution to this problem is equivalent to one where a quantity is chosen and an allocation is implemented with a Vickrey auction.

Equation (1.4) shows that f_C and $\tau(c)$ are sufficient statistics for social welfare and landowner choices when p is the only instrument available to allocate landowners to contracts. More generally, f_C and $\tau(c)$ are sufficient statistics for social welfare for any incentive compatible mechanism.[9] Contracting with a landowner with cost c_i increases expected social surplus when:

$$B \cdot \tau(c_i) - c_i \geq 0. \tag{1.5}$$

Our interest in this stylized framework is in when an allocation that maximizes expected social surplus, $B \cdot \tau(c) - c$, is implementable. We will refer to this allocation as the efficient allocation.[10]

1.2.2 Graphical Analysis

We analyze the efficiency of allocations in the market graphically, plotting markets with different $F(\theta)$ in Figures 1.1a and 1.1b. Each figure plots two curves: one based on f_C and one based on $\tau(c)$. The first curve is the inverse distribution function of contracting costs, $F_C^{-1}(q)$, or the *marginal cost curve* (MC), where the horizontal axis q is the share of the population ranked by contracting costs. The second curve is the value of contracting at each quantile of the distribution of contracting costs, $B \cdot \tau$, or the *contract value curve*.[11] The contract value curve lies weakly below B reflecting the possibility that $a_{i0} > 0$ for some landowners.

Each panel in Figure 1.1 displays an upwards-sloping contract value curve ($\tau'(c) > 0$). This captures the fact that landowners' expectations about future payoffs to cropping influence both c_i and a_{i0}; landowners who expect to conserve may face a low cost of accepting a contract requiring conservation. In other words, there may be adverse selection in the market.[12] Modeling adverse selection with an upwards-sloping contract value curve builds on the widely-used graphical analysis of adverse selection in insurance markets developed in Einav et al. (2010a). We emphasize,

[9] See Lopomo et al. (2023) for more details and a proof. See also Einav et al. (2010a) on the use of similar sufficient statistics for the analysis of adverse selection in competitive insurance markets

[10] We focus on this benchmark, the maximum social welfare gain achievable given knowledge of c_i and the function $\tau(c)$ in the population, following Lopomo et al. (2023), as it maximizes expected social surplus given cost reports c_i. Implementing this allocation, however, will be possible only for some $F(\theta)$.

[11] This plots $B \cdot \tau\left(F_C^{-1}(q)\right)$. We conduct the change of variables (to q) so that the areas between the contract value and marginal cost curves are interpretable as magnitudes of social welfare gains (or losses).

[12] Some may argue that using the term "adverse selection" abuses terminology. This is an example of "selection on moral hazard" defined in Einav et al. (2013a), which also includes a discussion on terminology.

however, that Figure 1.1 is for illustration: $\tau(c)$ — including whether it is upwards-sloping — and $F_C^{-1}(q)$ are to be estimated.

The vertical distance between the contract value and marginal cost curves equals $B \cdot \tau(c) - c$, or the expected social surplus of contracting with all landowners with costs equal to c. From equation (1.5), it is efficient to contract only in regions where the contract value curve lies above the marginal cost curve.

In Figure 1.1a, the efficient allocation can be implemented with socially-optimal incentives, satisfying $p^* = B \cdot \tau(p^*)$, at the intersection of the contract value and marginal cost curves. This implements social welfare gains in triangle CDE. Setting $p = B$, the social value of the conservation action, can result in inefficient contracting and social welfare losses (triangle EFG). Socially-optimal incentives therefore require knowledge of both f_C and $\tau(c)$: the distribution of contracting costs and heterogeneous impacts of contracting along this distribution.

In Figure 1.1b, the efficient allocation cannot be implemented. In fact, in Figure 1.1b, it is socially-optimal not to offer a market, despite the existence of landowners for whom contracting is socially desirable. In the distribution of landowner types illustrated in Figure 1.1b, the contract value curve lies below the marginal cost curve at low contracting costs (low q). This represents landowners that have low but positive costs of contracting — due to some option value of cropping and/or hassle costs — but a high likelihood of conserving without the contract. In this market, a regulator cannot implement the efficient allocation (triangle EFG), as any incentive that is attractive for landowners in triangle EFG is also attractive for landowners in CDE, where losses in CDE outweigh gains in EFG.

The difference between equations (1.2) and (1.3) causes the inefficiency in Figure 1.1b. The regulator can only affect allocations based on landowner costs, c_i, and the incentive p, but social surplus depends also on the impact of contracting, or a_{i0}. In contrast to standard markets, the relationship between social surplus and landowner costs may not be monotonically decreasing. Because $B \cdot \tau(c) - c$, the vertical distance between the contract value and marginal cost curves, crosses zero from below in Figure 1.1b, no mechanism can implement the efficient allocation (triangle EFG) as it would require an allocation rule that is not monotonically decreasing in landowner costs (see Myerson (1981a); Lopomo et al. (2023)). The difference between Figures 1.1a and 1.1b thus has implications beyond the illustrative posted prices mechanism in this Section. In a procurement auction with a quantity constraint — our empirical setting — if $B \cdot \tau(c) - c$ is not decreasing in c, the efficient allocation (subject to the quantity constraint) may not be implementable by any mechanism. In Figure 1.1a, $B \cdot \tau(c) - c$ is decreasing in c, but in Figure 1.1b, it is not.

If contracts are traded in competitive markets, which we term offset markets,[13] adverse selection

[13] We will refer to competitive markets with price-taking buyers as offset markets, though we model buyers as

can also prevent the competitive equilibrium price from implementing an efficient allocation, even when it is implementable with the price that solves equation (1.4) (Akerlof, 1970a). Price-taking buyers in the market take expectations over the additionality of *all* market participants, not only those contracting at the margin. We define a competitive market price p^c by the equilibrium condition: $p^c = \mathbb{E}[B \cdot \tau(c_i) \,|\, c_i \leq p^c]$.[14] Figure 1.1c adds the curve defined by $\mathbb{E}[B \cdot \tau(c_i) \,|\, c_i \leq p]$ to the population of landowners presented in Figure 1.1a.[15] Its intersection with the marginal cost curve defines the competitive market equilibrium, which differs from the socially-optimal price. In the presence of adverse selection, trade in competitive (offset) markets will be limited and efficient contracting, with social welfare gains represented in triangle EFG, will not occur.

Empirical Questions Figure 1.1 illustrates that the welfare implications of additionality depend on f_C and $\tau(c)$. The goal of our empirical analysis is therefore to estimate f_C and $\tau(c)$. But this stylized model was limited in its tools. Our empirical analysis will include a richer set of contracts and observable characteristics. We will then investigate both the possibility of social welfare losses when market incentives do not implement the efficient allocation and social welfare gains from alternative market designs.

1.3 Setting and Data

1.3.1 The Conservation Reserve Program

Our empirical setting is the Conservation Reserve Program (CRP), a Payments for Ecosystem Services (PES) scheme incentivizing conservation on agricultural land administered by the United States Department of Agriculture (USDA). Established in 1985, the CRP pays landowners between $1.6 and $1.8 billion per year to retire erodible and other environmentally sensitive cropland and adopt additional conservation actions for a contract duration of 10 years. The CRP is one of the largest and most mature PES schemes in the world. It is also a major source of expenditures on environmental policy in the United States; the CRP is one of several conservation programs at the USDA with a combined budget of $8 billion.[16] Moreover, the structure of the CRP and its

valuing all of the social benefits of the conservation action B, not only an emissions offset.

[14] We focus on the social welfare losses from supply-side adverse selection. We abstract away from the possibility that buyer valuations may diverge from B, that buyers may not know the distribution $F(\theta)$, or that buyers may not value additionality.

[15] This curve is defined as $\int_0^q B \cdot \tau\left(F_C^{-1}(\tilde{q})\right) d\tilde{q}$.

[16] By comparison, the Superfund program and Weatherization Assistance Programs have annual budgets of $1.1 billion, and the total Environmental Protection Agency (EPA) budget is $12 billion. See the USDA FY 2023 Budget Summary, FY 2023 EPA Budget in Brief, and NASCP Weatherization Assistance Program Funding Report for FY 2019 for more details.

incentivized activities are similar to other government financed PES schemes,[17] to offset contracts traded in voluntary markets,[18] and most specifically, to a burgeoning private agricultural offset market in the US. There is substantial policy interest in growing this market. The Growing Climate Solutions Act of 2021 includes provisions for the creation of a USDA-regulated agricultural offset market, in which CRP-style contracts would be sold to private buyers.[19]

Unlike the uniform pricing problem in Section 1.2, the USDA awards CRP contracts via a complex auction mechanism. This adds richness to both the strategic and contracting environment that we will leverage empirically. Under the CRP's General Enrollment mechanism, eligible landowners bid for heterogeneous contracts in a discriminatory, asymmetric, scoring auction.[20] Contracts are differentiated by conservation actions that "top up" the base action of land retirement. These actions include planting specific grass mixes, planting or maintaining trees, and establishing or restoring pollinator or rare habitats.

Bids are scored according to a known scoring rule that awards bidders points for the environmental sensitivity of their land, their chosen contract (described above), and their bid rental rate, a payment per acre per year. Characteristics of the land that determine points for environmental sensitivity include erodibility, importance for habitats, potential for water and air pollution, and carbon sequestration potential. Bid rental rates are subject to a bid cap based on the average land rental rate in the county and soil productivity estimates. Appendix 1.10.1 describes the auction mechanism and scoring rule in more detail.

The aggregate acreage to be awarded contracts is determined by Congress in the Farm Bill, which in turn determines the threshold score for contract awards. All bidders with scores above the threshold score are awarded a contract.[21] Given uncertainty in both the aggregate acreage and their opposing bidders, bidders face uncertainty over the threshold score at the time of bidding. Bids are prepared with the assistance of staff at Farm Services Agency county offices, who help landowners trade-off different contracts and bid rental rates.

Auctions occur once every 1-4 years. Landowners are eligible to bid if they meet erosion standards, are in a national or state conservation priority area, and either had cropped at least four years in a 5-10 year window preceding the auction or are re-enrolling CRP land.[22] This eligibility

[17] See Kinzig et al. (2011) and Salzman et al. (2018) for overviews.

[18] Over 50% of contracts traded in voluntary offset markets are related to land use and management. See the Voluntary Registry Offsets Database at the Berkeley Carbon Trading Project for more details.

[19] See S. 1251 and H.R. 2820 for more details.

[20] In addition to the General Enrollment mechanism, there is a posted-price Continuous Enrollment mechanism for targeted land. Historically, 75% of CRP acreage has been contracted via the General Enrollment mechanism (Hellerstein, 2017a).

[21] There is an additional constraint that no more than 25% of a county's total acreage can be in a CRP contract. This constraint essentially never binds.

[22] The fact that eligibility is determined in a window five years preceding bidding is designed to eliminate any

requirement is designed to limit participation to additional landowners. Landowners face steep penalties, refunding all payments since enrollment plus a 25-percent penalty, if they exit early or fail to comply with the rules of the program.[23]

Research quantifying the value of the CRP has documented improvements in wildlife habitats, erosion control, water quality, and carbon sequestration from cropland retirement (Feather et al., 1999; Hansen, 2007a; FAPRI-MU, 2007; Allen and Vandever, 2012; Johnson et al., 2016a; Hellerstein, 2017a). However, these analyses are typically conducted using models that ignore counterfactual land use. Perhaps motivated by eligibility requirements designed to restrict to additional landowners, the scoring rule is constructed under the assumption that land would crop in the absence of the program (Claassen et al., 2018). Because the primary environmental gains from the CRP accrue from land retirement, relative to cropping, the possibility that some landowners conserve absent the CRP $(a_{i0} = 1)$ presents the additionality concern.

1.3.2 Data

Our dataset links bids to a panel of landowners' land use to measure additionality.

Data on Bids We obtain data on all components of the bid, including the bid rental rate, the bid contract, and the characteristics of landowners that impact the score. Our data cover all seven auctions that occurred from 2009 to 2021. We also obtain data on all CRP contracts.

Each landowner is defined by a collection of fields, or Common Land Units, the smallest agricultural unit with a common land use. CRP contracts typically cover only a subset of a landowner's total fields. Our data include the geolocation of all bidding landowners for all auctions as well as identifiers for the specific fields offered into the mechanism ("bid fields") for one auction (in 2016).

Data on Land Use We link bidders, and for the purpose of comparison, agricultural non-bidders, to a panel of land use outcomes. The primary land use outcome of interest is whether land is cropped versus retired, as this is the behavior incentivized by the CRP. We use both remote sensing and administrative datasets due to their complementary strengths.

Our primary dataset is the Cropland Data Layer (CDL), a remote-sensing product from the National Agricultural Statistics Service (NASS). This dataset provides land cover classifications at

perverse incentives to crop land to in order to become eligible or maintain eligibility for the CRP. Activities in the 1-5 years preceding bidding have no impact on CRP eligibility.

[23]The USDA has occasionally allowed for voluntary contract extensions or automatic re-enrollment. No such initiatives were implemented during our main period of study.

30m by 30m resolution (roughly a quarter acre) from 2009-2020. The binary indicator of crop versus non-crop — our primary outcome of interest — is rarely misclassified (Lark et al., 2021).[24] However, as in other satellite-derived products, non-classical measurement error can generate biases in assessing land-use change (Torchiana et al., 2022; Alix-Garcia and Millimet, 2022).

Our second dataset is field-level administrative data on land use that agricultural landowners report to the USDA in "Form 578" for 2013-2019. These data are accurate and comprehensive for cropped land because crop insurance payouts are dependent on these reports, but have two limitations. Landowners are only incentivized to report Form 578 if fields are insured by crop insurance, and landowners with CRP contracts are mechanically classified as non-cropped.

Our final land-use dataset is a collection of high-resolution satellite imagery (1m) of contracted land collected under the National Agriculture Imagery Program (NAIP) from 2017-2021. We use these images to observe and confirm compliance with CRP rules.

While accurate to assess agricultural land retirement — the main incentivized activity of the CRP — these datasets cannot measure the different "top-up" actions that differentiate the heterogeneous contracts in the mechanism (e.g. specific species). Our main estimates of additionality will focus on the measure that we can observe and the principal goal of the program: the binary outcome of retiring versus cropping land.

Appendix 1.10.2 provides more details about agricultural land units, the construction of our dataset, and the use of aerial photographs to confirm compliance with CRP rules.

Summary Statistics Table 1.1 presents summary statistics. Columns (1)-(2) summarize all agricultural landowners in the US, including CRP-eligible and ineligible landowners. Columns (3)-(4) summarize all land among bidders in our sample and columns (5)-(6) summarize bid fields.

Panel A presents land use outcomes in the year prior to bidding. Approximately 21% of bidders' land is cropped (18-21% on bid fields) versus 28-30% nationwide. The majority of the remainder is in natural vegetation and grassland. Corn and soybean cultivation account for two-thirds of all cropping. The remote sensing and administrative data generally align, but do not match exactly.

Bidders have lower USDA-constructed estimates of soil productivity (Panel B), are larger, and are more environmentally sensitive as measured by the scoring rule than the average agricultural landowner. These differences, along with the differences in land use in Panel A, are likely driven in part by eligibility requirements that columns (1) and (2) are not conditioning on.

[24] The superclass accuracy of cropland in the Cropland Data Layer has user (probability that a classification of crop is true crop) and producer (probability that true crop is classified as crop) accuracy of over 95% from 2008-2016 (Lark et al., 2021)

The average bidder offers 84.1 acres into the CRP mechanism (33% of a bidder's land) for a rental rate per acre per year of $83. Two-thirds of bidders bid on a contract that includes a grassland-planting action, 21% choose a wildlife habitat action, and 12% choose a tree-planting action. 70% of bidders are re-contracting after their initial 10-year contract expired.[25] 81% of bidders are awarded contracts across the auctions in our sample. The average auction includes 36,763 bidders.

1.4 Evidence: Additionality and Asymmetric Information

In this section, we estimate the extent of additionality in the CRP and test for heterogeneity in and asymmetric information about additionality.

1.4.1 Regression Discontinuity Estimates of Additionality

Estimates of additionality are important inputs into the evaluation and design of markets for environmental services but require a credible empirical strategy. We exploit the sharp discontinuity in CRP contract awards at the winning score threshold to evaluate the treatment effect of a CRP contract in a regression discontinuity (RD) design.

Empirical Strategy Our RD specification pools all auctions in the sample, normalizes each landowner's score relative to her auction's win threshold, and evaluates how land use outcomes differ over time around this threshold.

Our main specification estimates, for landowner (or bidder) i, in auction g, and year t:

$$y_{igt} = \beta_{r(i,t)} \cdot \mathbb{1}\{S_{ig} \geq \underline{S_g}\} + f_{r(i,t)}(S_{ig} - \underline{S_g}) + \nu_{igt}, \tag{1.6}$$

where S_{ig} is landowner i's score in auction g, $\underline{S_g}$ is the winning score threshold in auction g, $r(i,t) = t - t_{g(i)}$ normalizes the time dimension relative to the year of each auction $(t_{g(i)})$, and $f_{r(i,t)}(S_{ig} - \underline{S_g})$ are relative-year-specific local-linear regressions in the MSE-optimal bandwidth (Calonico et al., 2014) allowed to differ for positive and negative values of $S_{ig} - \underline{S_g}$. $\beta_{r(i,t)}$ estimates time-varying RD coefficients around the year of the auction. We also estimate the following pooled specification:

$$y_{igt} = \beta \cdot \mathbb{1}\{S_{ig} \geq \underline{S_g}\} + f(S_{ig} - \underline{S_g}) + \nu_{igt}. \tag{1.7}$$

[25] Re-contracting bidders are treated identically to new bidders by the scoring rule.

Restricted to $r(i,t) \leq 0$, equation (1.7) provides a test of validity. Restricted to $r(i,t) > 0$, β provides an estimate of the treatment effect at the margin of contract awards pooled over auctions and post-auction years.

We estimate equations (1.6) and (1.7) at the landowner level to accommodate the possibility of spillovers across bid and non-bid fields. We cluster standard errors at the landowner level.

Validity The RD design is valid under the assumption that bidders possess information about the winning score threshold's distribution, but not its exact ex-post realization. Testing this assumption translates to standard smoothness and manipulation tests for RD analyses (McCrary, 2008); if bidders knew the threshold score, they would optimally bid just above it. Figure 1.2a presents a histogram of the score distribution normalized relative to the threshold score, $S_{ig} - \underline{S_g}$, or the running variable of the RD. Bidders with positive values are awarded contracts and bidders with negative values are not. Figure 1.2a confirms the lack of bunching at the winning score threshold. Figure 1.2b also shows no differential cropping at the discontinuity before the auction, plotting (binned) raw data and estimates of equation (1.7) for $r(i,t) \leq 0$.

Interpretation of equations (1.6) and (1.7) also requires an estimate of the magnitude of the first stage. Figure 1.2c plots the share of bidders with a CRP contract after the auction around the award threshold (equation (1.7) for $r(i,t) > 0$) and demonstrates a first stage close to one. We will therefore interpret the RD coefficients from equations (1.6) and (1.7) as the impact of a CRP contract.

Results Figure 1.3 presents raw data and estimates of equation (1.7) for $r(i,t) > 0$. As the CRP's primary goal is to incentivize agricultural land retirement, our outcome of interest in Figure 1.3a is the share of each bidder's land that is cropped, which we measure in the remote sensing data. Figure 1.3a demonstrates that CRP contracts impact land use at the margin of contract awards. Landowners above the score threshold, who are awarded a contract, crop eight percentage points less land than the marginal landowners who are not awarded contracts. This land is instead put into natural vegetation and grassland (trees, shrubs, wetlands and grasses), as incentivized by the CRP (Figure 1.3b). Because we present estimates at the bidder level, cropping is not zero for winners who typically only contract on a subset of their land.

To analyze the time path of effects and estimate the extent of additionality, Figure 1.4 presents coefficient estimates of $\beta_{r(i,t)}$ from equation (1.6) and compares them to a 100% additional ($a_{i0} = 0$ for all i) benchmark. We estimate equation (1.6) using both the remote sensing data (used in Figures 1.2b and 1.3) and the administrative data to ensure that results are consistent across the two datasets. The 100% additional ($\tau = 1$) benchmark in Figure 1.4 is calculated as the share

of marginal bidders' land in a CRP contract. If contracting induced all bidders to change land use, treatment effects would equal the $\tau = 1$ line on Figure 1.4. Dashed lines represent pooled post-period estimates (equation (1.7)) in each dataset.

Figure 1.4 presents four facts. First, consistent with the pre-period placebo test in Figure 1.2b, coefficient estimates are zero before the auction. Because the estimates in Figure 1.4 are relative-year-specific RD coefficients, pre-period effects are identified in levels. Second, post-period effect sizes and time-trends are similar in both datasets, confirming that results are not driven by either non-classical measurement error in the remote sensing data or misreporting in the administrative data. Third, while treatment effects grow in the first two years as land transitions, effects are constant over the ten year contract period. Opportunities to rebid, which would cause treatment effects to decrease over time, do not drive down average treatment effects.[26]

Finally, the main result in Figure 1.4 is that over the 10-year contract, the magnitude of the treatment effect of a CRP contract on land use is substantially smaller than the $\tau = 1$, or 100% additional, benchmark. Figure 1.4 demonstrates that approximately one in four bidders is additional. Conversely, three of four bidders conserve even absent a CRP contract. Figure 1.4 provides evidence on the relevance of additionality in the CRP.

Table 1.2 summarizes results from Figures 1.2, 1.3, and 1.4, presenting estimates for the pooled specification (equation (1.7)) in both datasets. The main results in Table 1.2 quantify the additionality estimates from Figure 1.4: depending on the specification and data, we estimate rates of additionality at the margin of contract awards between 21% and 31%, with a mean and median effect size of 26%. Panel B presents estimates on other land use outcomes.[27]

Discussion Estimates of additionality at the margin provide information about the contract value ($B \cdot \tau$) curve in Figure 1.1. First, it lies below B, as many landowners are not additional. The results in Figure 1.4 also highlight the need to estimate, rather than assume, the $\tau(c)$ function. If alternatively, costs and additionality could be summarized by a single index, in which bidders with positive costs are additional and bidders with costs equal to zero are not, then at the margin additionality should be either zero or one. Estimates of additionality at the margin reject both of these hypotheses.[28] This will motivate our modeling and estimation approach: neither estimates of

[26] We observe limited rebidding. Appendix Figure A.7 plots the hazard rate of rebidding following a failed initial bid. Even five years following the initial bid, only approximately 20% of losers have rebid and fewer than 15% have won (despite multiple opportunities). This is consistent with the magnitude of the first stage in Figure 1.2c and the institutions of the setting. The CRP is so mature that the General Enrollment mechanism is shrinking over time; acreage contracted in later auctions is lower than acreage contracted in earlier auctions over our time period of analysis.

[27] Appendix Figures A.4, A.5, and A.6 present corresponding RD figures. Appendix Table A.6 replicates Table 1.2 restricting to bids of more than five acres, following Lark et al. (2017).

[28] This interpretation is slightly complicated by bidder asymmetry, contract choices, and the pooling of auctions with different thresholds. Appendix Table A.5 presents RD estimates split by the location of the win threshold param-

costs nor additionality alone are sufficient to estimate social welfare under current or counterfactual market designs.

Mechanisms We argue that the estimates in Figure 1.4 are driven by heterogeneous land use absent the contract (a_{i0}) specifically on the land bid into the mechanism. Panel C of Table 1.2 (and Appendix Figure A.4) documents the absence of any spillovers onto non-bid fields. This could occur either via a leakage mechanism, by which landowners reduce cropping on bid fields but increase it on other fields, or if there are complementarities to cropping multiple fields. We observe no evidence in support of either of these hypotheses.

In theory, both conservation without a CRP contract and cropping with a CRP contract (non-compliance) could contribute to the result in Figure 1.4. We assess the CRP's compliance regime by inspecting high resolution aerial photographs of over 1,000 contracted fields.[29] As described in more detail in Appendix 1.10.2, we find no evidence of non-compliance.

Beyond the RD Together, these two results — no spillovers and compliance — provide a basis for empirical analysis beyond the RD. Among rejected bidders, we observe a realization of each bidder's a_{i0} on bid fields. With knowledge that $a_{i1} = 1$, we therefore observe $1 - a_{i0}$ for each landowner that loses the auction. In other words, we simplify to a "selective labels" problem (Lakkaraju et al., 2017; Chan et al., 2022b; Arnold et al., 2022).

1.4.2 Testing for Asymmetric Information

Empirical Strategy We use observations of landowners' realized additionality ($1 - a_{i0}$) and bids to conduct a test for asymmetric information in the spirit of Chiappori and Salanie (2000) and Hendricks and Porter (1988). We estimate the following regression specification:

$$1 - a_{i0} = \beta \cdot \mathbf{b}_i + \pi \cdot \mathbf{z}_i + h(\mathbf{z}_i^s) + \nu_i, \tag{1.8}$$

where $1 - a_{i0}$ is measured as the share of landowner i's bid fields that are cropped, observed only for landowners without a contract award (those rejected by the auction), $\mathbf{b}_i$ represents characteristics of i's bid, $h(\mathbf{z}_i^s)$ are controls for characteristics that enter the scoring rule, and $\mathbf{z}_i$ are other landowner characteristics. Every specification of equation (1.8) includes controls for the scoring rule, which

eterized by the amount a bidder would need to bid for the base contract to achieve $\underline{S_g}$ and finds that $0 < \tau < 1$ across groups.

[29] We use aerial photographs because any measurement error in the remote sensing classifications will mechanically bias toward finding non-compliance.

impacts the strategic environment facing bidders. These include estimates of groundwater quality, surface water quality, wind and water erosion (deciles), air quality impacts, and whether a bidder is in a Wildlife Priority Zone or Air Quality Zone. We restrict to the one auction where we observe the delineations of bid fields (the 2016 auction); this is required to construct $1 - a_{i0}$. This auction is also the most restrictive auction in our sample: $1 - a_{i0}$ is observed for 82% of bidders. We will address the complication that equation (1.8) is estimated in the selected sample of rejected bidders in Section 1.5.

Following the logic of Chiappori and Salanie (2000) and Hendricks and Porter (1988), a positive correlation between bids and $1 - a_{i0}$ is indicative of asymmetric information about additionality. In the context of the stylized model in Section 1.2, a positive correlation implies an upwards-sloping contract value curve, or adverse selection in the market.

Results We first document evidence of adverse selection in the market. Figure 1.5a presents a binned scatterplot of the correlation between additionality and the bid per acre-year (the bid rental rate), residualized of $h(\mathbf{z}_i^s)$. Figure 1.5a demonstrates substantial heterogeneity in additionality and a systematic positive relationship between higher bids — reflecting higher costs of contracting — and additionality. The interpretation of Figure 1.5a is intuitive: bidders with low costs of contracting have low costs in part because of private information that they would be likely to conserve even without a CRP contract. Figure 1.5b shows that bids remain correlated with additionality, capturing residual private information, conditional on other observables including prior land use interacted with estimates of the soil productivity of the bidders' land

Next, we show that choices of contracts in the mechanism are systematically correlated with additionality. Figure 1.5c replaces $\mathbf{b}_i$ with a vector of chosen contract indicators — the submitted bid on the menu of contracts — and documents substantial adverse selection (low additionality) on tree-related contracts, relative to the base category of introduced grasses. Figure 1.5c highlights that contract choices reveal information about additionality and that alternative menus of contracts may lead to different outcomes in the market.

Finally, we present evidence that observable characteristics are predictive of additionality. Figure 1.5d plots additionality by decile of landowner predicted soil productivity, conditional on $h(\mathbf{z}_i^s)$ but excluding any endogenous bid choices from the regression specification. These estimates of soil productivity are collected by the USDA and are designed to approximate the amount that a landowner would be able to earn on a given parcel of land. This characteristic is an ideal predictor of additionality in theory, and Figure 1.5d demonstrates that it is predictive of additionality in practice. This result highlights the potential to differentiate incentives using this characteristic, which is not currently incorporated in the scoring rule.

Discussion The analysis in Sections 1.4.1 and 1.4.2 provide evidence on the extent of additionality, the presence of asymmetric information in the market, and the availability of tools to differentiate landowners by their additionality. However, the welfare and market design implications of these facts require a quantitative economic framework. In the next section, we develop an empirical approach to obtain the sufficient statistics for welfare presented in Section 1.2. Relative to that stylized set-up, our empirical model will incorporate heterogeneity across contracts and observable characteristics to capture a richer empirical setting for analysis and market design.

1.5 Empirical Model of Bidding and Additionality

We develop a joint model of bidding and additionality. We use this model to estimate (i) the distribution of landowner costs of contracting over a menu of contracts and (ii) additionality as a function of landowner costs and observable characteristics. Together with estimates of the social benefits of CRP actions, which we take from the literature, the model facilitates analysis of social welfare under current and counterfactual market designs.

Landowners are characterized by a vector of private costs and bid on discrete contracts, differentiated by heterogeneous conservation actions, in response to a non-linear scoring rule. Landowners also differ in their additionality, which we model with a conditional expectation function that depends on both observable characteristics and bidders' vector of costs. Our empirical strategy first uses the optimality of bidding in the auction to estimate bidders' costs by revealed preferences and then estimates expected additionality as a function of costs and landowner characteristics by matching moments of the observed joint distribution of land use, landowner characteristics, and optimal bids (presented in Section 1.4).

1.5.1 Model

Landowners There are N landowners, indexed by i, and J contracts, indexed by j. Each landowner is characterized by (i) her costs (c_i, κ_i) for $\kappa_i = \{\kappa_{ij}\}$, and (ii) her action a_{i0} absent the CRP. Extending the model in Section 1.2, re-define $F(\theta)$ as the cumulative distribution function of landowner types $\theta_i = ((c_i, \kappa_i), a_{i0})$ and $F_{c,\kappa}$ as the marginal on (c_i, κ_i).

The vector (c_i, κ_i) defines landowner i's costs of contracting. A landowner's cost of contract j is $c_i + \kappa_{ij}$, where c_i is the base cost of contracting, common across contracts, and κ_{ij} is the top-up cost associated with contract j. We assume (c_i, κ_i) are drawn independently across landowners conditional on observables $\mathbf{z}_i$.

It will again be useful to define the function:

$$\tau(\mathbf{z}_i, c_i, \kappa_i) = \mathbb{E}\left[1 - a_{i0} \,|\, \mathbf{z}_i, c_i, \kappa_i\right], \tag{1.9}$$

or the expected difference in conservation with a CRP contract versus without a CRP contract given observable characteristics $\mathbf{z}_i$ and landowner costs (c_i, κ_i).

Auction Landowners (bidders) submit a two-part bid $\mathbf{b}_i = (r_i, \mathbf{x}_i)$. $\mathbf{x}_i$ is a contract vector, with $x_{ij} = 1$ if the j-th contract is chosen and $x_{ij} = 0$ otherwise. Landowners choose a single contract so $\sum_j x_{ij} = 1$. If landowner i wins the auction, $\mathbf{b}_i$ describes the terms of her CRP contract: she performs the action defined in $\mathbf{x}_i$ and receives a payment of r_i dollars per acre-year. Each bid $\mathbf{b}_i$ is converted into a score according to a known scoring rule $s(\mathbf{b}_i, \mathbf{z}_i^s)$. All landowners above a winning threshold score $\underline{S}$ are awarded a contract.

Landowner i forms beliefs about the probability of winning the auction with a score S given uncertainty over her competitors and the acreage limit of the auction.[30] We assume that landowner i does not observe the number or characteristics of her competitors, and all landowners form the same beliefs about the distribution of the threshold score $\underline{S}$.[31] Define $G(S) = \Pr\{S \geq \underline{S}\}$.

Optimal Bidding Each landowner i solves:

$$\mathbf{b}_i^* = \underset{(r,\mathbf{x})}{\operatorname{argmax}} \left\{ \underbrace{(r - c_i - \mathbf{x} \cdot \kappa_i)}_{\text{Payoff to } i \text{ conditional on bid } (r,\mathbf{x})} \times \underbrace{G\left(s\left((r,\mathbf{x}), \mathbf{z}_i^s\right)\right)}_{\text{Probability of } i \text{ winning with bid } (r,\mathbf{x})} \right\}, \tag{1.10}$$

where a landowner chooses her bid $\mathbf{b}_i = (r_i, \mathbf{x}_i)$ to maximizes her payoff conditional on winning, multiplied by the probability of winning with that bid, given her costs (c_i, κ_i).

Additionality In the contract period, landowners make land use decisions. If awarded a contract, $a_{i1} = 1$. If not, landowners choose a_{i0}, which is not contractible. At the time of bidding, equation (1.9) is the population expectation of $1 - a_{i0}$ for landowners with observable characteristics $\mathbf{z}_i$ and contracting costs (c_i, κ_i). We estimate the function $\tau(\mathbf{z}_i, c_i, \kappa_i)$, instead of modeling the choice of a_{i0}, because $\tau(\mathbf{z}_i, c_i, \kappa_i)$ and $F_{c,\kappa}$ are sufficient statistics for social welfare under current and alternative market designs (see Section 1.2).

[30] The acreage limit is determined by Congress in the Farm Bill. Appendix Figure A.9 documents evidence consistent with quantity uncertainty.

[31] This is motivated by the fact that bidding is decentralized among thousands of bidders across the US.

Remarks Landowners compete on both r and $\mathbf{x}$ in equation (1.10). This captures the importance of competition on contracts in reality and allows for counterfactuals that re-design the menu or incentives across contracts in the scoring rule.[32] Although landowners are competing on multiple dimensions, whether the landowner wins against her competitors depends only on the choice of score. The bidding problem can be solved as an "inner problem" of a single-agent discrete choice and an "outer problem" of a one-dimensional game, building on Asker and Cantillon (2008b) and Che (1993). Each score induces a menu of payoffs from winning the auction for each contract for each bidder (illustrated in Appendix Table A.3). The resulting discrete choice problem is the bidder's "inner problem." Then, the choice of the optimal score, given the optimally chosen contract, defines the bidder's "outer problem."

The bidding problem in equation (1.10) incorporates two important simplifications. First, bidding is costless and there is no selection into bidding.[33] Second, the bidding problem in equation (1.10) is static. The CRP is so mature that the quantity procured in the auction is in decline. This limits the option value to rebid, and is reflected in the fact that the vast majority of bidders do not re-bid upon losing (see Appendix Figure A.7). However, in a dynamic framework, the cost parameters estimated from equation (1.10) can be interpreted as *pseudo-costs* that are the result of mapping a dynamic program with sequential auctions into a static game (Jofre-Bonet and Pesendorfer, 2003). Counterfactuals that do not condition on prior actions and hold the sequence of future auctions fixed will not be biased by the static formulation of equation (1.10).

Although the mechanism is more complex, the market failure is the same as in Section 1.2. Landowner choices and allocations depend only on the scoring rule $s(\mathbf{b}_i, \mathbf{z}_i^s)$ and costs of contracting (c_i, κ_i), but the efficient allocation depends also on $\tau(\mathbf{z}_i, c_i, \kappa_i)$.

1.5.2 Identification and Estimation

Identification Because observed bids are discrete, we cannot invert them using bidders' first order conditions to point identify costs as in Guerre et al. (2000a). We instead obtain inequalities from the observed $\mathbf{b}_i^* = (r_i, \mathbf{x}_i)$ revealed preferred from optimal bidding in equation (1.10) that define identified sets containing the true values of (c_i, κ_i) (Agarwal et al., 2023). Instruments in the scoring rule $s((r, \mathbf{x}), \mathbf{z}_i^s)$, which vary the relative payoffs across contracts $\mathbf{x}$ but are conditionally independent of costs, narrow the bounds on the identified sets. Variation in $s((r, \mathbf{x}), \mathbf{z}_i^s)$ traces out

[32] The EBI Factsheets provided to landowners state: "The single most important producer decision involves determining which cover practice to apply to the acres offered. Planting or establishing the highest scoring cover mixture is the best way to improve the chances of offer acceptance."

[33] This is a simplifying assumption. Hellerstein (2017a) makes the point that many eligible landowners do not bid. We assume that non-bidders are invariant to changes in the mechanism.

the distribution of (c_i, κ_i) conditional on observable characteristics $\mathbf{z}_i$, $F_{c,\kappa|\mathbf{z}}$. Appendix Figure A.10 provides a graphical explanation; Agarwal et al. (2023) provides a proof.

For identification of $\tau(\mathbf{z}_i, c_i, \kappa_i)$, suppose bidders truthfully report (c_i, κ_i) to the mechanism, but no bidders are awarded contracts. $F(\theta)$ is point identified from observing the joint distribution of (c_i, κ_i) in bids and a_{i0} in the land use data. Our setting differs from this ideal: (i) the discrete choice in equation (1.10) results in only identified sets for (c_i, κ_i) from observed bids, and (ii) a_{i0} is not observed for the 18% of bidders who win the auction and are awarded a contract. Identification of the function $\tau(\mathbf{z}_i, c_i, \kappa_i)$ uses the observed joint distribution of $1 - a_{i0}$, characteristics $\mathbf{z}_i$, and optimal bids $\mathbf{b}_i^* = (r_i, \mathbf{x}_i)$ and instruments that shift $s((r, \mathbf{x}), \mathbf{z}_i^s)$. $\tau(\mathbf{z}_i, c_i, \kappa_i)$ is the conditional expectation function that rationalizes this joint distribution. With full support, instruments in the scoring rule that shift payoffs across contracts and the probability of winning replicate the "ideal experiment" described above.

We use three sources of variation in the scoring rule as instruments. Two shift relative payoffs across contracts (illustrated in Figure A.11). One shifts only the level of the score. We assume that all three sources of variation are conditionally independent of (c_i, κ_i) and a_{i0}. The first source of variation is a mid-mechanism policy change: after bids were initially collected in 2021, Climate Smart Practice Incentives — additional payments dependent on contracts' carbon sequestration potential — were announced and bids were recollected under the new scoring rule. We obtained the bids submitted in both the interim and final mechanisms, which provides variation in the relative payoffs across contracts for the same bidders and same contract period.[34] The second source of variation comes from the fact that bidders in Wildlife Priority Zones (WPZs) face different payoffs across contracts both cross-sectionally and over time. The third source of variation is similar, whether a bidder is in an Air Quality Zone (AQZ), but shifts only the level of the score. The use of these instruments are justified by the fact that WPZ and AQZ status are based on state priorities and the sensitivity of wildlife or the importance of air quality, not characteristics of landowners or their land.

Parameterization Although with sufficient variation in the scoring rule, the model is non-parametrically identified, to take it to the data, we parameterize $F_{c,\kappa|\mathbf{z}}$ and $\tau(\mathbf{z}_i, c_i, \kappa_i)$. Landowners make a discrete choice across contracts with heterogeneous features, so we parameterize (c_i, κ_i) with a characteristics model:

$$c_i \sim N\left(c(\mathbf{z}_i), \sigma_c^2(\mathbf{z}_i)\right) \quad \kappa_{ij} = p_j(\mathbf{z}_i) + u_j(\mathbf{z}_i) + \varepsilon_{ij} \quad \varepsilon_{ij} \overset{iid}{\sim} N\left(0, \sigma_\kappa^2(\mathbf{z}_i)\right). \tag{1.11}$$

[34] We can use this policy experiment to directly test that landowners are competing on contracts; 8% of landowners change their contract choice under the new scoring rule.

c_i and κ_{ij} are drawn from independent normal distributions with means and variances that depend on observable characteristics, $\mathbf{z}_i$. κ_{ij} costs are differentiated by contract features, p_j and u_j. p_j defines mean costs for a vector of primary covers, which vary by the left-most four categories in Figure 1.5c, relative to the base category of introduced grasses (normalized to zero). u_j is a vector of upgrade covers, which varies by the right-most two categories in Figure 1.5c plus the no-upgrade option, normalized to zero. The parameterization in equation (1.11) parsimoniously captures differences across contracts.[35]

We also parameterize

$$\tau(\mathbf{z}_i, c_i, \kappa_i) = \pi \cdot \mathbf{z}_i + \beta \cdot c_i + \alpha \cdot \kappa_i. \tag{1.12}$$

This specification allows $\tau(\mathbf{z}_i, c_i, \kappa_i)$ to depend on observable characteristics, $\mathbf{z}_i$, and unobserved bidder costs (c_i, κ_i), where we align the dimension of α with the primary and upgrade parameterization of κ_{ij}, i.e. we impose that $\alpha_j = \alpha_{j'}$ if $p_j = p_{j'}$ and $u_j = u_{j'}$.

Estimation Estimation proceeds in three steps and closely follows the identification argument. In the first step, we estimate landowner beliefs $G(S)$ via simulation. In the second step, we estimate the parameters of $F_{c,\kappa|\mathbf{z}}$ via revealed preferences in observed optimal bids (equation (1.10)). In the third step, we estimate $\tau(\mathbf{z}_i, c_i, \kappa_i)$ using the Step 2 estimates of $F_{c,\kappa|\mathbf{z}}$ and optimal bidding in equation (1.10) to simulate and match land-use moments of the joint distribution of $1 - a_{i0}$, $\mathbf{b}_i^* = (r_i, \mathbf{x}_i)$, and $\mathbf{z}_i$. Steps 1 and 2 are a common approach to the empirical analysis of auctions (Guerre et al., 2000a; Hortaçsu and McAdams, 2010; Agarwal et al., 2023) and Steps 2 and 3 are a common approach to the empirical analysis of selection markets (Bundorf et al., 2012a; Tebaldi, 2022). Appendix 1.10.4 provides additional details.

Step 1: Simulate $G(S)$ First, we estimate $G(S)$ by simulation following Hortaçsu (2000); Hortaçsu and McAdams (2010). We fit Beta distributions to the number of bidders and acreage limits across auctions. We supplement our primary dataset with additional historic data on acreage limits and the numbers of bidders for all auctions from 2000 to 2021 to fit these distributions. Then, we simulate the numbers of bidders and the acreage thresholds and re-sample from the observed joint distribution of the scores and acreages of bidders within each auction.

Step 2: Estimate Costs The next step estimates $F_{c,\kappa|\mathbf{z}}$ using the parameterization in (1.11) and the optimality of observed bids from equation (1.10). We classify bidders into 32 categories of

[35] Landowners face a discrete choice over each of the primary and upgrade covers, but primary and upgrade covers can be combined. There are 36 total possible contracts, reflecting finer categorizations of primary covers beyond the five dimensions in p_j (twelve total) that each can be combined with an upgrade option. See Appendix 1.10.1 for more details.

observable types $\mathbf{z}_i$ that parameterize $F_{c,\kappa|\mathbf{z}}$ based on interactions of quartiles of soil productivity, prior CRP status, and prior land use status.

We estimate the parameters of $F_{c,\kappa|\mathbf{z}}$ via Maximum Simulated Likelihood (MSL). This estimator maximizes the likelihood of each bidder's observed score-contract combination. Estimation poses a computational challenge because the combined discrete-continuous bidding problem makes choice sets extremely large without allowing for an inversion. We address this challenge in two ways. First, we coarsen the bid space used to construct each bidder's likelihood contribution, maintaining the full dimensionality of the problem when solving equation (1.10).[36] Second, we use a change of variables and importance sampling (following Ackerberg (2009)) to reduce the computational burden associated with searching over a high dimensional bid space for each simulation draw.

Step 3: Estimate Additionality The third step estimates the parameters in equation (1.12), (π, β, α), via the Method of Simulated Moments (MSM). This estimator searches for the parameters (π, β, α) that rationalize moments of the joint distribution of $1 - a_{i0}$, $\mathbf{b}_i^* = (r_i, \mathbf{x}_i)$, and $\mathbf{z}_i$ observed in the data (key moments are illustrated in Figures 1.5). Specifically, we simulate $\left(c_i^k, \kappa_i^k\right)$ from $F_{c,\kappa|\mathbf{z}}$ estimated in Step 2, solve for the optimal $\mathbf{b}_i^{*k}$ (equation (1.10)) for each simulation draw k, and search for the parameters (π, β, α) that match: (i) additionality at the winning score threshold, (ii) additionality among all rejected bidders and by observable characteristics, (iii) the covariance between additionality and chosen scores, and (iv) the additionality among all landowners choosing a given contract.

We measure additionality as $1 - a_{i0}$ among rejected bidders in the remote sensing land use data, as in Section 1.4.2. All moments condition on optimal bids that are below the score threshold. This feature of the estimator accounts for the sample selection in Figure 1.5. The observables $\mathbf{z}_i$ in $\tau(\mathbf{z}_i, c_i, \kappa_i)$ are the 32 observable bidder types that parameterize $F_{c,\kappa|\mathbf{z}}$ and the remaining components of the scoring rule.

1.5.3 Parameter Estimates

Costs Figures 1.6a and 1.6b plot the estimated distributions of c_i and κ_{ij}.[37] A large share of landowners have low values of c_i, below \$50 per acre, per year with a tail of bidders with higher c_i. Top-up costs κ_{ij} are mostly positive; most contracts are more costly than the normalized category of introduced grasses. Table 1.3 summarizes mean costs across contract features and highlights

[36] We coarsen the observed continuous choice of score into deciles of the score distribution and the choice of contract into seven categories corresponding to the seven dimensions of p_j and u_j. See Appendix 1.10.4.

[37] Appendix Table A.7 presents parameter estimates for select cells of $\mathbf{z}_i$ and standard errors.

observable heterogeneity along landowner soil productivity. Relative costs across contracts are generally intuitive. Higher soil productivity bidders have higher costs for primary covers, but lower costs for upgrade covers.

Because costs are estimated using only revealed preferences in bids, Figure 1.6c examines whether estimated costs correlate with land use. Figure 1.6c presents a binned scatterplot of $1 - a_{i0}$ against mean base costs c_i among rejected bidders. Figure 1.6c documents that landowners with higher mean base costs are more additional. This both validates the revealed preference estimates and indicates the presence of adverse selection in the market mediated by observables.

Appendix 1.10.4 examines model fit and compares estimated top-up costs to some limited administrative data on costs submitted to the USDA. Our fit is reasonable and model-implied costs are similar in rank and in magnitude to the administrative data.

Additionality Table 1.4 presents select parameter estimates in $\tau(\mathbf{z}_i, c_i, \kappa_i)$ that describe how additionality varies with (c_i, κ_i). The remaining parameters estimate how additionality varies across observable characteristics.

Each column of Table 1.4 presents a different specification of $\tau(\mathbf{z}_i, c_i, \kappa_i)$. Columns (1) and (2) restrict to a correlation between additionality and c_i and impose that $\alpha = 0$. Column (1) includes observable characteristics from the scoring rule and column (2) adds observable characteristics that parameterize $F_{c,\kappa|\mathbf{z}}$ (interactions of soil productivity and prior land use). Consistent with the positive correlation between bids and additionality in Figures 1.5a and 1.5b, the positive coefficients in columns (1) and (2) of Table 1.4 indicate that landowner additionality is systematically correlated with costs, conditional on observable characteristics. The magnitude of the coefficients presented in Table 1.4 imply that a one standard deviation increase in c_i is associated with an eight percentage point increase in additionality. The coefficient estimates in Table 1.4 reflect the adverse selection presented in Figures 1.5a and 1.5b.

Columns (3) and (4) allow additionality to also depend on κ_i. Column (3) only allows tree-related action costs to impact additionality. The coefficient on tree-related κ_{ij} is positive and large, while the coefficient on c_i is reduced, but still positive. Column (4) allows additionality to vary with κ_i more flexibly, and the residual correlation between costs and additionality loads onto κ_i. The largest coefficient remains on tree-related contracts.

The model-implied estimate of additionality at the RD margin is between 22-23%, within our range of estimates of 21%-31%. This is expected, as our estimation strategy matches land use moments directly.

1.5.4 From Additionality to Contract Value

To calculate social welfare, we require estimates of the social benefits of contracted actions, $B_j(\mathbf{z}_i^s)$. We now index $B_j(\mathbf{z}_i^s)$ by j to account for heterogeneous social benefits across contracts and allow $B_j(\mathbf{z}_i^s)$ to depend on observable characteristics in the scoring rule $\mathbf{z}_i^s$ that capture heterogeneity in the environmental sensitivity of landowners. We take average estimates of the value of CRP actions from literature that quantifies the benefits from habitat restoration and reductions in erosion, water and air pollution, and greenhouse gas emissions from the CRP (Johnson et al., 2016a; Feather et al., 1999; Hansen, 2007a). We take relative valuations across landowners with characteristics $\mathbf{z}_i^s$ and across contracts j from the scoring rule. We therefore consider social welfare under valuations $B_j(\mathbf{z}_i^s)$ revealed preferred by the USDA. See Appendix 1.10.5 for more details.

One remaining detail concerns the fact that additionality is one-dimensional (land-retirement), but the menu of contracts is multi-dimensional. This is due to fundamental data limitations, the substantial simplification that focusing on only this one dimension affords, and the fact that the primary incentivized activity of the CRP is land retirement. Our baseline specification follows Section 1.2 and calculates contract value as $B_j(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i)$. In Appendix 1.10.6, we present and discuss results under an alternative assumption where additionality only affects the base contract. This assumes a valuation of contracts equal to $B_0(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i) + B^j(\mathbf{z}_i^s)$, where $B_0(\mathbf{z}_i^s)$ is the value of the base action and $B^j(\mathbf{z}_i^s)$ is the incremental value of the top-up action.

1.6 Social Welfare and Alternative Market Designs

With estimates of the costs (c_i, κ_i) and expected social benefits $B_j(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i)$ of contracting, we turn to analyzing the social welfare and market design consequences of additionality. Define the expected social surplus of contracting with landowner i for contract j as:

$$B_j(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i) - c_i - \kappa_{ij}. \tag{1.13}$$

Equation (1.13) is based on $\tau(\mathbf{z}_i, c_i, \kappa_i)$ not $1 - a_{i0}$. Because current and counterfactual mechanisms screen only on $(\mathbf{z}_i, c_i, \kappa_i)$, using equation (1.13) for comparisons of social welfare under current and alternative market designs is without further loss.

We first examine allocative efficiency and pricing in the context of our graphical framework with a uniform price and a single contract. In these analyses, we collapse heterogeneity to one dimension of cost for a single contract and the expected additionality at each point in this one-dimensional cost distribution (as in Section 1.2). We then build on the graphical analysis — incorporating

heterogeneity across landowners and contracts — to investigate the performance and design of current and counterfactual auctions and competitive markets for conservation (offset markets).

1.6.1 Graphical Analysis

Base Contract Figure 1.7 presents the empirical analogue of Figure 1.1, graphing the marginal cost (MC), contract value ($B \cdot \tau$), and average contract value curves for the base contract. We simulate the minimum cost to landowners of fulfilling the base contract to construct the MC curve. Then, we calculate the average $B_0(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i)$, where $B_0(\mathbf{z}_i^s)$ denotes the social benefit of the base action, at each quantile of this cost distribution to obtain the contract value curve. The two facts from Section 1.4 are reflected in Figure 1.7. The contract value curve lies below B, representing landowners who are not additional (Section 1.4.1), and is upwards-sloping, representing adverse selection (Section 1.4.2). Figure 1.7 offers three conclusions about the welfare implications of these two facts in the context of our stylized framework.

First, the contract value curve crosses the marginal cost curve from above: the empirical market described by Figure 1.7 is similar to Figure 1.1a, not Figure 1.1b. The socially-optimal uniform price implements the one-dimensional efficient allocation defined by equation (1.5) with social surplus equal to the vertical distance between the contract value and marginal cost curves. This leads to social welfare gains of \$14.66 per acre-year in region CDG. Figure 1.7 demonstrates that the potential market failure introduced by additionality does not lead the market to completely fail.

Second, Figure 1.7 illustrates inefficient contracting when prices are set at B (the average $B_0(\mathbf{z}_i^s)$ across landowners), ignoring counterfactual land use. Pricing at B leads to social welfare losses of \$11.79 per acre-year in triangle GHI, 80% of the gains realized in triangle CDG. These social welfare losses underscore the importance of quantitative analysis of the joint distribution of the costs of contracting and additionality to set socially-optimal incentives to implement efficient allocations.

Third, Figure 1.7 illustrates the trade-limiting effects of adverse selection in competitive (offset) markets with price-taking buyers. We isolate the effect of supply-side adverse selection by assuming buyers in competitive markets possess the same full-information preferences as the USDA, but take expectations over the additionality of all market participants. This demand curve is illustrated with the gray average contract value curve in Figure 1.7. Adverse selection would limit trade in a competitive market to the equilibrium quantity $q^c = 0.58$, a 15% reduction from the socially optimal quantity $q^* = 0.68$. Triangle EFG represents social welfare gains from contracting that are not realized in competitive markets. The magnitude of triangle EFG is 4% of the (one-dimensional) efficient allocation, triangle CDG. Though the adverse selection introduced by additionality exists

in the market, limits trade, and reduces social welfare, it does not unravel the market.

Overall, Figure 1.7 presents a relatively optimistic view of markets for environmental services, which contrasts with arguments that establishing these markets is a futile endeavor (Anderson, 2012). Figure 1.7 also illustrates why. Perhaps surprisingly, the contract value curve is flat for landowners with low contracting costs.

Heterogeneity Figure 1.8 uses the estimated heterogeneity to examine differences in the graphical analysis across observable characteristics and contracts. This heterogeneity will serve as a basis for counterfactual market designs.

Figures 1.8a and 1.8b replicate Figure 1.7 in sub-populations split by estimated soil productivity. Both contract value and marginal cost curves differ in the lowest versus high quintile of the soil productivity distribution, implying different socially-optimal prices.

Figure 1.8c examines heterogeneity across contracts. We focus on tree contracts due to the evidence of substantial adverse selection in Table 1.4 and the prevalence of tree-related PES programs and offset contracts. Figure 1.8c calculates the marginal cost curve as the minimum cost required to fulfill any tree-related contract and plots the average $B_j\left(\mathbf{z}_i^s\right) \cdot \tau\left(\mathbf{z}_i, c_i, \kappa_i\right)$ at each quantile of this distribution. The exercise requires substantial out-of-sample extrapolation, but it illustrates how alternative landowner type distributions across important classes of contracts in our setting can generate different conclusions.

In Figure 1.8c, the contract value curve crosses the marginal cost curve from below, leading to social welfare losses at low q. The socially-optimal uniform price cannot implement the one-dimensional efficient allocation (DE) defined in equation (1.5). This is because, as in Figure 1.1b, the ordering of landowners by social surplus (the vertical distance between the contract value and marginal cost curves) differs from the ordering of landowners by contracting costs (the construction of the x-axis, q). Because they are less additional, the lowest cost landowners are not the highest social value.

Figure 1.8c also illustrates that supply-side adverse selection would cause a competitive (offset) market for tree contracts to unravel.

1.6.2 Alternative Auctions

The standard auction design problem focuses on cost-minimizing procurement auctions, but the objective of payments for ecosystem services mechanisms is to *impact outcomes* (conservation) at lowest cost. Many other incentive-based public policies face similar objectives. However, standard

mechanisms focused on cost-minimization, which consider reports of (c_i, κ_i) but not the effect of contracting on conservation $\tau(\mathbf{z}_i, c_i, \kappa_i)$, may not advance this goal.

We simulate bidding and additionality under status quo and counterfactual auctions to investigate this possibility and the performance of alternative designs. Figure 1.9 and Table 1.5 present results. Figure 1.9 plots social welfare under each allocation:

$$\sum_i \sum_j \left(B_j(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i) - c_i - \kappa_{ij} \right) \cdot x_{ij}. \tag{1.14}$$

Table 1.5 tabulates the bars in Figure 1.9 and reports additional details: USDA spending, landowner surplus, the value of environmental benefits $\sum_i \sum_j B_j(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i) \cdot x_{ij}$, average additionality, and the share of bidders allocated a contract. Each bar of Figure 1.9 corresponds to the same numbered column in Table 1.5.

The Status Quo Auction versus an Efficient Benchmark Because social welfare depends on additionality but the design of the status quo auction does not, the social value of the CRP is ambiguous. We document social welfare gains of $126 million per auction in the status quo (bar (1) of Figure 1.9). This is calculated by simulating optimal bidding in the mechanism with the estimated distribution of (c_i, κ_i) and beliefs $G(S)$.

However, social welfare under the status quo auction is only 15% of an efficient benchmark. This efficient benchmark is defined as the allocation that uses all observables $\mathbf{z}_i$ and the full vector of costs (c_i, κ_i) to maximize social welfare (equation (1.14)) subject to two constraints. First, each landowner must obtain at most one contract $\sum_j x_{ij} \leq 1$. Second, no more landowners are allocated contracts than in the status quo. Because many landowners are not additional, the efficient allocation involves contracting with fewer landowners than the status quo and the quantity constraint does not bind (column (2) of Table 1.5).

This efficient allocation may not be implementable in an incentive compatible auction if social surplus and allocations are not monotone in bidder costs (Myerson, 1981a). This complication is relevant because of adverse selection; once the mechanism's impact on conservation (additionality) is considered, the lowest cost landowners may not be the highest social value. This issue is illustrated in principle in Figure 1.1b and based on our estimates in Figure 1.8c.

Alternative Auctions: Vickrey Auctions with Scoring The status quo auction underperforms the efficient allocation in part because it does not consider additionality in its design. Implementing the efficient allocation may be impossible, but changes in design to incorporate additionality may close the gap.

In practice, a common approach to additionality is to define eligibility requirements that determine who and what is allowed to trade.[38] Emphasis is placed on identifying additional participants who are then allowed to participate in the market.

We consider a more flexible approach that treats landowners asymmetrically by their expected additionality. Contracting with a low expected additionality landowner could be justified at sufficiently low cost. Conversely, landowners who are likely to be additional may still counterfactually conserve with some probability. This approach accommodates a minimum standard — incentives could be zero for some participants or some conservation actions — but it also allows incentives to differ across landowner observables and contracts in the market. We implement this approach in counterfactual scoring auctions that construct scoring rules based on predictions of $B_j(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i)$. These auctions build directly on the status quo, which uses a scoring rule based on $B_j(\mathbf{z}_i^s)$.

Define a *contract value scoring rule* $s_j(\mathbf{z}_i)$ to parameterize the (score-implied) expected social benefit of contracting with a bidder with characteristics $\mathbf{z}_i$ for contract j. We focus on linear rules based on (a simplified version of) the functional form of the status quo scoring rule,[39]

$$s_j(\mathbf{z}_i) = \omega_{\mathbf{z}} \cdot \mathbf{z}_i + \omega_j, \tag{1.15}$$

where $\omega_{\mathbf{z}}$ parameterizes scores across observables $\mathbf{z}_i$ (asymmetry terms) and ω_j parameterizes scores across contracts j.

We implement allocations defined by the status quo and alternative scoring rules with a Vickrey-Clarke-Groves (VCG) mechanism.[40] We term these auctions "Vickrey auctions with scoring," which maximize a definition of social welfare implied by the scoring rule $s_j(\mathbf{z}_i)$:

$$\sum_i \sum_j \left(s_j(\mathbf{z}_i) - c_i - \kappa_{ij}\right) \cdot x_{ij}. \tag{1.16}$$

Bidders are treated asymmetrically by $s_j(\mathbf{z}_i)$ not $\tau(\mathbf{z}_i, c_i, \kappa_i)$: compare equations (1.14) and (1.16). In Vickrey auctions with scoring, bidders truthfully report their vector of (c_i, κ_i), then are ranked by $\max_j s_j(\mathbf{z}_i) - c_i - \kappa_{ij}$.[41] The highest scoring bidders subject to the auction's quantity threshold

[38] For examples, see the Clean Development Mechanism Methodology Booklet, the REDD+ eligibility requirements, and the Verified Carbon Standard.

[39] We simplify the status quo scoring rule by eliminating heterogeneity across contracts based on WPZ status and non-linearities in the scoring rule. See Appendix 1.10.1 for more details.

[40] A VCG mechanism is a generalization of a second price auction. Many of the well-known undesirable properties of VCG mechanisms do not apply in our setting because bidders have substitutes preferences (see Ausubel and Milgrom (2005)).

[41] The term $\max_j s_j(\mathbf{z}_i) - c_i - \kappa_{ij}$ is a bidder's pseudo-type in the terminology of Asker and Cantillon (2008b). It is the maximum level of scoring-rule-implied social surplus that bidder i can generate.

are allocated the contract $\arg\max_j s_j(\mathbf{z}_i) - c_i - \kappa_{ij}$.[42] Unlike in Asker and Cantillon (2008b) and Che (1993), the scoring rule may not capture all heterogeneity in $\tau(\mathbf{z}_i, c_i, \kappa_i)$. $s_j(\mathbf{z}_i)$ depends only on observable characteristics. Moreover, some characteristics may not be used to avoid introducing perverse incentives to game the scoring rule (e.g. prior land use).

Vickrey auctions with scoring have three advantages. First, they focus attention on the design of the scoring rule because they implement an allocation that maximizes scoring-rule-implied social welfare (equation (1.16)). Second, allocations are not computationally demanding to calculate. Finally, they are simple: the market designer needs only to compute $s_j(\mathbf{z}_i)$.[43]

Social Welfare Under Status Quo and Alternative Scoring Rules Bars (3)-(6) of Figure 1.9 adjust the scoring rule $s_j(\mathbf{z}_i)$, holding the number of awarded contracts fixed at the status quo quantity. Each bar evaluates the allocation implemented by the auction with equation (1.14). Additional details are reported in the corresponding columns of Table 1.5.

Bar (3) maintains the status quo scoring rule, $s_j(\mathbf{z}_i) = B_j(\mathbf{z}_i^s)$, but changes the auction mechanism to VCG. This counterfactual isolates the impact of a scoring rule that is naive to additionality (bar (2) versus bar (3)) and provides a basis for further comparisons that change only the rule $s_j(\mathbf{z}_i)$ but hold constant the VCG auction mechanism. If all landowners were additional, the scoring rule defined by $s_j(\mathbf{z}_i) = B_j(\mathbf{z}_i^s)$ would implement the efficient allocation and dominate the status quo. Instead, it slightly *underperforms* it. Correcting inefficient design features of the status quo auction, e.g. bid caps, does not increase social welfare when the scoring rule does not incorporate additionality. The comparison of bar (3) to bars (1) and (2) illustrates that the poor performance of the status quo, relative to the efficient allocation, is because $\tau(\mathbf{z}_i, c_i, \kappa_i)$ is not incorporated into the mechanism.

Bars (4)-(6) in Figure 1.9 adjust the scoring rule $s_j(\mathbf{z}_i)$ defined in equation (1.15) based on predictions of additionality. First, we adjust the scoring of the menu of contracts, ω_j. In bar (4), ω_j is calculated to maximize equation (1.14), holding $\omega_{\mathbf{z}}$ constant at the status quo rule.[44] This change to the scoring rule doubles the social welfare gains of the auction relative to the status quo ($128

[42] The VCG incentive payment that implements this allocation is:

$$\sum_j s_j(\mathbf{z}_i) \cdot x^*_{ij} + \sum_{i' \neq i} \sum_{j'} \left(s_{j'}(\mathbf{z}_{i'}) - c_{i'} - \kappa_{i'j'} \right) \cdot x^*_{i'j'} - \sum_{i' \neq i} \sum_{j'} \left(s_{j'}(\mathbf{z}_{i'}) - c_{i'} - \kappa_{i'j'} \right) \cdot x^{-i}_{i'j'},$$

where $\{x^*_{ij}\}$ denotes the allocation that maximizes $\sum_{i'} \sum_{j'} \left(s_{j'}(\mathbf{z}_{i'}) - c_{i'} - \kappa_{i'j'} \right)$ given all reports of $(c_{i'}, \kappa_{i'j'})$ in the population and $\{x^{-i}_{i'j'}\}$ denotes the allocation that solves $\max_x \sum_{i' \neq i} \sum_{j'} \left(s_{j'}(\mathbf{z}_{i'}) - c_{i'} - \kappa_{i'j'} \right) \cdot x^{-i}_{i'j'}$.

[43] This is in contrast to alternatives that could incorporate randomization, as in Lopomo et al. (2023).

[44] We solve for the ω_j that maximize equation (1.14) given simulations of landowner (c_i, κ_i), estimates of $\tau(\mathbf{z}_i, c_i, \kappa_i)$, calibrations of $B_j(\mathbf{z}_i^s)$, and the allocation rule, holding $\omega_{\mathbf{z}}$ fixed at the status quo scoring rule.

million per auction). Relative to the status quo rule, re-weighting ω_j accounts for both heterogeneity in $\tau(\mathbf{z}_i, c_i, \kappa_i)$ as a function of κ_i — for example, a landowner's choice of a tree-related contract reveals that her conservation is unlikely to be additional — and the fact that the full social benefits across actions is not realized when conservation would have counterfactually occurred.

Next, we adjust the bidder asymmetry terms across observables, $\omega_{\mathbf{z}}$, in equation (1.15). Displayed in bars (5) and (6) in Figure 1.9, this change leads to a further \$46 million of social welfare gains per auction (37% of the status quo). Bar (5) first re-weights the existing characteristics in the scoring rule, $\mathbf{z}_i^s$. In the status quo, asymmetry is based only on environmental sensitivity; re-weighting $\omega_{\mathbf{z}}$ across $\mathbf{z}_i^s$ based also on additionality contributes two-thirds of the social welfare gains from adjusting the bidder asymmetry terms. The final adjustment (bar (6)) adds an additional characteristic to the rule: an "additionality factor" $\hat{\tau}(\mathbf{z}_i)$. This builds on the scoring rule's design, which adds together many "factors" to construct a composite score of bidder characteristics.[45] We calculate $\hat{\tau}(\mathbf{z}_i)$ by projecting $\tau(\mathbf{z}_i, c_i, \kappa_i)$ on immutable characteristics of landowners already collected by the USDA, but not all incorporated in the status quo scoring rule: deciles of soil productivity and wind and water erosion. Then, we calculate the social-welfare maximizing score using both $\mathbf{z}_i^s$ and $\hat{\tau}(\mathbf{z}_i)$ as asymmetry terms in equation (1.15). The simple change of adding this single "additionality factor" to the scoring rule increases social welfare by a further 12% of the status quo.

Figure 1.9 illustrates that simple changes to the scoring rule can lead to significant social welfare gains. In contrast to standard cost-minimizing procurement auctions, these auctions are designed to *impact conservation* at lowest cost. A scoring rule that incorporates landowners' expected additionality balances the objectives of allocating contracts to both low cost and high social benefit, additional landowners.

Social Welfare Under Alternative Market Sizes Beyond the allocation rule, additionality also impacts the socially-optimal size of the market. This also contributes to the gap between bars (1) and (2) but was ignored in the prior counterfactuals, which held the number of contract awards constant at the status quo.

Because many landowners are not additional, the status quo quantity procured is higher than is socially-optimal. Bar (7) in Figure 1.9 keeps the scoring rule $s_j(\mathbf{z}_i)$ of bar (6) but awards contracts only to landowners with $\max_j s_j(\mathbf{z}_i) - c_i - \kappa_{ij} \geq 0$. This reduction in market size increases social welfare by a further \$110 million dollars per auction.

Together, simple adjustments to both the size of the market and the scoring rule based on predicted additionality closes the gap between the status quo (bar (1)) and the efficient allocation (bar (2))

[45] For more details on these factors, see an example EBI Factsheet here.

by 41% (an increase of $284 million per auction).[46] Each component of the mechanism — adjusted incentives across contracts, across characteristics, and the overall size of the market — is a quantitatively important contribution to this improvement.

Mechanisms: Additionality in the Scoring Rule The alternative auctions in Figure 1.9 outperform the status quo by adjusting the scoring rule to reflect the social benefit of contracting, which depends on additionality. This occurs via two channels.

First, the status quo scoring rule $B_j(\mathbf{z}_i^s)$ over-weights asymmetry across characteristics $\mathbf{z}_i^s$ and contracts j. The heterogeneous social benefits of conservation $B_j(\mathbf{z}_i^s)$ are not fully realized when conservation would have counterfactually occurred. An auction that treats bidders and contracts asymmetrically by $B_j(\mathbf{z}_i^s)$ may not implement an efficient allocation. This relates to the social welfare losses from pricing at B illustrated in Figure 1.7.

Second, as highlighted in Figure 1.8, bidders may be systematically heterogeneous in their additionality, which can be exploited in the scoring rule. Adjusting the scoring rule based on heterogeneity in additionality — using observable characteristics and choices of contracts — more closely aligns the allocation implemented by the auction with the socially-optimal allocation that considers heterogeneity in both additionality and costs.

Figure 1.10 explores these two mechanisms. Figure 1.10 holds quantity constant and plots social welfare under the status quo auction (bar (1)), a Vickrey auction with a scoring rule $s_j(\mathbf{z}_i) = \theta \cdot B_j(\mathbf{z}_i^s)$ for the (single) multiplier θ chosen to maximize equation (1.14) (bar (2)), and a Vickrey auction with a scoring rule that adjusts ω_j and $\omega_{\mathbf{z}}$ to maximize equation (1.14) (bar (3), which replicates the auction in bar (6) in Figure 1.9). Bar (2) examines the social welfare gains achieved with only a uniform instrument to adjust the scoring rule for additionality. Bar (3) further adjusts ω_j and $\omega_{\mathbf{z}}$ to reflect heterogeneously additional landowners.

Figure 1.10 demonstrates that adjusting the scoring rule based on heterogeneity in additionality yields substantial social welfare gains of $23 million per auction (bar (3)), but that a uniform adjustment for additionality (bar (2)) achieves a large share of the gains relative to the status quo. An auction that incorporates asymmetry in the social benefits of actions must also reflect the additionality of landowners in its design (see bar (2)). Moreover, further adjusting asymmetry based on heterogeneity in additionality yields further social welfare gains, equivalent to 18% of the status quo (see bar (3)).

[46] Further differences between bar (7) and bar (2) reflect (i) $\mathbf{z}_i$ that are not incorporated into the scoring rule to avoid perverse incentives to game the rule, (ii) private landowner costs in $\tau(\mathbf{z}_i, c_i, \kappa_i)$, and (iii) the functional form of equation (1.15) relative to $B_j(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i)$.

USDA Spending Beyond social welfare, the budgetary implications of alternative auctions may be relevant in practice. Among implementable auctions we consider, the auction with the greatest social welfare gains also reduces USDA spending relative to the status quo (column (7) of Table 1.5). This occurs because the status quo auction contracts with too many landowners. Reducing the size of the market, and therefore total USDA spending, increases social welfare. Appendix 1.10.6 also evaluates social welfare with a cost of public funds, motivated by the need to finance expenditures with distortionary taxation. Social welfare with a cost of public funds is negative under the status quo, but becomes positive and substantial under alternative designs.

However, Table 1.5 also demonstrates that in all auctions, government spending exceeds the value of environmental services procured, $\sum_i \sum_j B_j(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i) \cdot x_{ij}$. This is due to the presence of adverse selection in the market: the marginal landowner has a higher value of $\tau(\mathbf{z}_i, c_i, \kappa_i)$ than the inframarginal landowner.

1.6.3 Offset Market Design

We conclude with the implications of supply-side adverse selection for the performance and design of competitive (offset) markets for environmental services. We continue to isolate the effect of supply-side adverse selection. We assume that buyers have the same full-information preferences as the USDA and form expectations over the value of any contract given the equilibrium price(s). We ask two questions motivated by the analysis in Section 1.6.1. First, should offset markets be differentiated? And second, which markets risk unravelling?

The effect of differentiation on social welfare in competitive markets is ambiguous (Einav and Finkelstein, 2011). We analyze this market design choice empirically in Figure 1.11a, restricting analysis to the base contract. Figure 1.11a plots the percent reduction in quantities traded and social welfare in a competitive market, relative to with socially-optimal prices, under uniform and differentiated markets. In the uniform market, there is only a single socially-optimal price and market-clearing condition. In the differentiated market, we project $B_0(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i)$ onto immutable observable characteristics ($\mathbf{z}_i^s$, soil productivity, and erosion) and then segment the market into deciles of predicted contract value. This "certification scheme" is similar in structure to existing rating schemes in environmental markets.[47] Figure 1.11a also presents social welfare per acre-year under each of these offset market designs.

Differentiation reduces social welfare losses from adverse selection in competitive markets from 5% to less than 1% and increases social welfare by 15% overall via more efficient trades in the

[47] See, for example, Carlyx Global, BeZero Ratings, and Sylvera.

market. The gains from differentiation are high even in the ex-ante ambiguous competitive market setting, supporting on-going efforts to collect detailed information to predict additionality in environmental markets.[48]

Next, we investigate which contracts can be successfully traded in competitive markets, motivated by Figures 1.7 and 1.8c. We consider uniform markets for grass-, tree-, and habitat-related contracts. Figure 1.11b plots the reduction in social welfare relative to the socially optimal uniform price in each of these three hypothetical markets. Tree-related contracts unravel, but social welfare losses for the remaining contracts are limited to at most 3%.

Figure 1.11 presents a relatively optimistic view of offset markets and actionable insights for market design. We offer three ideas about features of our setting that contribute to this conclusion. First, the eligibility requirements for the CRP are stringent enough that there is some probability that landowners are additional even at the bottom of the contracting cost distribution. Second, hassle costs and long-term contracts mute the extent of adverse selection, which limits unravelling. Finally, agricultural decisions are simple to predict, offering covariates to differentiate landowners and increase social welfare.

1.7 Conclusion

Additionality is a central challenge to environmental market design. It undermines the appeal of market-based mechanisms if incentives attract the least additional landowners.

This paper combines data and theory to document this potential market failure, quantify its implications for social welfare, and evaluate alternative market designs in the largest auction mechanism for ecosystem services in the world. Linking satellite data to auction bids, we use a regression discontinuity design to demonstrate that only one quarter of landowners are additional. Moreover, heterogeneity in counterfactual land use introduces adverse selection in the market. To quantify the implications of these facts and test possible remedies, we develop and estimate a joint model of multi-dimensional bidding and land use that incorporates adverse selection on additionality.

With socially-optimal incentives, the market can deliver social welfare gains, but the lowest cost providers of environmental services are not always the highest social value. Re-designing the auction's scoring rule to incorporate predicted additionality substantially outperforms the status quo, and a simple differentiation scheme also increases social welfare in competitive offset markets.

A common market design solution to the issue of additionality is to define eligibility requirements that restrict who and what can trade; in this paper, we propose a more flexible approach. Because

[48] See, for example, Google, Microsoft, and the platform NCX.

many markets will inevitably attract landowners who are with some probability not additional, allocation mechanisms should consider this dimension of heterogeneity in their design. We show how auctions can be used to cost-effectively *impact conservation,* selecting participants based on both expected additionality and costs, despite the existence of many landowners in the market that are not additional. Segmenting offset markets yields social welfare gains via similar mechanisms.

Our analysis focused only on the supply-side market failure of additionality. Investigating other features of offset markets, including demand, the incentives of platforms and certifiers that facilitate trade, and both of their interactions with supply-side additionality and adverse selection are interesting and impactful avenues for future research.

More broadly, our results highlight that successful market design depends not only on market participants' private costs, but also on whether their behavior in the market advances a socially desirable outcome. Developing empirical approaches to apply this idea to the design of other markets and policy objectives is a rich and exciting area for research.

1.8 References

Ackerberg, Daniel A., "A new use of importance sampling to reduce computational burden in simulation estimation," *QME*, December 2009, *7* (4), 343–376.

Agarwal, Nikhil, Charles Hodgson, and Paulo Somaini, "Choices and Outcomes in Assignment Mechanisms: The Allocation of Deceased Donor Kidneys," November 2020.

_ , Pearl Li, and Paulo Somaini, "Identification using Revealed Preferences in Linearly Separable Models," October 2023.

Akerlof, George A., "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 1970, *84* (3), 488–500. Publisher: Oxford University Press.

Alix-Garcia, Jennifer and Daniel Millimet, "Remotely Incorrect? Accounting for Nonclassical Measurement Error in Satellite Data on Deforestation," *Journal of the Association of Environmental and Resource Economists*, December 2022.

Alix-Garcia, Jennifer M., Katharine R. E. Sims, and Patricia Yañez-Pagans, "Only One Tree from Each Seed? Environmental Effectiveness and Poverty Alleviation in Mexico's Payments for Ecosystem Services Program," *American Economic Journal: Economic Policy*, November 2015, *7* (4), 1–40.

_ and Michael Greenstone, "Is There an Energy Efficiency Gap?," *Journal of Economic Perspectives*, February 2012, *26* (1), 3–28.

_ and _ , "Measuring the Welfare Effects of Residential Energy Efficiency Programs," May 2017.

Allen, Arthur W. and Mark W. Vandever, "Conservation Reserve Program (CRP) contributions to wildlife habitat, management issues, challenges and policy choices–an annotated bibliography," USGS Numbered Series 2012-5066, U.S. Geological Survey, Reston, VA 2012.

Allen, Jason, Robert Clark, Brent Hickman, and Eric Richert, "Resolving Failed Banks: Uncertainty, Multiple Bidding, and Auction Design," *The Review of Economic Studies*, 2023.

Anderson, Kevin, "The inconvenient truth of carbon offsets," *Nature*, April 2012, *484* (7392), 7–7.

Anderson, Terry L. and Gary D. Libecap, *Environmental Markets: A Property Rights Approach*, New York: Cambridge University Press, May 2014.

Arnold, David, Will Dobbie, and Peter Hull, "Measuring Racial Discrimination in Bail Decisions," *American Economic Review*, September 2022, *112* (9), 2992–3038.

Aronoff, Daniel and Will Rafey, "Conservation priorities and environmental offsets: Markets for Florida Wetlands," July 2023.

_ and _ , "Properties of scoring auctions," *The RAND Journal of Economics*, 2008, *39* (1), 69–85.

_ **and** _ , "Procurement when price and quality matter," *The RAND Journal of Economics*, 2010, *41* (1), 1–34.

Ausubel, Lawrence M. and Paul Milgrom, "The Lovely but Lonely Vickrey Auction," in Peter Cramton, Yoav Shoham, and Richard Steinberg, eds., *Combinatorial Auctions*, The MIT Press, December 2005, p. 0.

_ **and** _ , "Scaling Auctions as Insurance: A Case Study in Infrastructure Procurement," *Econometrica*, 2023, *91* (4), 1205–1259.

Borenstein, Severin, "The Private and Public Economics of Renewable Electricity Generation," *Journal of Economic Perspectives*, February 2012, *26* (1), 67–92.

Borrelli, Pasquale, David A. Robinson, Larissa R. Fleischer, Emanuele Lugato, Cristiano Ballabio, Christine Alewell, Katrin Meusburger, Sirio Modugno, Brigitta Schütt, Vito Ferro, Vincenzo Bagarello, Kristof Van Oost, Luca Montanarella, and Panos Panagos, "An assessment of the global impact of 21st century land use change on soil erosion," *Nature Communications*, December 2017, *8* (1), 2013.

Bundorf, M. Kate, Jonathan Levin, and Neale Mahoney, "Pricing and Welfare in Health Plan Choice," *American Economic Review*, December 2012, *102* (7), 3214–3248.

Calel, Raphael, Jonathan Colmer, Antoine Dechezlepretre, and Mattieu Glachant, "Do carbon offsets offset carbon?," November 2021.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Carril, Rodrigo, Andres Gonzalez-Lira, and Michael S. Walker, "Competition under Incomplete Contracts and the Design of Procurement Policies," March 2022.

Chan, David C., Matthew Gentzkow, and Chuan Yu, "Selection with Variation in Diagnostic Skill: Evidence from Radiologists," *The Quarterly Journal of Economics*, May 2022, *137* (2), 729–783.

Che, Yeon-Koo, "Design Competition Through Multidimensional Auctions," *The RAND Journal of Economics*, 1993, *24* (4), 668–680.

Chiappori, Pierre-Andre and Bernard Salanie, "Testing for Asymmetric Information in Insurance Markets," *Journal of Political Economy*, 2000, *108* (1), 56–78.

Claassen, Roger, Eric N. Duquette, and David J. Smith, "Additionality in U.S. Agricultural Conservation Programs," *Land Economics*, February 2018, *94* (1), 19–35.

Decarolis, Francesco, "Awarding Price, Contract Performance, and Bids Screening: Evidence from Procurement Auctions," *American Economic Journal: Applied Economics*, January 2014, *6* (1), 108–132.

Dirzo, Rodolfo, Hillary S. Young, Mauro Galetti, Gerardo Ceballos, Nick J. B. Isaac, and Ben Collen, "Defaunation in the Anthropocene," *Science*, July 2014, *345* (6195), 401–406.

_ , _ , **and Mark R. Cullen**, "Estimating Welfare in Insurance Markets Using Variation in Prices*," *The Quarterly Journal of Economics*, August 2010, *125* (3), 877–921.

_ , _ , **Stephen P. Ryan, Paul Schrimpf, and Mark R. Cullen**, "Selection on Moral Hazard in Health Insurance," *American Economic Review*, 2013, *103* (1), 178–219.

_ , _ , **Yunan Ji, and Neale Mahoney**, "Voluntary Regulation: Evidence from Medicare Payment Reform*," *The Quarterly Journal of Economics*, February 2022, *137* (1), 565–618.

_ **and** _ , "Selection in Insurance Markets: Theory and Empirics in Pictures," *Journal of Economic Perspectives*, March 2011, *25* (1), 115–138.

Engel, Stefanie, Stefano Pagiola, and Sven Wunder, "Designing payments for environmental services in theory and practice: An overview of the issues," *Ecological Economics*, May 2008, *65* (4), 663–674.

FAPRI-MU, *Estimating Water Quality, Air Quality, and Soil Carbon Benefits of the Conservation Reserve Program* 2007.

Feather, Peter, Daniel Hellerstein, and LeRoy T. Hansen, "Economic Valuation of Environmental Benefits and the Targeting of Conservation Programs: The Case of the CRP," *Agricultural Economic Reports*, 1999.

Filewod, Ben, "Why REDD will Fail," *Journal of Forestry*, July 2017, *115* (4), 323.

Friedlingstein, Pierre, Michael O'Sullivan, Matthew W. Jones, Robbie M. Andrew, Luke Gregor, Judith Hauck, Corinne Le Quéré, Ingrid T. Luijkx, Are Olsen, Glen P. Peters, Wouter Peters, Julia Pongratz, Clemens Schwingshackl, Stephen Sitch, Josep G. Canadell, Philippe Ciais, Robert B. Jackson, Simone R. Alin, Ramdane Alkama, Almut Arneth, Vivek K. Arora, Nicholas R. Bates, Meike Becker, Nicolas Bellouin, Henry C. Bittig, Laurent Bopp, Frédéric Chevallier, Louise P. Chini, Margot Cronin, Wiley Evans, Stefanie Falk, Richard A. Feely, Thomas Gasser, Marion Gehlen, Thanos Gkritzalis, Lucas Gloege, Giacomo Grassi, Nicolas Gruber, Özgür Gürses, Ian Harris, Matthew Hefner, Richard A. Houghton, George C. Hurtt, Yosuke Iida, Tatiana Ilyina, Atul K. Jain, Annika Jersild, Koji Kadono, Etsushi Kato, Daniel Kennedy, Kees Klein Goldewijk, Jürgen Knauer, Jan Ivar Korsbakken, Peter Landschützer, Nathalie Lefèvre, Keith Lindsay, Junjie Liu, Zhu Liu, Gregg Marland, Nicolas Mayot, Matthew J. McGrath, Nicolas Metzl, Natalie M. Monacci, David R. Munro, Shin-Ichiro Nakaoka, Yosuke Niwa, Kevin O'Brien, Tsuneo Ono, Paul I. Palmer, Naiqing Pan, Denis Pierrot, Katie Pocock, Benjamin Poulter, Laure Resplandy, Eddy Robertson, Christian Rödenbeck, Carmen Rodriguez, Thais M. Rosan, Jörg Schwinger, Roland Séférian, Jamie D. Shutler, Ingunn Skjelvan, Tobias Steinhoff, Qing Sun, Adrienne J. Sutton, Colm Sweeney, Shintaro Takao, Toste Tanhua, Pieter P. Tans, Xiangjun Tian, Hanqin Tian, Bronte Tilbrook, Hiroyuki Tsujino, Francesco Tubiello, Guido R. van der Werf, Anthony P. Walker, Rik Wanninkhof, Chris Whitehead, Anna Willstrand Wranne, Rebecca Wright, Wenping Yuan, Chao Yue, Xu Yue, Sönke Zaehle, Jiye Zeng, and Bo Zheng, "Global Carbon Budget 2022," *Earth System Science Data*, November 2022, *14* (11), 4811–4900.

Guerre, Emmanuel, Isabelle Perrigne, and Quang Vuong, "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 2000, *68* (3), 525–574. Publisher: [Wiley, Econometric Society].

Hagerty, Nick, "Adaptation to Surface Water Scarcity in Irrigated Agriculture," November 2022.

Hanazono, Makoto, Yohsuke Hirose, Jun Nakabayashi, and Masanori Tsuruoka, "Theory, Identification, and Estimation for Scoring Auctions," August 2020.

Hansen, LeRoy, "Conservation Reserve Program: Environmental Benefits Update," *Agricultural and Resource Economics Review*, October 2007, *36* (2), 267–280.

Harstad, Bård, "The market for conservation and other hostages," *Journal of Economic Theory*, November 2016, *166*, 124–151.

_ **and Torben K. Mideksa**, "Conservation Contracts and Political Regimes," *The Review of Economic Studies*, October 2017, *84* (4), 1708–1734.

Haupt, Andreas, Nicole Immorlica, and Brendan Lucier, "Voluntary Carbon Market Design," August 2023.

Hellerstein, Daniel M., "The US Conservation Reserve Program: The evolution of an enrollment mechanism," *Land Use Policy*, April 2017, *63*, 601–610.

Hendricks, Kenneth and Robert Porter, "An Empirical Study of an Auction with Asymmetric Information," *American Economic Review*, 1988, *78* (5), 865–83.

Hortaçsu, Ali, "Mechanism Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis Of the Turkish Treasury Auction Market," November 2000.

_ **and David McAdams**, "Mechanism Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis of the Turkish Treasury Auction Market," *Journal of Political Economy*, October 2010, *118* (5), 833–865.

_ **,** _ **,** **and** _ , "Selection on Welfare Gains: Experimental Evidence from Electricity Plan Choice," *American Economic Review*, November 2023, *113* (11), 2937–2973.

Jack, B. Kelsey, "Private Information and the Allocation of Land Use Subsidies in Malawi," *American Economic Journal: Applied Economics*, July 2013, *5* (3), 113–135.

_ **and Seema Jayachandran**, "Self-selection into payments for ecosystem services programs," *Proceedings of the National Academy of Sciences*, March 2019, *116* (12), 5326–5333.

Jayachandran, Seema, Joost de Laat, Eric F. Lambin, Charlotte Y. Stanton, Robin Audy, and Nancy E. Thomas, "Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation," *Science*, July 2017, *357* (6348), 267–273.

Jofre-Bonet, Mireia and Martin Pesendorfer, "Estimation of a Dynamic Auction Game," *Econometrica*, 2003, *71* (5), 1443–1489.

Johnson, Kris A., Brent J. Dalzell, Marie Donahue, Jesse Gourevitch, Dennis L. Johnson, Greg S. Karlovits, Bonnie Keeler, and Jason T. Smith, "Conservation Reserve Program (CRP) lands provide ecosystem service benefits that exceed land rental payment costs," *Ecosystem Services*, April 2016, *18*, 175–185.

Jones, Julia P. G. and Simon L. Lewis, "Forest carbon offsets are failing," *Science*, August 2023, *381* (6660), 830–831.

Kinzig, A. P., C. Perrings, F. S. Chapin, S. Polasky, V. K. Smith, D. Tilman, and B. L. Turner, "Paying for Ecosystem Services—Promise and Peril," *Science*, November 2011, *334* (6056), 603–604.

Kong, Yunmi, Isabelle Perrigne, and Quang Vuong, "Multidimensional Auctions of Contracts: An Empirical Analysis," *American Economic Review*, May 2022, *112* (5), 1703–1736.

Lakkaraju, Himabindu, Jon Kleinberg, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan, "The Selective Labels Problem: Evaluating Algorithmic Predictions in the Presence of Unobservables," *KDD: proceedings. International Conference on Knowledge Discovery & Data Mining*, August 2017, *2017*, 275–284.

Lark, Tyler J., Ian H. Schelly, and Holly K. Gibbs, "Accuracy, Bias, and Improvements in Mapping Crops and Cropland across the United States Using the USDA Cropland Data Layer," *Remote Sensing*, January 2021, *13* (5), 968.

_ , Richard M. Mueller, David M. Johnson, and Holly K. Gibbs, "Measuring land-use and land-cover change using the U.S. department of agriculture's cropland data layer: Cautions and recommendations," *International Journal of Applied Earth Observation and Geoinformation*, October 2017, *62*, 224–235.

Lewis, Gregory and Patrick Bajari, "Procurement Contracting with Time Incentives: Theory and Evidence," *The Quarterly Journal of Economics*, 2011, *126* (3), 1173–1211.

Li, Wanyi Dai, Itai Ashlagi, and Irene Lo, "Simple and Approximately Optimal Contracts for Payment for Ecosystem Services," *Management Science*, February 2022.

Lopomo, Giuseppe, Nicola Persico, and Alessandro T. Villa, "Optimal Procurement With Quality Concerns," *American Economic Review*, 2023.

Manelli, Alejandro M. and Daniel R. Vincent, "Optimal Procurement Mechanisms," *Econometrica*, 1995, *63* (3), 591–620.

Maron, Martine, Christopher D. Ives, Heini Kujala, Joseph W. Bull, Fleur J. F. Maseyk, Sarah Bekessy, Ascelin Gordon, James E.M. Watson, Pia E. Lentini, Philip Gibbons, Hugh P. Possingham, Richard J. Hobbs, David A. Keith, Brendan A. Wintle, and Megan C. Evans, "Taming a Wicked Problem: Resolving Controversies in Biodiversity Offsetting," *BioScience*, June 2016, *66* (6), 489–498.

Marone, Victoria R. and Adrienne Sabety, "When Should There Be Vertical Choice in Health Insurance Markets?," *American Economic Review*, January 2022, *112* (1), 304–342.

Mason, Charles and Andrew J. Plantinga, "The additionality problem with offsets: Optimal contracts for carbon sequestration in forests," *Journal of Environmental Economics and Management*, 2013, *66* (1), 1–14.

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, February 2008, *142* (2), 698–714.

McKinsey Sustainability, "A blueprint for scaling voluntary carbon markets to meet the climate challenge," Technical Report January 2021.

_ , "The net-zero transition: what it would cost, what it could bring," Technical Report 2022.

Milgrom, Paul, "Adverse Selection Without Hidden Information," June 1987.

Montero, Juan-Pablo, "Voluntary Compliance with Market-Based Environmental Policy: Evidence from the U.S. Acid Rain Program," *Journal of Political Economy*, October 1999, *107* (5), 998–1033.

Myerson, Roger B., "Optimal Auction Design," *Mathematics of Operations Research*, 1981, *6* (1), 58–73.

Ribaudo, Marc O., Dana L. Hoag, Mark E. Smith, and Ralph Heimlich, "Environmental indices and the politics of the Conservation Reserve Program," *Ecological Indicators*, August 2001, *1* (1), 11–20.

_ **,** _ **, and David Arnold**, "Land Use Impacts of the Conservation Reserve Program: An Analysis of Rejected CRP Offers," January 2022.

Salzman, James, Genevieve Bennett, Nathaniel Carroll, Allie Goldstein, and Michael Jenkins, "The global status and trends of Payments for Ecosystem Services," *Nature Sustainability*, March 2018, *1* (3), 136–144.

Samuelson, Paul A., "The Pure Theory of Public Expenditure," *The Review of Economics and Statistics*, 1954, *36* (4), 387–389.

Sant'Anna, Marcelo Castello Branco, "Empirical analysis of scoring auctions for oil and gas leases," October 2017.

Scott, Paul T., "Dynamic Discrete Choice Estimation of Agricultural Land Use," December 2013.

_ **, Katie Hoover, and Jonathan Ramseur**, "Agriculture and Forestry Offests in Carbon Markets: Background and Selected Issues," Congressional Research Service Report November 2021.

Tebaldi, Pietro, "Estimating Equilibrium in Health Insurance Exchanges: Price Competition and Subsidy Design under the ACA," March 2022.

Teytelboym, Alexander, "Natural capital market design," *Oxford Review of Economic Policy*, January 2019, *35* (1), 138–161.

Torchiana, Adrian L., Ted Rosenbaum, Paul T. Scott, and Eduardo Souza-Rodrigues, "Improving Estimates of Transitions from Satellite Data: A Hidden Markov Model Approach," December 2022.

Train, Kenneth E., *Discrete Choice Methods with Simulation*, 2nd edition ed., Cambridge ; New York: Cambridge University Press, June 2009.

van Benthem, Arthur and Suzi Kerr, "Scale and transfers in international emissions offset programs," *Journal of Public Economics*, November 2013, *107*, 31–46.

Vörösmarty, C. J., P. B. McIntyre, M. O. Gessner, D. Dudgeon, A. Prusevich, P. Green, S. Glidden, S. E. Bunn, C. A. Sullivan, C. Reidy Liermann, and P. M. Davies, "Global threats to human water security and river biodiversity," *Nature*, September 2010, *467* (7315), 555–561.

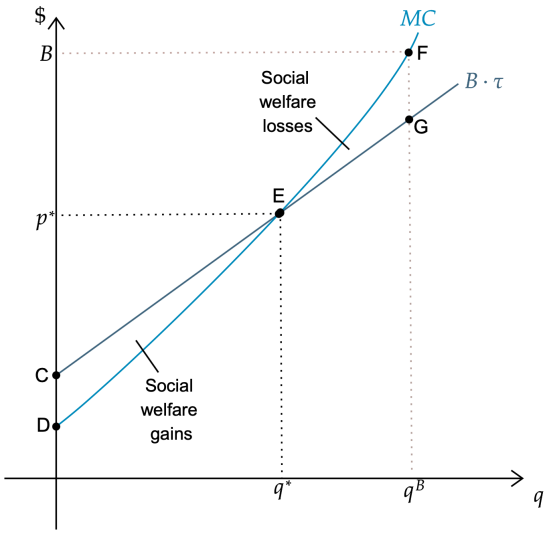
West, Thales A. P., Jan Börner, Erin O. Sills, and Andreas Kontoleon, "Overstated carbon emission reductions from voluntary REDD+ projects in the Brazilian Amazon," *Proceedings of the National Academy of Sciences*, September 2020, *117* (39), 24188–24194.

_ , Sven Wunder, Erin O. Sills, Jan Börner, Sami W. Rifai, Alexandra N. Neidermeier, Gabriel P. Frey, and Andreas Kontoleon, "Action needed to make carbon offsets from forest conservation work for climate change mitigation," *Science*, August 2023, *381* (6660), 873–877.

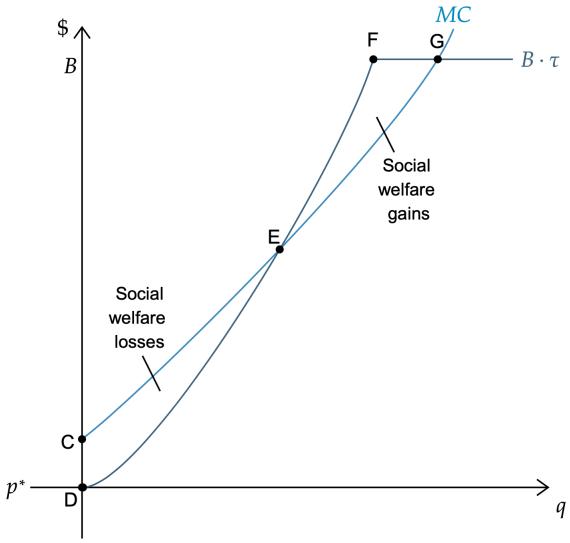
1.9 Figures and Tables

Figure 1.1: Graphical Analysis

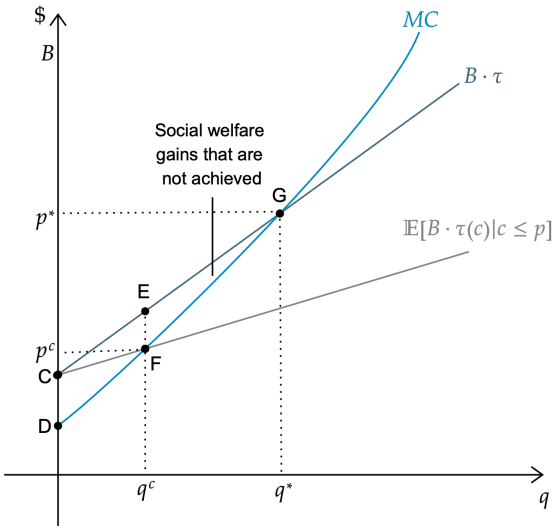
(a) Efficient allocation can be implemented



(b) Efficient allocation cannot be implemented



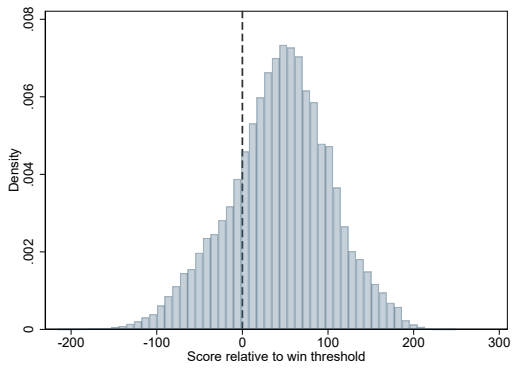
(c) Inefficiency in competitive markets



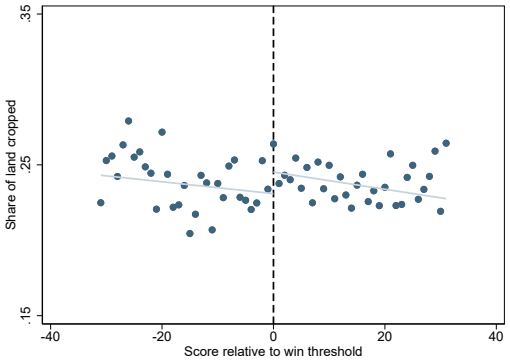
Notes: Figure describes markets characterized by marginal cost ($MC = F_C^{-1}(q)$) and contract value ($B \cdot \tau$) curves. The horizontal axis is the share of the population ordered by costs of contracting c_i. B denotes the social benefits of $a_{i1} = 1$. $B \cdot \tau$ denotes the incremental social benefits of contracting (contract value), relative to no contract, at each quantile of landowner costs of contracting. The vertical distance between the $B \cdot \tau$ and MC curves represents the social surplus from contracting. Upwards-sloping $B \cdot \tau$ curves illustrate markets with adverse selection. Panel (a) documents a population distribution in which the efficient allocation (defined in equation (1.5)) can be implemented with the socially optimal uniform price p^* and panel (b) documents a population distribution in which it cannot. Panel (a) also demonstrates the social welfare losses from mis-pricing (at B) (triangle EFG). Panel (c) includes a curve defining the average contract value of all landowners selecting into the market at any given price p, $\mathbb{E}[B \cdot \tau(c) \,|\, c \leq p]$. This defines the value of a contract to a price-taking buyer in a stylized competitive (offset) market. The intersection of the MC and average contract value curves define a competitive market equilibrium price p^c. In panel (c), adverse selection limits trade in competitive markets with social welfare gains in triangle EFG that are not achieved.

Figure 1.2: Regression Discontinuity Validity and First Stage

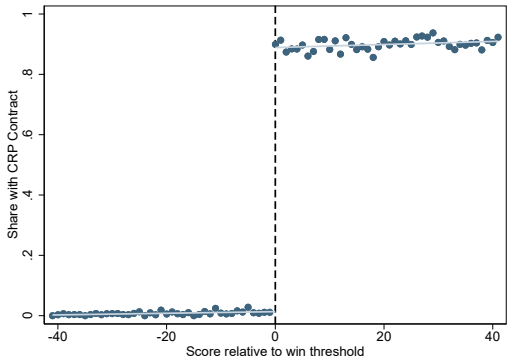
(a) Histogram of running variable



(b) Placebo: pre-auction



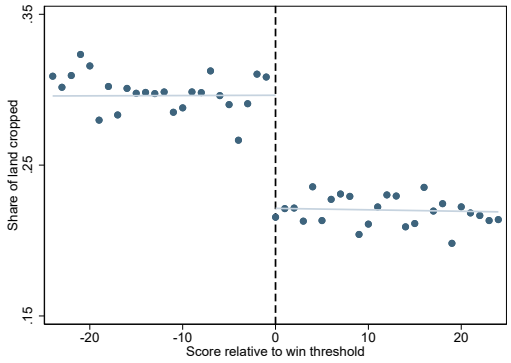
(c) First stage: contracting



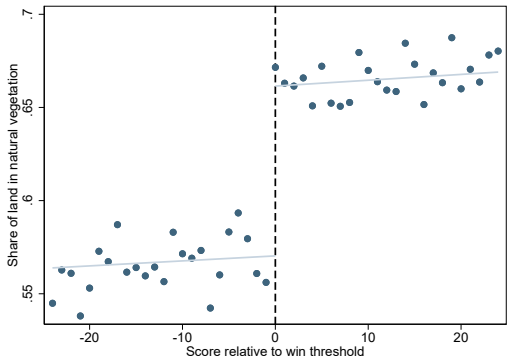
Notes: Panel (a) presents a histogram of bidders' scores in the auction relative to that auction's win threshold, $S_{ig} - S_g$, pooled across auctions. This is the running variable for the regression discontinuity design: bidders above zero win the auction. Panels (b) and (c) present raw data and estimates from equation (1.7). Panel (b) is estimated for $r(i,t) \leq 0$ (pre-auction), and panel (c) is estimated for $r(i,t) > 0$ (post-auction). The outcome in panel (b) is the share of the bidder's land that is cropped, measured in the remote sensing data. The outcome in panel (c) is an indicator for a bidder obtaining a CRP contract. Positive numbers on the x-axis correspond to winning scores, negative numbers correspond to losing scores. In panel (a), each observation is a bidder, in panels (b) and (c), each observation is a bidder-year.

Figure 1.3: The Effect of a CRP Contract on Land Use

(a) Share of land cropped

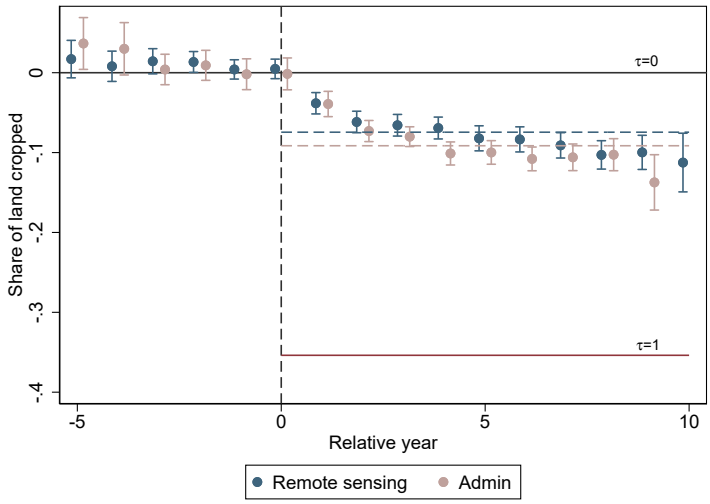


(b) Share of land in natural vegetation



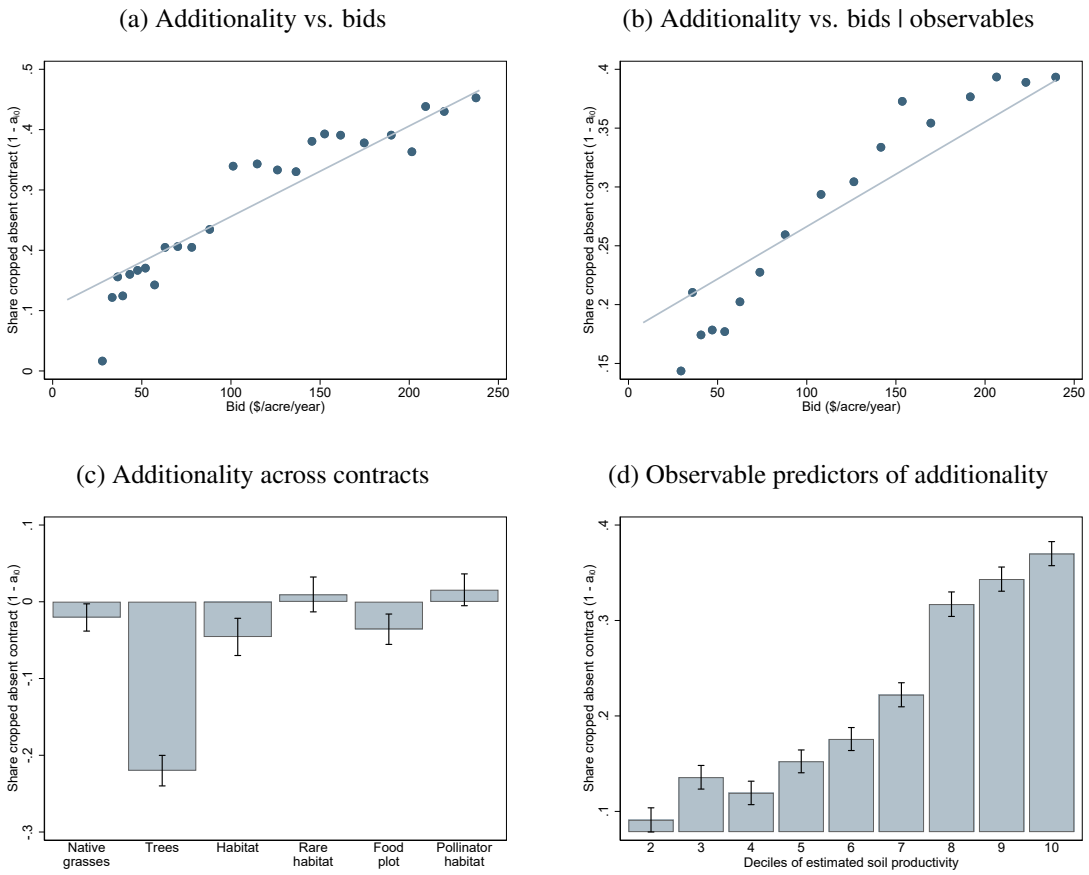
Notes: Panels (a) and (b) present raw data and estimates from equation (1.7) for $r(i,t) > 0$ (post-auction). Outcomes are the share of the bidder's land that is cropped (a) and the share of the bidder's land that is in natural vegetation (trees, grassland, shrubs, and wetlands) (b), both measured in the remote sensing data. The running variable is the difference between each bidder's score and the threshold score. Positive numbers on the x-axis correspond to winning scores, negative numbers correspond to losing scores. Each observation is a bidder-year. Appendix Figure A.6 provides corresponding figures with outcomes measured in the administrative data. Corresponding coefficient estimates and standard errors presented in Table 1.2.

Figure 1.4: Regression Discontinuity Estimates of Additionality



Notes: Figure plots coefficient estimates from equation (1.6). The outcome is the share of each bidder's land that is cropped, measured with both remote sensing and administrative datasets. The x-axis is the year relative to the year of each bidder's auction: $r(i,t) = t - t_{g(i)}$. Positive years correspond to post-auction years. Each point is a regression discontinuity coefficient. Dashed lines indicate the pooled post-auction treatment effects (equation (1.7) estimated for $r(i,t) > 0$). The black line at 0 ($\tau = 0$) and red line at -.35 ($\tau = 1$) indicate the implied effect size if $a_{i0} = 1\ \forall i$ and $a_{i0} = 0\ \forall i$, respectively. $\tau = 1$ represents a benchmark where all landowners are additional. This is calculated as the share of land contracting in the MSE-optimal bandwidth (Calonico et al., 2014) used to estimate the RD. Each observation is a bidder-year. Standard errors are clustered at the bidder level. Ten years is the full duration of a CRP contract. Corresponding coefficient estimates and standard errors presented in Table 1.2.

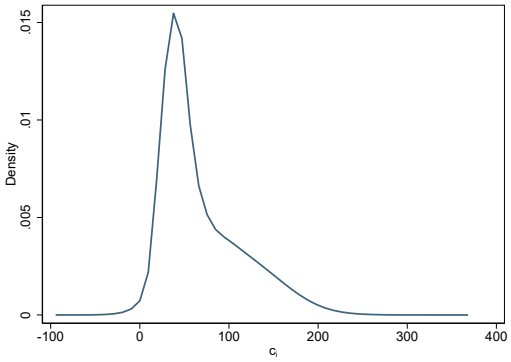
Figure 1.5: Testing for Asymmetric Information



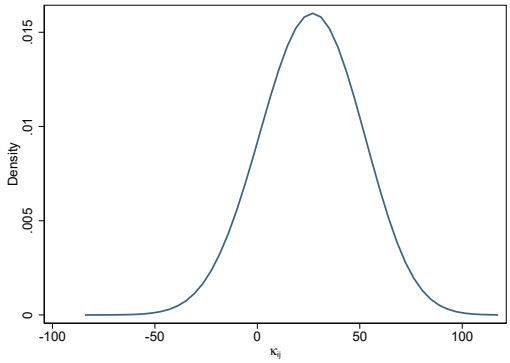
Notes: Figures present visual representations of estimates of equation (1.8). All regressions control for landowner characteristics in the scoring rule: whether a bidder is in a wildlife priority zone, estimates of groundwater quality, estimates of surface water quality, estimates of wind and water erosion (deciles), air quality impacts, and whether or not a bidder is in an air quality zone. The outcome variable in all panels is a landowner-specific measure of additionality ($1 - a_{i0}$). This is calculated as the share of all fields bid into the CRP mechanism that are cropped post auction for rejected landowners. The sample is restricted to the 2016 auction, in which 82% of bidders are rejected and the delineations of bid fields are observed. Cropping on bid fields is measured in 2017-2020 in the remote sensing data (see Figure A.8 for corresponding figures using the administrative data). Panel (a) is a binned scatterplot correlating the dollar bid (per acre, per year) with additionality, conditional on characteristics included in the scoring rule. Panel (b) adds controls for interaction terms of prior land use (quartiles of share of land cropped prior to bidding and re-enrolling CRP status) and deciles of estimated soil productivity. Panel (c) plots relative additionality by the chosen contract in the bid, relative to an omitted category of introduced grasses. Panel (d) plots relative additionality by deciles of estimated soil productivity. Standard errors clustered at the bidder level.

Figure 1.6: Estimated Landowner Cost Distribution

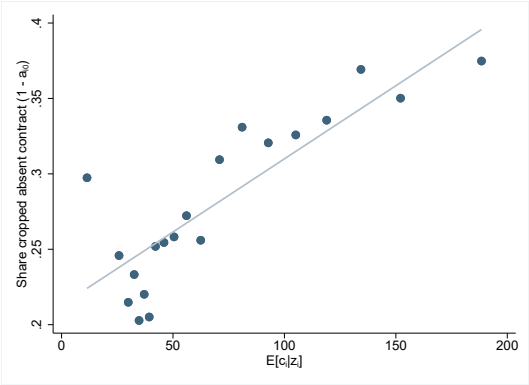
(a) Base cost c_i



(b) Top-up costs κ_{ij}

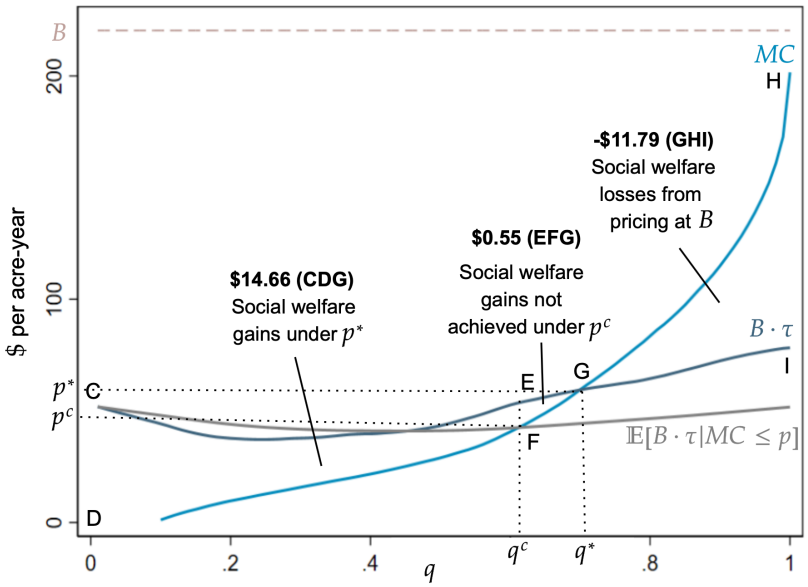


(c) Land use vs. revealed preference cost estimates



Notes: Panels (a) and (b) present kernel density plots of estimates of the base cost c_i (a) and top-up cost κ_{ij} (b) of contracting. Panels (a) and (b) pool bidders across auctions. Costs are estimated using revealed preferences in optimal bidding (equation (1.10)). See Section 1.5.2 for estimation details. Panel (c) correlates expected base costs, c_i, conditional on observable characteristics $\mathbf{z}_i$, with land use outcomes measuring landowner additionality in the remote sensing data. Panel (c) is restricted to the 2016 auction and the 82% of bidders who lose (see Section 1.4.2 for more details). $\mathbf{z}_i$ includes interactions of soil productivity, prior CRP, and prior land use. Costs are reported in dollars per acre per year.

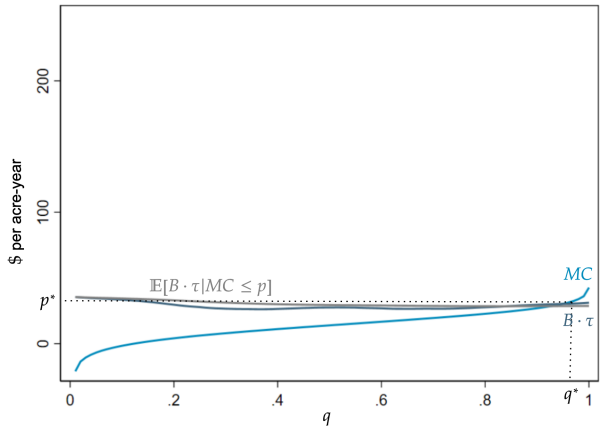
Figure 1.7: Empirical Graphical Analysis



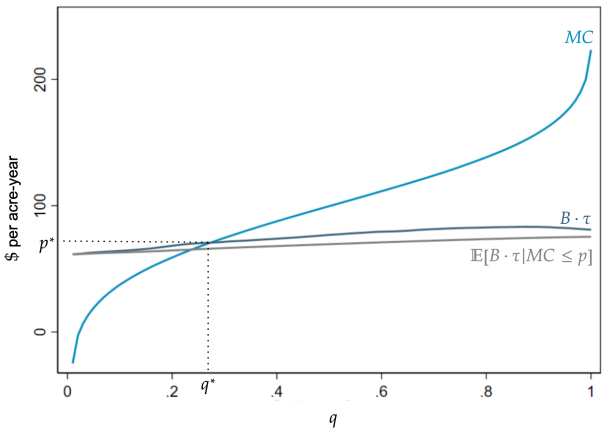
Notes: Figure presents the empirical version of Figure 1.1 for the base contract. The horizontal axis is the share of the population, ordered by contracting costs. Values are reported in dollars per acre per year. The *MC* curve is the inverse distribution function of the minimum cost to fulfill the base contract. B denotes the average value of the base contract action, calculated as described in Appendix 1.10.5. $B \cdot \tau$ denotes the incremental value of contracting, relative to no contract, averaged at each quantile of the population distribution of the base costs of contracting. The vertical distance between the $B \cdot \tau$ and *MC* curves represents the social surplus from contracting at each quantile of the population distribution of contracting costs. The upwards-sloping $B \cdot \tau$ curve illustrates the presence of adverse selection in the market. The intersection of the *MC* and $B \cdot \tau$ curve denotes the socially-optimal uniform price, p^*. Triangle CDG represents social welfare gains under the socially-optimal price. The triangle GHI represents social welfare losses from mispricing at B. The average contract value curve calculates the average $B \cdot \tau$ of all landowners selecting into the market at any given price p, $\mathbb{E}[B \cdot \tau | MC \leq p]$. This defines the value of a contract to a price-taking buyer in a stylized competitive (offset) market. The intersection of the *MC* and average contract value curves define a competitive market equilibrium price p^c. Adverse selection limits trade in competitive markets leading to social welfare gains that are not realized (triangle EFG).

Figure 1.8: Empirical Graphical Analysis: Heterogeneity Across Observables and Contracts

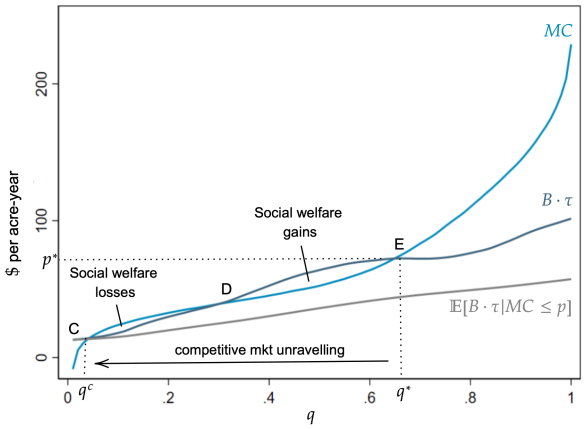
(a) Lowest quintile soil productivity



(b) Highest quintile soil productivity

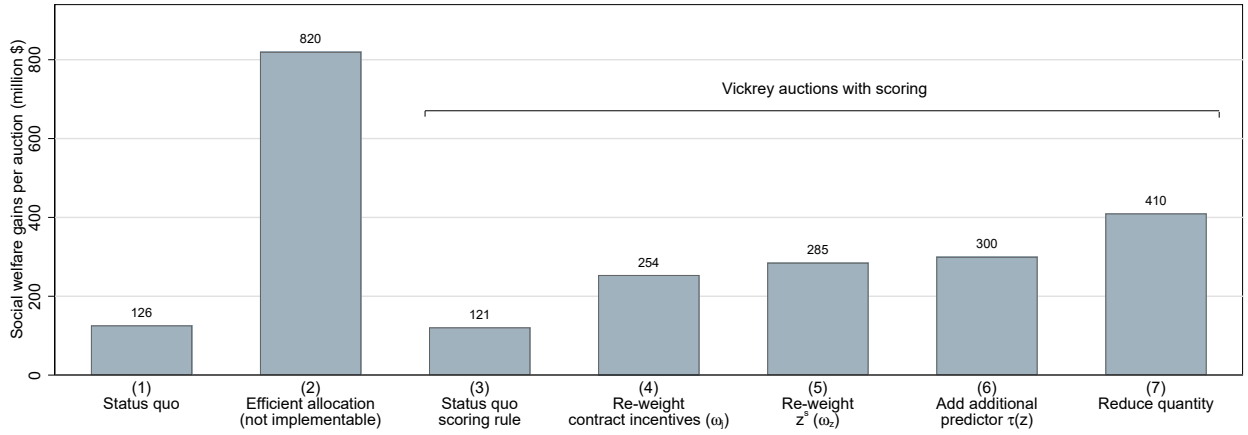


(c) Tree planting and maintenance contracts



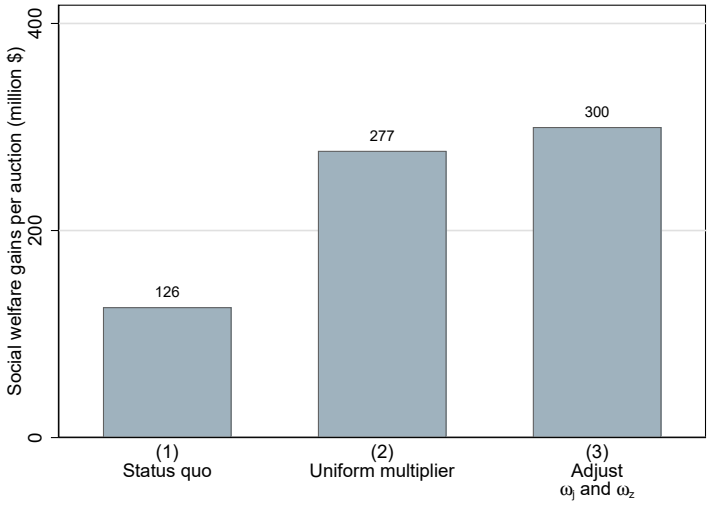
Notes: Figures presents empirical version of Figure 1.1. Panels (a) and (b) calculate the *MC* curve as the inverse distribution function of the minimum cost to fulfill the base contract, split by whether landowners are in the lowest or highest quintile of soil productivity. In panel (c), the *MC* curve is calculated as the inverse distribution function of the minimum cost to fulfill a tree planting and maintenance contract. The horizontal axis is the share of the population, ordered by contracting costs for the base contract in each sub-population ((a) and (b)) and for tree planting and maintenance contracts (c). $B \cdot \tau$ denotes the incremental value of contracting, relative to no contract, averaged at each quantile of contracting costs for the base contract in each sub-population ((a) and (b)) and for tree planting and maintenance contracts (c). The vertical distance between the $B \cdot \tau$ and *MC* curves represents the social surplus from contracting at each quantile of contracting costs. p^* denotes the socially-optimal price, set at the intersection of the $B \cdot \tau$ and *MC* curves. The average contract value curve (gray) calculates the average $B \cdot \tau$ of all landowners selecting into the market at any given price p, $\mathbb{E}[B \cdot \tau | MC \leq p]$. This defines the value of a contract to a price-taking buyer in a stylized competitive (offset) market. In panel (c), the efficient allocation defined in equation (1.5) cannot be implemented. The stylized competitive market also unravels.

Figure 1.9: Social Welfare Under Alternative Auctions



Notes: Figure presents estimates of the social welfare gains (defined in equation (1.14)) under status quo and alternative auctions. Results reported in million dollars per auction. All auctions impose that each landowner obtains at most one contract and that total contracts awarded cannot exceed the status quo. Bar (1) simulates the status quo. Bar (2) calculates the social welfare gains under an efficient allocation that allocates contracts using all $\mathbf{z}_i$ and (c_i, κ_i) to maximize equation (1.14). Due to adverse selection, this allocation may not be implementable. Bars (3)-(7) calculate social welfare under alternative Vickrey auctions with scoring (see Section 1.6.2 for more details). Bars (3)-(6) hold quantity (the number of landowners allocated contracts) constant at the status quo and change the scoring rule $s_j(\mathbf{z}_i)$ defined in equation (1.15). Bar (3) uses the existing scoring rule $s_j(\mathbf{z}_i) = B_j(\mathbf{z}_i^s)$. Bar (4) uses a scoring rule with the social-surplus maximizing incentives across contracts (ω_j). Bar (5) uses a scoring rule with the social-surplus maximizing asymmetry across bidders using characteristics already in the scoring rule $(\mathbf{z}_i^s)$. Bar (6) adds an additional characteristic to the scoring rule, a prediction of $\tau(\mathbf{z}_i, c_i, \kappa_i)$ based on immutable characteristics of landowners already collected by the USDA (deciles of soil productivity and wind and water erosion). Bar (7) uses the same scoring rule as bar (6) but reduces the number of contracts allocated to landowners: only landowners with positive scoring-rule-implied social surplus $\max_j s_j(\mathbf{z}_i) - c_i - \kappa_{ij} \geq 0$ are awarded contracts. See each bar's corresponding column in Table 1.5 for more details.

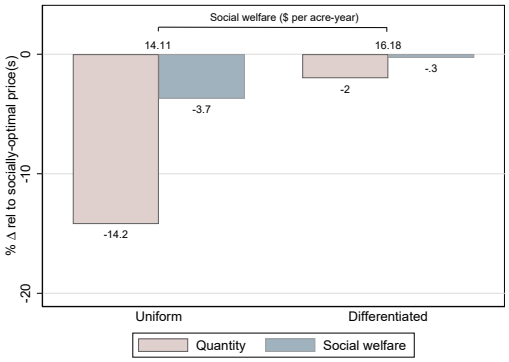
Figure 1.10: Mechanisms: Uniform vs. Heterogeneous Scoring Rule Adjustments



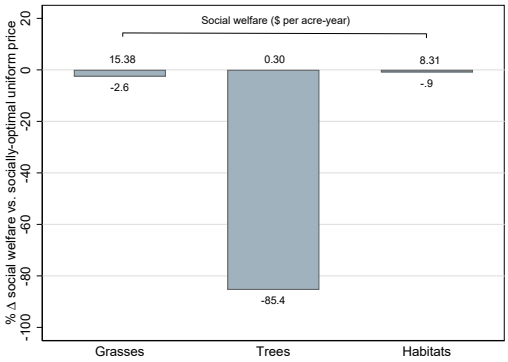
Notes: Figure presents estimates of the social welfare gains (defined in equation (1.14)) under status quo and alternative auctions. Results reported in million dollars per auction. All auctions impose that each landowner obtains at most one contract. All auctions hold constant the total number of landowners awarded contracts at the status quo. Bar (1) simulates the status quo. Bars (2) and (3) simulate Vickrey auctions with scoring (see Section 1.6.2 for more details). Bar (2) uses a scoring rule $s_j(\mathbf{z}_i^s) = \theta \cdot B_j(\mathbf{z}_i^s)$ for a uniform multiplier θ that maximizes equation (1.14). Bar (3) corresponds to Bar (6) in Figure 1.9: it uses the social welfare maximizing $\omega_j$and $\omega_{\mathbf{z}}$ using all characteristics in the rule and a prediction of $\tau(\mathbf{z}_i, c_i, \kappa_i)$ based on immutable characteristics of landowners already collected by the USDA (deciles of soil productivity and wind and water erosion).

Figure 1.11: Offset Market Design

(a) Uniform vs. differentiated market



(b) By contract type



Notes: Figures describe social welfare and quantities traded under a stylized competitive offset equilibrium versus under socially-optimal prices. The competitive market equilibrium is calculated under the assumption that buyers have the same full-information preferences as the USDA and form expectations over the value of any contract given the equilibrium price(s). Panel (a) restricts to the base contract and reports quantities traded and social welfare under a competitive market equilibrium relative to socially-optimal prices in a uniform and a differentiated market. In the uniform market, there is only a single socially-optimal price and market-clearing condition. In the differentiated market, the market is segmented into deciles of predicted contract value (based on $\mathbf{z}_i^s$, soil productivity, and erosion). Panel (b) plots the percent reduction in social welfare under a stylized competitive market equilibrium relative to socially-optimal prices under three different hypothetical markets, each with only one contract traded at a uniform price. The numbers above the bars in panels (a) and (b) tabulate total social welfare (per acre-year) in each competitive market.

Table 1.1: Summary Statistics

	All agricultural land		Bidders		Bid fields	
	Remote sensing	Admin	Remote sensing	Admin	Remote sensing	Admin
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Land use						
Share cropped	0.30	0.28	0.21	0.21	0.21	0.18
Share corn	0.11	0.11	0.07	0.08	0.07	0.06
Share soybean	0.11	0.10	0.06	0.07	0.07	0.07
Share fallow	0.02	0.01	0.03	0.01	0.05	0.03
Share nat. veg. or grassland	0.55		0.70		0.65	
Panel B. Land characteristics						
Size (acres)	160.7		250.6			
	(2690.7)		(506.5)			
Soil productivity ($/acre)	92.4		86.9			
	(63.2)		(58.5)			
Enviro sensitivity (points)	53.5		86.5			
	(29.8)		(33.7)			
Panel C. Bid characteristics						
Rental rate ($/acre/year)			83.0			
			(56.4)			
Acres bid			84.1			
			(136.3)			
Share re-contracting			0.70			
Contract action = grasses			0.67			
Contract action = trees			0.12			
Contract action = habitat			0.21			
Share contracting			0.81			
N bidders / auction			36,763			
N	7,890,426		258,286		61,703	

Notes: Table presents summary statistics of all agricultural landowners (columns (1)-(2)), bidding landowners (columns (3)-(4)), and bid fields (columns (5)-(6)), defined as the delineated land area entered into the mechanism to be awarded a CRP contract (observed only for bidders in the 2016 auction). Standard deviations in parenthesis. Panel A reports land use outcomes in the remote sensing (CDL) and admin (Form 578) data. All land use outcomes are reported for the year prior to bidding among bidders. Years in columns (1)-(2) are re-weighted to match the distribution of bidder-years. Columns (1) and (2) includes both eligible non-bidders and ineligible land. Land use categories follow Lark et al. (2017). Crop outcomes exclude alfalfa and hay. Soil productivity is calculated by NASS and is reported in dollars per acre. Environmental sensitivity points are the points given for characteristics of land in the scoring rule. Rental rate is reported in dollars per acre per year and is the dollar component of the bid in the auction. Acres bid is the total acreage entered into the auction to be awarded a CRP contract. Grasses, trees, and habitat contract indicators are aggregated over the menu of possible contracts.

Table 1.2: Regression Discontinuity Coefficient Estimates

	Remote sensing	Admin
	(1)	(2)
Panel A: Main outcome: share of land cropped		
Pre-auction (placebo)	0.014	0.009
	(0.007)	(0.006)
Post-auction (pooled sign-ups)	-0.075	-0.091
	(0.007)	(0.006)
Implied additionality	21%	26%
Post-auction (full contract duration: 2010-2020)	-0.109	
	(0.020)	
Implied additionality	31%	
Panel B: Other outcomes		
Corn	-0.015	-0.023
	(0.003)	(0.003)
Soybean	-0.018	-0.026
	(0.003)	(0.003)
Fallow	-0.008	-0.011
	(0.002)	(0.001)
Natural vegetation or grassland	0.091	
	(0.007)	
Panel C: Spillovers to non-bid fields		
Share of non-bid fields cropped	-0.001	-0.000
	(0.015)	(0.015)
N bidders	258,286	258,286
N bidder-years	3,099,432	1,808,002

Notes: Table presents coefficient estimates from equation (1.7) with land use outcomes measured in the remotely sensed (column (1)) and administrative (column (2)) data. All results use a local linear regression on either side of the win threshold in the MSE-optimal bandwidth (Calonico et al., 2014). The full-contract duration specification restricts to the 2009 auction, others pool all auctions with post-period data (2009, 2011, 2012, 2013, and 2016). The pooled post-period includes an average of 7-8 post-auction years. Natural vegetation or grassland is only observed in remotely sensed data. Calculations of implied additionality divide the treatment effect estimates by the amount of land contracting among winning bidders in the MSE-optimal bandwidth. Panel C estimates the effect of a CRP contract on non-bid, and therefore non-contracting, fields to test for spillovers. This analysis is restricted to the 2016 auction. Standard errors are clustered at the bidder level.

Table 1.3: Mean Landowner Costs of Contracting

	All	Landowners with above median soil productivity
	(1)	(2)
Base cost (c_i)	67.49	87.05
Top-up cost (κ_{ij})		
Introduced grasses (normalized)	0.	0.
Native grasses	0.11	3.38
Trees	24.41	26.65
Habitat	14.87	17.49
Rare habitat	15.33	17.98
Wildlife food plot	18.58	15.32
Pollinator habitat	18.03	17.54

Notes: Table presents estimated mean landowner costs of contracting for the base cost c_i and top-up cost κ_{ij} reported in dollars per acre per year. The cost of each contract j is defined as $c_i + \kappa_{ij}$. Costs are estimated using revealed preferences in optimal bidding (equation (1.10)). See Section 1.5.2 for estimation details. Column (1) presents mean costs for all bidders across all auctions, and column (2) restricts to landowners with above median soil productivity. See Appendix Table A.8 for a comparison to administrative data.

Table 1.4: Additionality as a Function of Landowner Costs

	Estimates of $\tau(\mathbf{z}_i, c_i, \kappa_i)$			
	(1)	(2)	(3)	(4)
β: coefficient on base cost (c_i)	0.0018	0.0020	0.0007	-0.0002
	(0.0002)	(0.0002)	(0.0003)	(0.0004)
α: coefficient on top-up cost (κ_{ij})				
Trees			0.0035	0.0046
			(0.0002)	(0.0005)
Native grasses				-0.0011
				(0.0006)
Habitat				-0.0004
				(0.0005)
Rare habitat				0.0027
				(0.0007)
Wildlife food plot				0.0031
				(0.0006)
Pollinator habitat				0.0010
				(0.0005)
Includes $\mathbf{z}_i^s$	✓	✓	✓	✓
Includes prior land use and soil prod. $\mathbf{z}_i$		✓	✓	✓

Notes: Table presents select coefficient estimates of $\tau(\mathbf{z}_i, c_i, \kappa_i)$ (equation (1.12)). Coefficients measure how additionality varies with a \$1 per acre, per year change in costs. Positive coefficients indicate a positive correlation between costs of contracting and additionality, or adverse selection in the market. Parameter estimates obtained via the Method of Simulated Moments estimator described in Section 1.5.2. This estimator matches moments of land use in the remote sensing data (for losing bidders in the 2016 auction) and bids given simulated (c_i, κ_i) and optimal bidding in equation (1.10). All specifications include flexible controls for the components of the scoring rule excluding landowners' Wildlife Priority Zone and Air Quality Zone status. Columns (2)-(4) control for the 32 cells of soil productivity, prior CRP status, and prior cropping status that parameterize bidder costs. Standard errors are calculated using 100 bootstrap draws and do not (yet) account for variance in the Step 2 estimates.

Table 1.5: Outcomes Under Alternative Auctions

	Status quo	Efficient allocation	Vickrey auctions with scoring				
			Status quo rule	Re-weight contracts (ω_j)	Re-weight $\mathbf{z}_i^s$ ($\omega_{\mathbf{z}}$)	Add $\hat{\tau}(\mathbf{z}_i)$ (incl. $\mathbf{z}_i$ not in status quo rule)	Reduce quantity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Welfare and spending (million $ per auction):							
Social welfare	126	820	121	254	285	300	410
USDA spending	1,323		2,033	1,760	1,703	1,724	936
Landowner surplus	546		906	1,176	1,127	1,147	580
Environmental value	902	1,239	1,130	838	861	876	766
Panel B. Other outcomes							
Additionality | contract	0.206	0.424	0.199	0.200	0.209	0.213	0.215
Share awarded contract	0.81	0.55	0.81	0.81	0.81	0.81	0.70

Notes: Table presents results under current and alternative auctions. Panel A tabulates social welfare (equation (1.14)), USDA spending, landowner surplus, and environmental value $\sum_i \sum_j B_j(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i) \cdot x_{ij}$ in million dollars per auction. Panel B tabulates average additionality of contracting landowners and the share of landowners with a contract. All auctions impose that each landowner obtains at most one contract and that total contracts awarded cannot exceed the status quo. Column (1) simulates the status quo. Column (2) simulates an efficient allocation that allocates contracts using all $\mathbf{z}_i$ and (c_i, κ_i) to maximize equation (1.14). Due to adverse selection, this allocation may not be implementable. Columns (3)-(7) simulate alternative Vickrey auctions with scoring (see Section 1.6.2 for more details). Columns (3)-(6) hold quantity (the number of landowners allocated contracts) constant at the status quo and change the scoring rule $s_j(\mathbf{z}_i)$ defined in equation (1.15). Column (3) uses the existing scoring rule $s_j(\mathbf{z}_i) = B_j(\mathbf{z}_i^s)$. Column (4) uses a scoring rule with the social-surplus maximizing incentives across contracts (ω_j). Column (5) uses a scoring rule with the social-surplus maximizing asymmetry across bidders using characteristics already in the scoring rule $(\mathbf{z}_i^s)$. Column (6) adds an additional characteristic to the scoring rule, a prediction of $\tau(\mathbf{z}_i, c_i, \kappa_i)$ using immutable characteristics of landowners already collected by the USDA but not all included in the status quo scoring rule (deciles of soil productivity and wind and water erosion). Column (7) uses the same scoring rule as column (6) but reduces the number of contracts allocated to landowners: only landowners with positive scoring-rule-implied social surplus $\max_j s_j(\mathbf{z}_i) - c_i - \kappa_{ij} \geq 0$ are awarded contracts. Each column corresponds to a bar in Figure 1.9.

1.10 Appendix

1.10.1 Institutional Appendix: The CRP Mechanism

The scoring rule depends on characteristics of the land, the conservation action defined in the contract, and the dollar component of the bid (the bid rental rate). We describe the details associated with each of these components below. The details of the scoring rule are published each year in EBI Factsheets.[49]

Land characteristics The characteristics that influence the scoring rule include:

- **Whether a bidder is in a Wildlife Priority Zone (WPZ)**, defined high priority wildlife geographic areas. 30 points.
- **Whether a bidder is in a Water Quality Zone (WQZ)**, areas with high value to improving ground or surface water quality. 30 points.
- **Groundwater quality**: an evaluation of the predominant soils, potential leaching of pesticides and nutrients into groundwater, and the impact to people who rely on groundwater as a primary source of drinking water. Continuous score: 0 to 25 points.
- **Surface water quality**: an evaluation of the amount of sediment (and associated nutrients) that may be delivered into streams and other water courses. Continuous score: 0 to 45 points.
- **Erosion potential**: Continuous score of 0 to 100 points depending on the Erodibility Index.
- **Air quality**: an evaluation of the air quality improvements by reducing airborne dust and particulate caused by wind erosion from cropland. Continuous score of 0 to 30 points depending on wind speed, wind direction, and the duration of wind events and soil erodibility.
- **Whether a bidder is in an Air Quality Zone (AQZ)**. 5 points.

These characteristics depend on a bidder's location and not their bid, i.e. they determine bidder asymmetry in the scoring rule. These characteristics are known for every agricultural field in the US.

[49] See an EBI Factsheet for an example.

Heterogeneous contracts defined by conservation actions Conservation actions can be grouped into two categories: a primary cover, described in Table A.1, which covers the total area offered into the CRP, and an (optional) additional upgrade action, described in Table A.2, which can be offered in addition to the primary cover on a smaller area. In total, there are 36 possible contracts: 12 primary covers interacted with three upgrade cover options (including no upgrade).

Table A.1: Contract Action Choices: Primary Covers

Short name	Description
Grasses 1	Permanent introduced grasses and legumes (CP1): Existing stand of one to three species or planting new stand of two to three species of an introduced grass species
Grasses 2	Permanent introduced grasses and legumes (CP1): Existing stand or planted mixture (minimum of four species) of at least 3 introduced grasses and at least one forb or legume species best suited for wildlife in the area.
Grasses 3	Permanent native grasses and legumes (CP2): Existing stand (minimum of one to three species) or planting mixed stand (minimum of three species) of at least two native grass species at least one forb or legume species beneficial to wildlife.
Grasses 4	Permanent native grasses and legumes (CP2): Existing stand or planting mixed stand (minimum of five species) of at least 3 native grasses and at least one shrub, forb, or legume species best suited for wildlife in the area.
Trees 1	Tree planting (softwoods) (CP3): Southern pines, northern conifers, or western pines – solid stand of pines/conifers/softwoods (existing, according to state developed standards, or planted at more than 550 (southern pines), 850 (northern conifers), or 650 (western pines) trees per acre).
Trees 2	Tree planting (softwoods) (CP3): Southern pines, northern conifers, or western pines – pines/confiers/softwoods existing or planted at a rate of 500-550 (southern pines), 750-850 (northern conifers), or 550-650 (western pines) per acre depending on the site index (state-developed standards) with 10-20% openings managed to a CP4D wildlife cover.
Trees 3	Hardwood tree planting (CP3A): Existing or planting solid stand of nonmast producing hardwood species.
Trees 4	Hardwood tree planting (CP3A): Existing or planting solid stand of single hard mast producing species.
Trees 5	Hardwood tree planting (CP3A): Existing or planting mixed stand (three or more species) or hardwood best suited for wildlife in the area or existing or planting stand of longleaf pine or atlantic white cedar – planted at rates appropriate for the site index.
Habitat 1	Permanent wildlife habitat, noneasement (CP4D): Existing stand or planting mixed stand (minimum of four species) of either grasses, trees, shrubs, forbs, or legumes planted in mixes, blocks, or strips best suited for various wildlife species in the area. A wildlife conservation plan must be developed with the participant.
Habitat 2	Permanent wildlife habitat, noneasement (CP4D): Existing stand or planting mixed stand (minimum of five species) or either predominantly native species including grasses, forbs, legumes, shrubs, or trees planted in mixes, blocks, or strips best suited to providing wildlife habitat. Only native grasses are authorized. A wildlife conservation plan must be developed with the participant.
Habitat 3	Rare and declining habitat restoration (CP25): Existing stand or seeding or planting will be best suited for wildlife in the area. Plant species selections will be based upon Ecological Site Description data.

Notes: Table describes the menu of primary cover actions.

Table A.2: Contract Action Choices: Upgrades

Short name	Description
No upgrade	Primary cover only
Wildlife food plot	Wildlife food plots are small plantings in a larger area
Pollinator habitat	Existing stand or planting (minimum of .5 acres) of a diverse mix of multiple species suited for pollinators

Notes: Table describes the menu of upgrade actions.

We obtain the points associated with each of the contract options, defined by the actions in Tables A.1 and A.2 from the EBI Fact Sheets. The point values assigned to the different contracts can vary across bidders based on whether or not a bidder is in a Wildlife Priority Zone (WPZ).

Bid rental rate The scoring rule is non-linear in r_i. The existence of bid caps make some choices infeasible if $r_i > \bar{r}_i$, where $\bar{r}_i$ denotes the i-specific bid cap. The scoring rule also includes non-linearities based on the amount a bidder bids below the bid cap with kinks at 10% and 15% below the bid cap.[50] The weight on this component is announced only after bids are collected, but it has remained essentially constant throughout our sample period, so we treat it as known.

An example menu The mechanism implies a menu of payments for each contract at each score. These menus differ by observable characteristics of landowners due to asymmetry in the existing rule. Table A.3 describes an example menu.

[50] We observe bunching at the kink points, suggesting that bidders understand the scoring rule and make sophisticated choices in the mechanism.

Table A.3: Payments (for a Target Score) and Market Shares Across Contracts

	Average payment at threshold score	Market share	Average payment at threshold score	Market share	Average payment at threshold score	Market share
	No upgrade		+ wildlife flood plot		+ pollinator habitat	
Intro Grasses 1	28.63	0.140	35.21	0.015	52.91	0.007
Intro Grasses 2	74.30	0.104	77.86	0.022	86.00	0.019
Native Grasses 1	43.64	0.067	49.37	0.005	64.68	0.009
Native Grasses 2	81.00	0.201	83.59	0.023	90.34	0.056
Trees 1	65.13	0.039	69.44	0.003	79.54	0.000
Trees 2	94.73	0.020	96.45	0.003	101.47	0.001
Trees 3	73.29	0.012	76.52	0.001	85.06	0.000
Trees 4	79.54	0.002	82.40	0.000	89.65	0.000
Trees 5	98.14	0.029	99.83	0.003	104.71	0.002
Habitat 1	75.29	0.032	78.72	0.006	86.60	0.001
Habitat 2	81.73	0.039	84.25	0.007	90.84	0.014
Rare Habitat	93.07	0.077	94.82	0.009	99.91	0.025

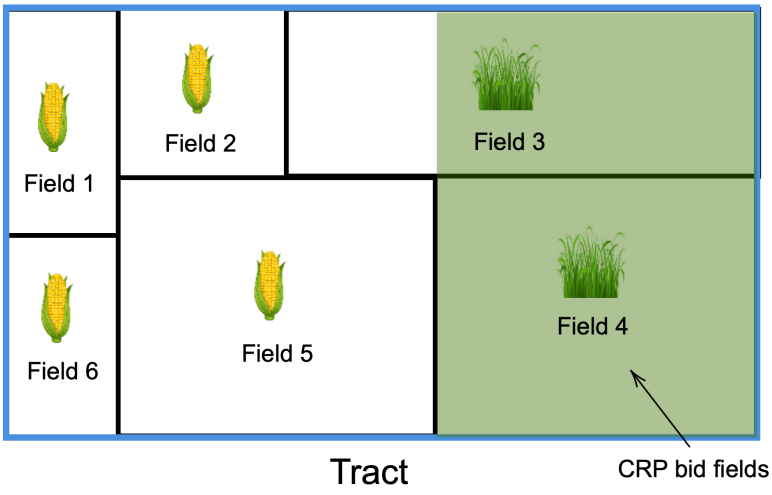
Notes: Table presents the menu of all 36 possible contracts, split into 12 primary covers and three upgrade options. Table reports payments across contracts, calculated as the rental rate per acre per year to reach a given score (held fixed in this table at the threshold score, $\underline{S}$) with a given contract. Payments vary across bidders with heterogeneous $\mathbf{z}_i^s$; this table calculates the averages across all bidders. Table also reports the market shares of each contract, pooled across auctions.

1.10.2 Data Appendix

Agricultural Units: Tracts and Fields Figure A.1 provides an illustrative example of the various agricultural land units.

All agricultural land is the US is divided into fields, or Common Land Units, by the USDA. A field is defined as the smallest unit of land that has: (i) a permanent, contiguous boundary, (ii) a common land cover and land management, and (iii) a common owner.[51] There are 37,480,917 fields in the US (as of 2016), with an average size of 33.82 acres. Each field, by definition, has a single land use.

Figure A.1: Example: Tract, Fields, and Bid Fields



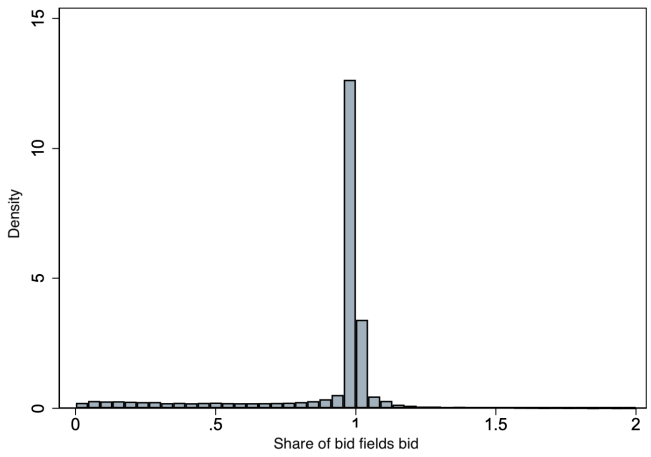
Notes: Figure explains the various geographic units in our dataset. The blue outline is a single tract: this is the unit of landowner (bidder) in our analysis. This tract contains six fields, these are administrative delineations of a tract, each with a single land use. The green shaded area represents an example area bid into the CRP. This could follow field boundaries (as for field 4) or cut into fields (as for field 3).

A tract is a collection of fields under one common ownership that is operated as a farm or part of a farm (a tract is a landowner, or bidder, in our setting). The average tract includes 4.75 fields. Each tract can submit at most one bid into a CRP auction. This bid can include any subset of a tract's fields. A bid is not constrained to bid only entire fields; in principle, a bidder can bid any subset of their land, regardless of field delineations. In practice, a large share of bids follow field boundaries, as illustrated by Figure A.2. In our analyses, we therefore treat bid fields as defining the land offered into the mechanism.

[51] See the Common Land Unit Information Sheet published by the USDA for more details.

Our dataset includes an identifier and the geolocation of each of the bidding tracts, and their subset fields, for all auctions. We only observe the identifiers of the bid fields in 2016.

Figure A.2: Share of Bid Fields Bid into the Mechanism



Notes: Figure shows a histogram of the share of the land of the bid fields that are bid into the CRP (the shaded green area as a share of the total area of fields 3 and 4 in Figure A.1). The mass point at one indicates that the vast majority of bidders bid the entire field.

Remote Sensing Data (CDL) Our first source of land use data is the Cropland Data Layer (CDL) from 2009 to 2020. The CDL is derived from annual satellite imagery at a 30m by 30m resolution (approximately one quarter acre) for the entire contiguous US. The dataset classifies each pixel into over 50 crop categories and over 20 non-crop categories. The CDL is produced by the National Agricultural Statistics Service (NASS), and is trained on administrative data submitted to the USDA for crop insurance purposes (Form 578, discussed in more detail below). The CDL has been used in prior economics research studying agriculture and land use (Scott, 2013; Hagerty, 2022).[52]

Our primary analysis aggregates CDL classifications into super-classes of crop versus non-crop, following (Lark et al., 2017). Also following Lark et al. (2017), our crop classification excludes alfalfa, hay, fallow, and idle cropland. The super-class accuracy of the CDL is high with > 95% average producer's (classified as cropped when truly cropped) and user's (truly cropped when classified as cropped) accuracy in the years 2008-2016 (Lark et al., 2017). Despite this high super-class accuracy, remote sensing classifications are subject to measurement error in classification (Alix-Garcia and Millimet, 2022; Torchiana et al., 2022), particularly when analyzing land use transitions. Moreover, in order to improve accuracy, some states in some years use prior years'

[52] See https://www.nass.usda.gov/Research_and_Science/Cropland/SARS1a.php for more details and Metadata.

CDL as an input into the training algorithm, providing a further source of bias stemming from the classification algorithm.

We merge the CDL to a shapefile of all agricultural fields in the US, which we can then aggregate to landowners (tracts) using USDA identifiers. We merge the CDL data to the geocoded location of the bidder, time-stamped at the point of bidding.

Calculating land use outcomes at the tract level as either the share of pixels that fall into the crop super-class, or a weighted average of field-level (binary) cropping indicators produce similar results. We use the former in our main specifications.

Form 578 Administrative Data Our second source of data (from Form 578) is new to economics research. It is the administrative data submitted to the USDA that the CDL is trained on. The data consist of annual field-level reports of total acreage cropped in detailed crop categories and enrollment in USDA programs. Though Form 578 is self-reported, crop insurance payouts depend on these reports. Unlike the CDL, which has coverage over the entire US, field-level data is only submitted if there is an incentive to do so, i.e. if it is cropped and covered by crop insurance. We assume that all non-reporting fields are not cropped. This is the primary limitation of the administrative data relative to the CDL.

We merge the Form 578 administrative data to bidders based on field and tract identifiers. We construct a panel that tracks changes in field identifiers over time using their geolocation.

NAIP Imagery Our final dataset is derived from the National Agriculture Imagery Program (NAIP) collected via Google Earth Engine. The NAIP is administered through the Forest Service Agency (FSA) of the USDA, and collects 0.6-1m resolution images of all agricultural land during growing season. We obtain NAIP images for enrolled land (the highlighted green area in Figure A.1) to assess compliance with CRP rules. We use high-resolution photographs as classification error in the derived (CDL) data product would mechanically bias toward finding non-compliance.

Figure A.3: Sample Images

(a) Enrolled field

(b) Cropped field

Notes: Example images for classification. For compliance, neither of these are actual images of CRP enrolled fields.

Validating Compliance To assess compliance, we hired and trained two MIT undergraduates (the "reviewers") to classify high resolution aerial photographs (NAIP images) of fields at 1m resolution (see Figure A.3 for examples). We focus on the 2016 auction and images taken between 2017 and 2021. Before asking the reviewers to classify any images, we provided them with a test set of hundreds of images of cropped and uncropped fields across the US. The reviewers used this "test set" to familiarize themselves with the visual patterns of cropped fields (see Figure A.3b). We then provided each of the reviewers with over 1,000 images of CRP enrolled fields and hundreds of placebo cropped fields as attention checks. The reviewers were blind to whether the images were of CRP enrolled fields or placebo cropped fields. Each of the two reviewers were provided with the same images.

Table A.4 presents results for the classification exercise. We restrict to the 83% of CRP images that the reviewers agreed upon for our assessment of compliance to minimize the potential for classification error. We find only 5% of fields to be out of compliance in all post-period years. Once we drop the two "transition" years from 2017-2018, we find even lower rates of non-compliance, and reject rates of non-compliance above 3%. We attribute the difference between columns (1) and

(2) to be driven by the fact that fields appear different when they are transitioning out of cropland, e.g. rows from row cropping may still be visible as new vegetation grows in. While not reported, rates of cropping are substantially higher, at approximately 40%, on placebo cropped fields; the reviewers were making meaningful classifications. We note, however, that this number is far below 100%. This is because we instructed the reviewers to be conservative in their assessment of non-compliance, operating under the (reasonable) null hypothesis that the program is in fact enforced.

Table A.4: Validation of Compliance: $a_{i1} = 1 \; \forall i$

	All post-period years	Drop first two years
	(1)	(2)
Share of enrolled fields classified as cropped	0.054	0.024
	(0.008)	(0.0085)
Upper bound of 95% CI	0.070	0.034
N fields classified (with agreement)	925	842
Rate of agreement across reviewers	0.824	0.863

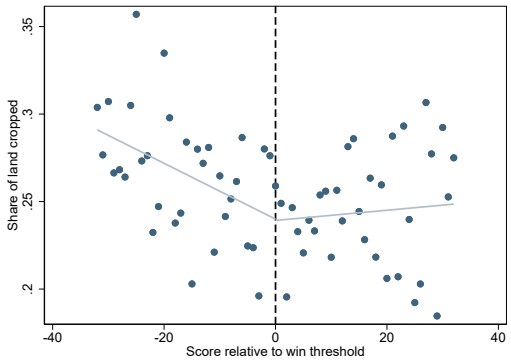
Notes: Table presents results from an exercise classifying aerial photographs of contracted fields as cropped or non-cropped among two reviewers, who also reviewed images of non-CRP fields and were blind to the distinction. Classification focuses on the 2016 auction. Column (1) includes photographs from 2017-2021. Column (2) includes only photographs from 2019-2021. Crop classifications are based on only fields in which the two reviewers agree (which occurred for 82-86% of fields). Fields more likely to be flagged as non-compliant (based on remote sensing data) were over-sampled, to be as conservative as possible.

This exercise only studies compliance on the base action, land retirement, not the top-up actions, which we cannot observe. We thus use this assessment of compliance to make an inference about the overall compliance regime across all actions.

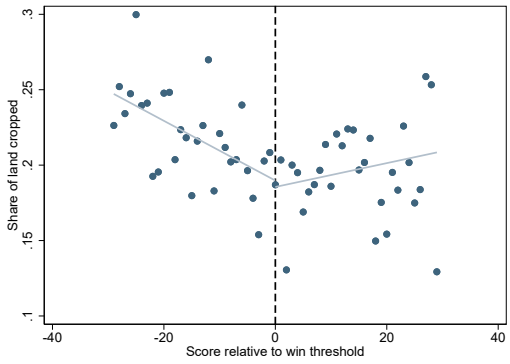
1.10.3 Supplemental Figures and Tables

Figure A.4: Spillovers: Cropping Effects on Non-Bid Fields

(a) Remote sensing



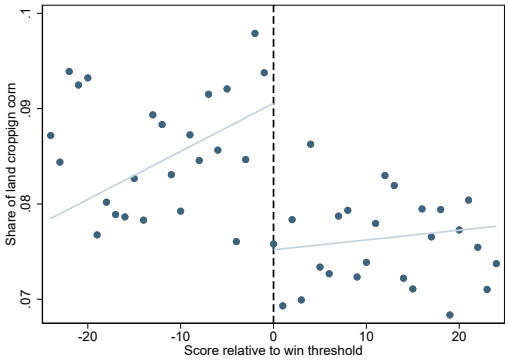
(b) Admin



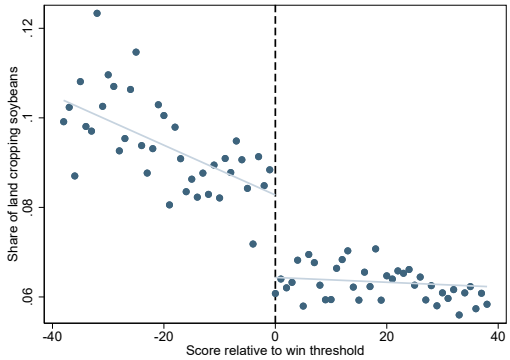
Notes: Panels (a) and (b) present raw data and estimates from equation (1.7) for $r(i,t) > 0$ (post-auction). Regression is estimated at the field level, restricting to non-bid fields for bidding landowners. Estimates are restricted to the 2016 auction where delineations of bid and non-bid fields are observed. Land-use outcomes are measured as the share of the bidding land that is cropped using the remote sensing data (a) and administrative data (b). The running variable is the difference between each bidder's score and the threshold score. Positive numbers on the x-axis correspond to winning scores, negative numbers correspond to losing scores. Each observation is a bidder-year. Corresponding coefficient estimates and standard errors presented in Table 1.2.

Figure A.5: Additional RD Plots: Remote-Sensing Data

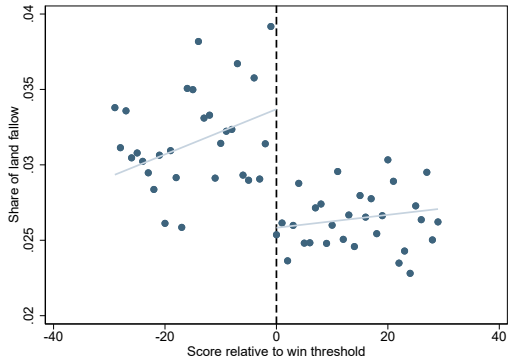
(a) Cropping corn



(b) Copping soybeans

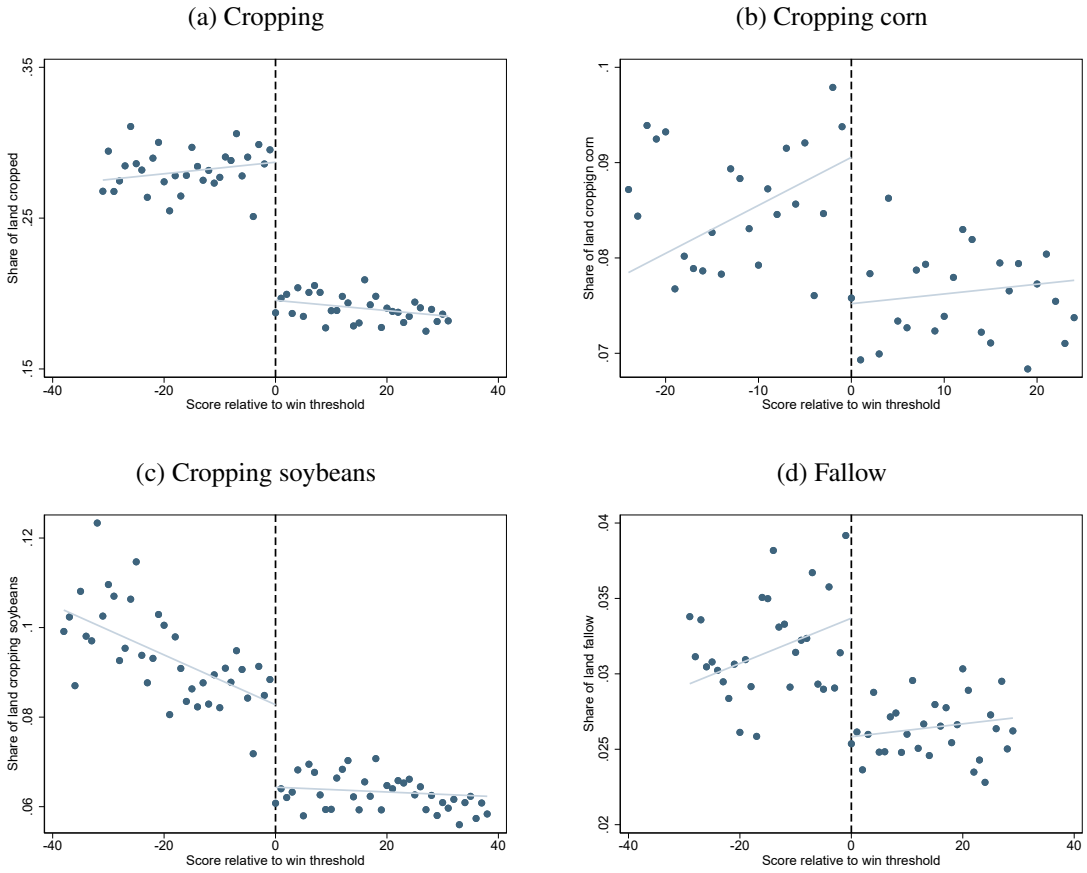


(c) Fallow



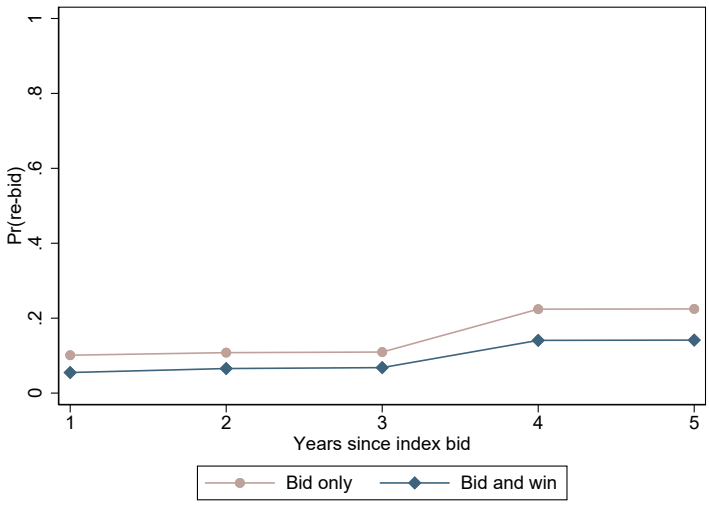
Notes: Figure presents raw data and estimates from equation (1.7) for $r(i,t) > 0$ (post-auction). Land-use outcomes are measured using crop classifications in the remote sensing data. The running variable is the difference between each bidder's score and the threshold score. Positive numbers on the x-axis correspond to winning scores, negative numbers correspond to losing scores. Each observation is a bidder-year. Corresponding coefficient estimates and standard errors presented in Table 1.2.

Figure A.6: Additional RD Plots: Admin Data



Notes: Figure presents raw data and estimates from equation (1.7) for $r(i,t) > 0$ (post-auction). Land-use outcomes are measured using crop classifications in the Form 578 data reported to the USDA. The running variable is the difference between each bidder's score and the threshold score. Positive numbers on the x-axis correspond to winning scores, negative numbers correspond to losing scores. Each observation is a bidder-year. Corresponding coefficient estimates and standard errors presented in Table 1.2.

Figure A.7: Rebidding Hazard



Notes: Figure plots the share of losing bidders who have rebid at least once in the years following an index auction, split by all bidders (beige) and successful bidders (blue).

Table A.5: RD Estimates: By Win Threshold of Bid Rental Rate for Base Contract

	Remote-sensing	Admin
	(1)	(2)
Quartile 1 threshold bid (lowest)	-0.039	-0.054
	(0.013)	(0.013)
Quartile 2 threshold bid	-0.059	-0.068
	(0.012)	(0.012)
Quartile 3 threshold bid	-0.031	-0.042
	(0.012)	(0.013)
Quartile 4 threshold bid (highest)	-0.075	-0.098
	(0.015)	(0.015)

Notes: Table presents pooled RD coefficients (Equation (1.7) for $r(i,t) > 0$ (post-auction) split by the bid rental rate required to achieve the threshold score with the base contract. This parameterizes heterogeneity in the location of the discontinuity across auctions and variation within auctions across bidders (based on $\mathbf{z}_i^s$). The outcome is the share of a bidder's land that is cropped, measured in the remotely sensed data. Standard errors clustered at the tract level.

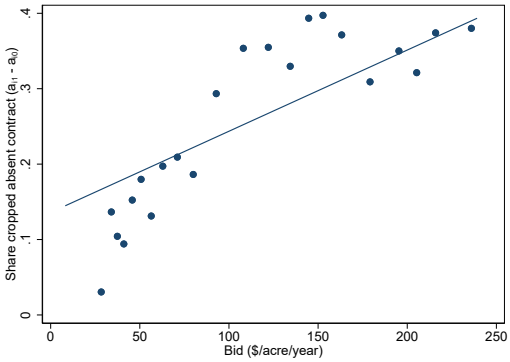
Table A.6: RD Coefficient Estimates | Bid $\geq$ Five Acres of Land

	Remote sensing	Admin
	(1)	(2)
Panel A: Main outcome: share of land cropped		
Pre-sign-up (placebo)	0.016	0.014
	(0.007)	(0.006)
Post-period (pooled sign-ups)	-0.076	-0.095
	(0.007)	(0.006)
Post-period (full contract duration: 2010-2020)	-0.117	
	(0.020)	
Panel B: Other outcomes		
Corn	-0.016	-0.023
	(0.003)	(0.003)
Soybean	-0.021	-0.027
	(0.003)	(0.003)
Fallow	-0.009	-0.011
	(0.002)	(0.001)
Natural vegetation or grassland	0.097	
	(0.007)	
Panel C: Spillovers to non-offered fields		
Share of non-offered fields cropped	-0.001	-0.000
	(0.015)	(0.015)
N bidders	236,593	236,593
N bidder-years	2,839,116	1,656,151

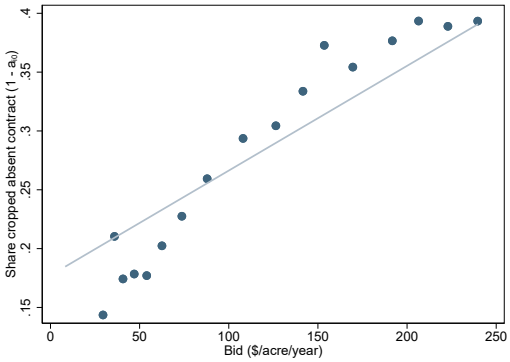
Notes: Table presents coefficient estimates from equation (1.7) with land use outcomes measured in the remotely sensed (column (1)) and administrative (column (2)) data, restricted to bidders who bid more then five acres into the mechanism (following Lark et al. (2017)). All results use a local linear regression on either side of the win threshold in the MSE-optimal bandwidth (Calonico et al., 2014). The full-contract duration specification restricts to the 2009 auction, others pool all auctions with post-period data (2009, 2011, 2012, 2013, and 2016). The pooled post-period includes an average of 7-8 post-auction years. Natural vegetation or grassland is only observed in remotely sensed data. Calculations of implied additionality divide the treatment effect estimates by the amount of land contracting among winning bidders in the MSE-optimal bandwidth. Panel C estimates the effect of a CRP contract on non-bid, and therefore non-contracting, fields to test for spillovers. This analysis is restricted to the 2016 auction. Standard errors are clustered at the bidder level.

Figure A.8: Testing for Asymmetric Information, Admin Data

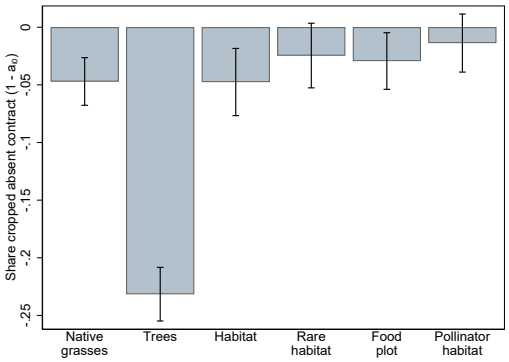
(a) Additionality vs. bids



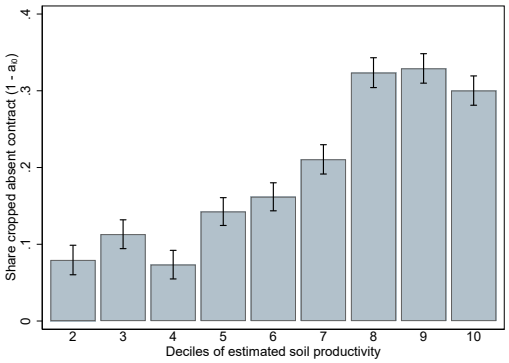
(b) Additionality vs. bids | observables



(c) Additionality across contracts



(d) Observable predictors of additionality



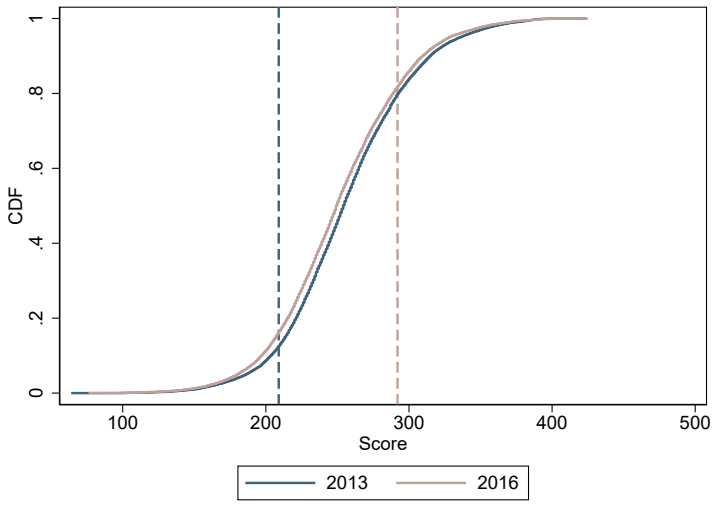
Notes: Figures present visual representations of estimates of equation (1.8). All regressions control for landowner characteristics in the scoring rule: whether a bidder is in a wildlife priority zone, estimates of groundwater quality, estimates of surface water quality, estimates of wind and water erosion (deciles), air quality impacts, and whether or not a bidder is in an air quality zone. The outcome variable in all panels is a landowner-specific measure of additionality ($1 - a_{i0}$). This is calculated as the share of all fields bid into the CRP mechanism that are cropped post auction for rejected landowners. The sample is restricted to the 2016 auction, in which 82% of bidders are rejected and the delineations of bid fields are observed. Cropping on bid fields is measured in 2017-2020 in the administrative data. Panel (a) is a binned scatterplot correlating the dollar bid (per acre, per year) with additionality, conditional on characteristics included in the scoring rule. Panel (b) adds controls for interaction terms of prior land use (quartiles of share of land cropped prior to bidding and re-enrolling CRP status) and deciles of estimated soil productivity. Panel (c) plots relative additionality by the chosen contract in the bid, relative to an omitted category of introduced grasses. Panel (d) plots relative additionality by deciles of estimated soil productivity. Standard errors clustered at the bidder level.

1.10.4 Model and Estimation Details

Information

Quantity uncertainty Figure A.9 provides empirical support for the uncertainty in quantity cleared based on the acreage limit of the auction (determined by the Farm Bill). The 2013 and 2016 auctions had very different quantity thresholds, and thus very different threshold scores — denoted by the dashed lines in blue and beige — but the cumulative distribution functions (CDFs) of bidder scores are similar.

Figure A.9: CDF of Scores versus Winning Thresholds: 2013 versus 2016



Notes: Figure presents ex-post win thresholds and cumulative distribution functions (CDFs) of ex-ante score distributions for the 2013 and 2016 auctions.

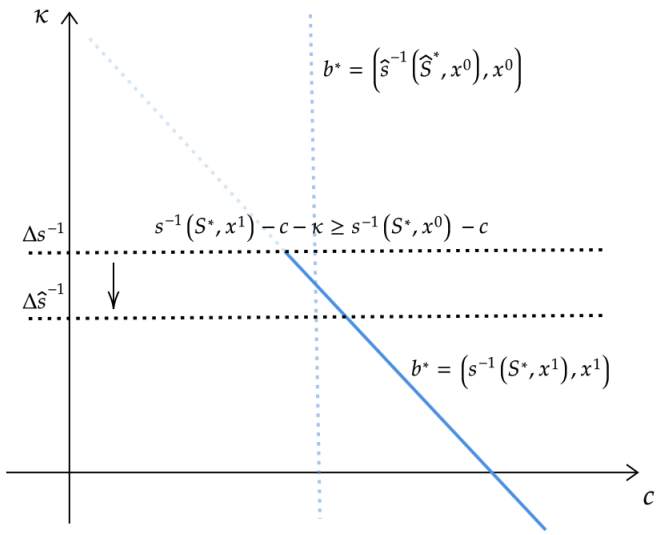
Identification Figure A.10 presents a graphical identification argument in the simple case with only two contract choices (one normalized to have $\kappa = 0$) and a quasi-linear scoring rule. $s^{-1}(S^*, x)$ describes the payment a bidder can receive to achieve score S^* with action x (see Table A.3 for an illustration of this function).

The choice to bid S^* and x^1 under scoring rule s defines the blue line segment containing the true parameters c and κ. S^* can be inverted as in Guerre et al. (2000a) to point identify $c + \kappa$ (the line containing the blue line segment in Figure A.10). The observation that x^1 was chosen (not x^0) to reach S^*, given the different payoffs associated with x^0 and x^1 in the scoring rule, bounds the magnitude of κ, defining the line segment. If κ were higher than the horizontal line in Figure A.10, it would have been optimal to reach score S^* with x^0 instead of x^1.

Variation in the scoring rule that shifts the payoffs to x^1 versus x^0, i.e. from s to $\hat{s}$, traces out the density of bidder costs as bidders' optimal choices change in response to the variation in the scoring rule. For example, the vertical dashed line documents a bidder who changes her optimal bid to x^0 with $\hat{S}^*$ under the new rule.

This argument extends to non-linearities in the scoring rule, a larger menu of contracts, and the fact that scores can only be integers. See (Agarwal et al., 2023) for more details.

Figure A.10: Graphical Identification Argument



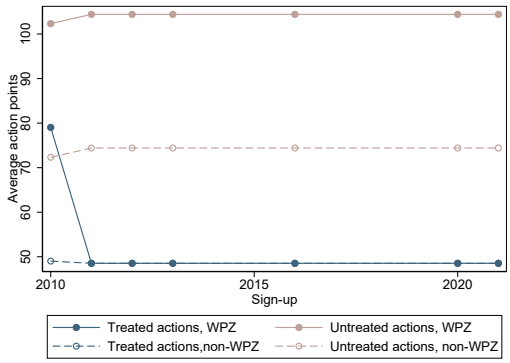
Notes: Figure presents a graphical identification argument.

As discussed in the main text, the final component of the model, $\tau(\mathbf{z}_i, c_i, \kappa_i)$, is identified by also observing a_{i0} jointly with optimal bids (including as they change with the variation from s to $\hat{s}$ described in Figure A.10).

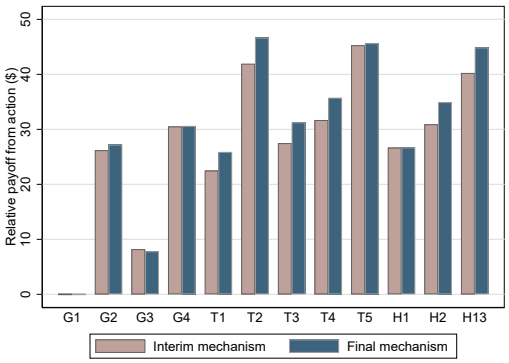
Figure A.10 clarifies the need for variation in the scoring rule to trace out the distribution of c and κ. Figure A.11 describes this variation in our context.

Figure A.11: Sources of Variation in the Scoring Rule

(a) Wildlife Priority Zone Variation



(b) Mid-Mechanism Policy Change

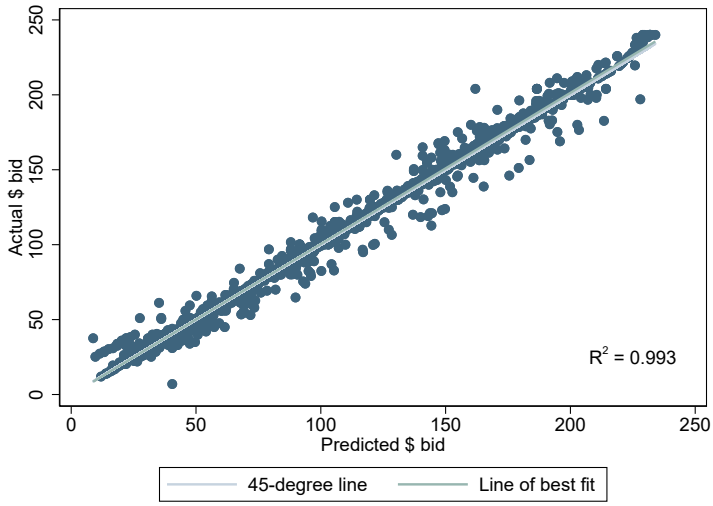


Notes: Figure presents sources of policy variation in the scoring rule that yield variation in payments across contracts differentiated by top-up conservation actions. Panel (a) plots average action points awarded for a set of "treated" actions, actions for which after the 2011 auction WPZ bidders no longer got WPZ points, and "untreated" actions, whose points remained the same, and the same average action points for non-WPZ bidders. Panel (b) plots the average rental rate that would be received for a target score (illustrated using the threshold score) among bidders under the interim mechanism before the introduction of Climate Smart Practice Incentives, and in the final mechanism after their introduction, for each of the twelve primary covers. G indicates grasses, T indicates trees, H indicates habitats.

Estimation

Step 0: Constructing the Scoring Rule We only observe scores for chosen bids $\mathbf{b}_i$, so we construct the function $s(\mathbf{b}_i, \mathbf{z}_i^s)$ from the EBI Factsheets. Figure A.12 confirms that our reconstruction performs well: at observed actions, our scoring-rule-implied required bid rental rate to achieve the score chosen in the data predicts the observed bid rental rate with an R^2 of over 0.99.

Figure A.12: True versus Predicted Bid Rental Rate at Observed Scores and Contracts



Notes: Figure presents a scatter plot of true versus predicted bid rental rates at observed contract and score choices to validate the construction of $s(\mathbf{b}_i, \mathbf{z}_i^s)$.

Step 1: Obtain Bidder Beliefs via Simulation Our resampling procedure to simulate the probability of winning at any score, $G(S)$ follows Hortaçsu (2000); Hortaçsu and McAdams (2010). Specifically, we:

1. Fit a Beta distribution to the observed distribution of acreage thresholds across auctions. For this step, we use additional historic data on auctions starting in 1999. This provides us with 12 auctions.

2. Fit a Beta distribution to the observed distribution of number of opposing bidders across auctions. For this step, we again use additional historic data on auctions starting in 1999. This provides us with 12 auctions.

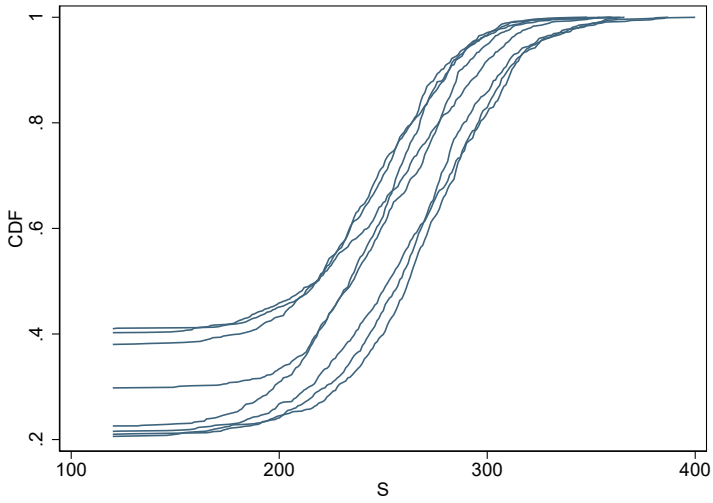
3. Draw an acreage threshold from the distribution fit in Step 1 and the number of opposing bidders, N, in Step 2. Then, for each auction g, sample with replacement N bidders from the

empirical distribution of bidders in that auction. Given the joint distribution of scores and acreage amounts among the N resampled bidders, and the drawn acreage threshold, find the winning score threshold S.

4. Repeat Step 3 to obtain an auction specific probability of winning at any given score $G_g(S)$.

Bidders form expectations about the distribution of competing scores without knowledge of their competitors' identities or characteristics, consistent with the large and decentralized bidding process, so $G_g(S)$ is the same for all bidders. Figure A.13 plots the output of our simulation procedure across all auctions.

Figure A.13: Probability of Winning at Score S



Notes: Figure presents CDFs of the simulated distribution of win probabilities at a given score across auctions.

Step 2: Estimate $F_{c,\kappa|\mathbf{z}}$ Our estimation procedure is as follows:

1. **Construct a proposal distribution.** We begin by constructing a proposal distribution from which to draw proposal (c_i, κ_i) draws. We obtain our proposal distribution by estimating a simplified version of the model. Bidders choose a score using only their *expectations* of their κ_{ij} draws, then given that score, choose an optimal contract. In this model, estimation of κ_{ij} and c_i can be separated into a discrete choice problem and an inversion. We obtain parameter estimates from this simplified model, then set our proposal distribution to be independent normals with the estimated means and variances of this simplified model (inflating the variance by 25%).

2. **Draw from proposal and solve the bidder's problem.** Following the approach of Ackerberg (2009), we use a change of variables to draw simulations of $\left(c_i^k, \kappa_i^k\right)$ from the proposal distribution and solve the bidder's problem in equation (1.10) for each bidder and each simulation draw. Bidders can only bid integer scores, so to solve equation (1.10), we search over all feasible score-contract combinations among integers in the support of observed scores. This change of variables allows us to solve the bidder's problem only $N \times K$ times, once for each bidder and each simulation draw, instead of $N \times K \times R$ times, for each evaluation of the objective function (R times).

3. **Coarsen choice probabilities.** Because the number of possible bids is large (on the order of 10,000 choices), we face the challenge that the probability of simulating each bid observed in the data is low. We address this challenge by coarsening the bidder's solution obtained in Step 2. We coarsen to the cartesian product of (i) deciles of the scoring rule and (ii) the five dimensions of p_j when u_j is the no upgrade option, plus the two upgrade options. Let $\widetilde{\mathbf{b}}_i^* = \left(\tilde{S}_i, \tilde{\mathbf{x}}_i\right)$ denote the optimal coarsened bid observed in the data.

4. **Reweight simulation draws.** We can then construct the importance sampling estimator by re-weighting simulation draws. The likelihood of observing the coarsened choice in the data, $\widetilde{\mathbf{b}}_i^* = \left(\tilde{S}_i, \tilde{\mathbf{x}}_i\right)$, given a parameter guess θ, is:

$$\mathscr{L}_i = \frac{1}{K} \sum \mathbb{1}\left(\widetilde{\mathbf{b}}_i^* = \widetilde{\mathbf{b}}_i^{*k} | \left(c_i^k, \kappa_i^k\right)\right) \frac{p\left(\left(c_i^k, \kappa_i^k\right) | \theta\right)}{g\left(\left(c_i^k, \kappa_i^k\right)\right)}, \tag{1.17}$$

   where $\widetilde{\mathbf{b}}_i^{*k}$ is the coarsened optimal bid given simulation draw $\left(c_i^k, \kappa_i^k\right)$, the solution to the bidder's problem in equation (1.10), and the coarsening described in Step 3. Equation (1.17) then re-weights simulation draws by $\frac{p\left(\left(c_i^k, \kappa_i^k\right) | \theta\right)}{g\left(\left(c_i^k, \kappa_i^k\right)\right)}$, where $p\left(\left(c_i^k, \kappa_i^k\right) | \theta\right)$ is the probability of observing simulation draw $\left(c_i^k, \kappa_i^k\right)$ given parameter guess θ, and $g\left(\left(c_i^k, \kappa_i^k\right)\right)$ the probability of observing $\left(c_i^k, \kappa_i^k\right)$ given the proposal distribution from Step 1.

5. **Find θ to maximize the log likelihood**. We suppressed dependence in (1.17) on $\mathbf{z}_i$. We estimate θ separately for each of the 32 cells of observable heterogeneity for a sample of 1,000 bidders in each cell in each auction (due to computational constraints on the USDA servers). An auxiliary benefit of the importance sampling approach of Ackerberg (2009) is that it yields a differentiable objective function.

6. **Repeat**. We repeat Steps 2-5 several times, using estimates from the solution to Step 5 as the new proposal distribution. Our final estimates use 10,000 simulation draws to mitigate simulation bias (Train, 2009).

Step 3: Estimate $\tau(\mathbf{z}_i, c_i, \kappa_i)$ Our final step involves estimating the conditional expectation function $\tau(\mathbf{z}_i, c_i, \kappa_i) = \mathbb{E}[1 - a_{i0} \,|\, \mathbf{z}_i, c_i, \kappa_i] = \pi \cdot \mathbf{z}_i + \beta \cdot c_i + \alpha \cdot \kappa_i$. We match model implied moments of additionality to observed moments of additionality, $1 - a_{i0}$, among bidders who lose the auction. We search for $\theta^\tau = (\pi, \beta, \alpha)$ that minimizes $\hat{g}(\theta^\tau)' A \hat{g}(\theta^\tau)$ for weight matrix A and $\hat{g}(\theta^\tau) = \hat{\mathbb{E}}\left[m_i - \frac{1}{K}\sum_k m_i\left(\theta^\tau | c_i^k, \kappa_{\mathbf{i}}^{\mathbf{k}}\right)\right]$, where $\hat{\mathbb{E}}$ denotes the sample expectation, for m_i equal to:

- Additionality at the award threshold: $(1 - a_{i0}) \cdot \mathbb{1}\left[\underline{S} - b < s(\mathbf{b}_i^*, \mathbf{z}_i^s) < \underline{S}\right]$ for bandwidth b.
- Additionality by observable characteristics: $(1 - a_{i0}) \cdot \mathbb{1}\left[s(\mathbf{b}_i^*, \mathbf{z}_i^s) < \underline{S}\right] \cdot \mathbf{z}_i$.
- Covariance between additionality and chosen scores: $(1 - a_{i0}) \cdot s(\mathbf{b}_i^*, \mathbf{z}_i^s) \cdot \mathbb{1}\left[s(\mathbf{b}_i^*, \mathbf{z}_i^s) < \underline{S}\right]$.
- Additionality within chosen contracts: $(1 - a_{i0}) \cdot \mathbb{1}\left[x_{ij} = 1\right] \cdot \mathbb{1}\left[s(\mathbf{b}_i^*, \mathbf{z}_i^s) < \underline{S}\right]$.

Our estimation approach follows the following steps:

1. Draws simulations $\left(c_i^k, \kappa_i^k\right)$ from $F_{c,\kappa|\mathbf{z}}$ estimated in Step 2.
2. Calculate optimal bids $\mathbf{b}_i^*$ given $\left(c_i^k, \kappa_i^k\right)$ using equation (1.10).
3. Calculate $m_i\left(\theta^\tau | c_i^k, \kappa_i^k\right)$ by replacing $1 - a_{i0}$ with $\pi \cdot \mathbf{z}_i + \beta \cdot c_i + \alpha \cdot \kappa_i$ and observed bids with simulated optimal bids for each simulation draw k and parameter guess θ^τ.
4. Minimize the objective $\hat{g}(\theta^\tau)' A \hat{g}(\theta^\tau)$.

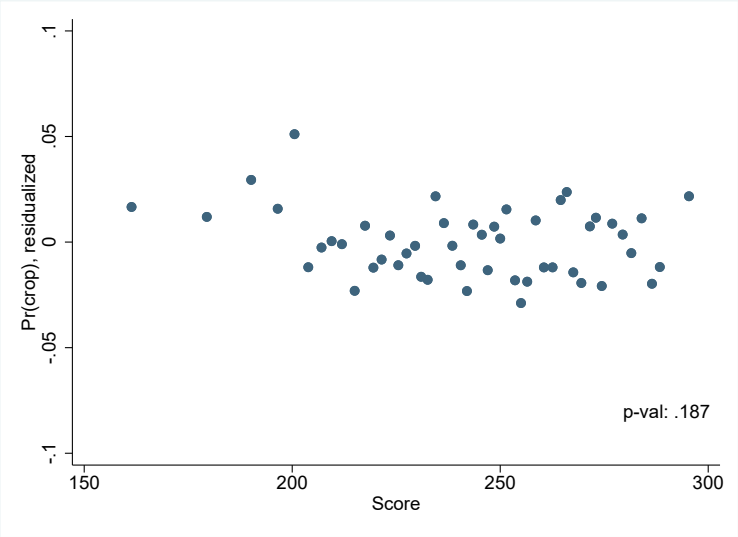
We use the two-step optimal weight matrix for the matrix A.

Because we require an observation of bid fields to calculate $1 - a_{i0}$, we estimate $\tau(\mathbf{z}_i, c_i, \kappa_i)$ using only the auction where we observe bid fields (2016). Our primary estimates use the remote-sensing data from 2017-2020 to measure $1 - a_{i0}$. We assume that the relationships estimated in $\tau(\mathbf{z}_i, c_i, \kappa_i)$ in this auction can be extrapolated to the other auctions in our sample, and that $\tau(\mathbf{z}_i, c_i, \kappa_i)$ can be estimated in only the three years following the auction. This may seem unappealing given the transition period in Figure 1.4, but we note that $1 - a_{i0}$ is calculated among losing bidders, not those transitioning into land retirement.

As discussed in the main text, we require instruments that shift $s(\mathbf{b}_i^*, \mathbf{z}_i^s)$ but that are conditionally independent of a_{i0}. We use landowners' Wildlife Priority Zone and Air Quality Zone status as instruments. We conduct a test to provide additional support for this assumption. Specifically, we estimate the simplified version of the model described in Step 1 of our Step 2 estimator, in which we can point identify c_i with an inversion. We show in Figure A.14 that cropping outcomes are

independent of the score after controlling for c_i and the remaining observables in $\tau(\mathbf{z}_i, c_i, \kappa_i)$. This suggests that the residual variation in the score is conditionally independent of a_{i0}.

Figure A.14: Residualized Correlation Between Scores and Cropping



Notes: Figure presents the relationship between a binary indictor for cropping, residualized of observable characteristics, a point-identified c_i estimate from an alternative model, and scoring rule characteristics except for Wildlife Priority Zone and Air Quality Zone. Estimated among losing bidders in the 2016 auction only.

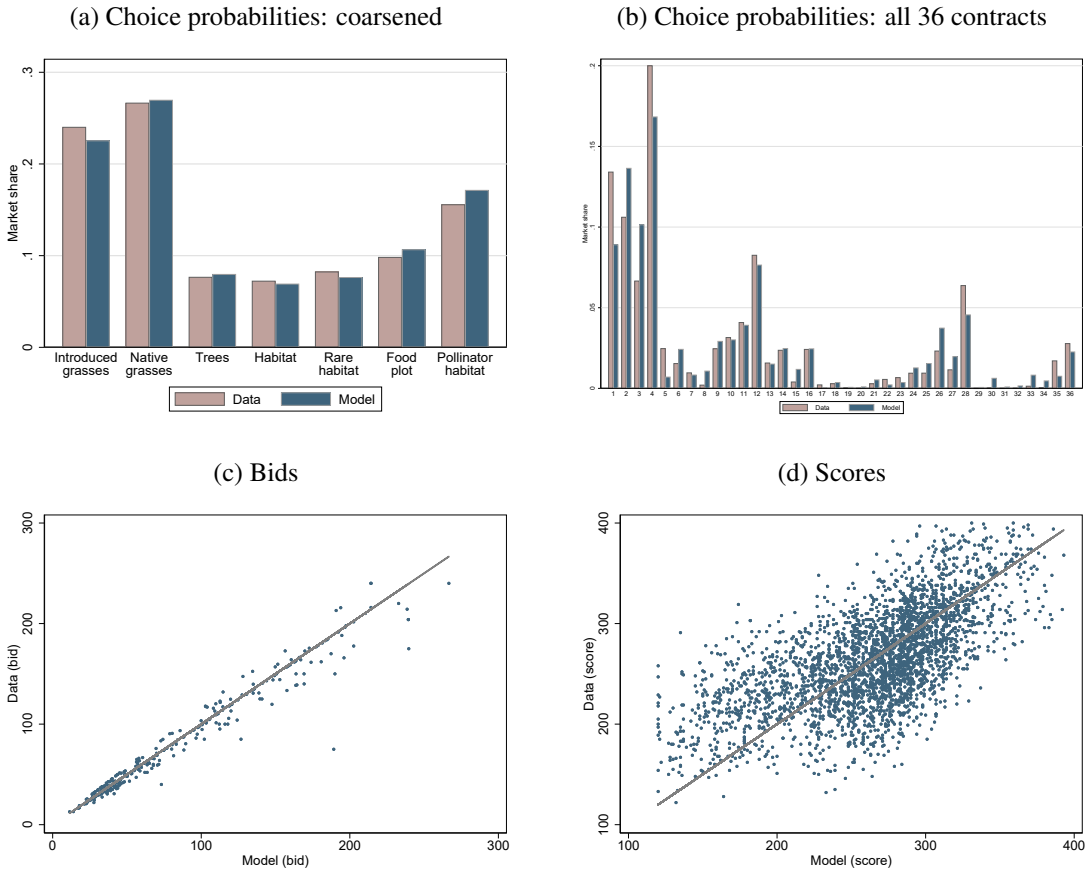
We calculate standard errors via bootstrapping. Our final procedure will bootstrap over the entire estimation procedure to incorporate estimation error in earlier steps. The current standard errors do not incorporate estimation error in (c_i, κ_i).

Table A.7: $F_{c,\kappa|\mathbf{z}}$ Parameter Estimates (Select $\mathbf{z}_i$)

	Former CRP = 0				Former CRP = 1			
Prior crop	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Soil prod.	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
c_i								
Mean	31.65	37.51	66.35	135.97	36.55	42.90	66.85	126.27
	(0.02)	(0.04)	(0.05)	(0.06)	(0.02)	(0.03)	(0.05)	(0.06)
Log σ_c	1.60	2.77	3.53	3.61	1.16	2.60	3.47	3.89
	(0.004)	(0.002)	(0.001)	(0.001)	(0.004)	(0.002)	(0.001)	(0.001)
κ_{ij}								
Means								
Native grasses	0.70	-4.46	3.96	-0.61	-2.59	-4.76	4.87	2.31
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Trees	28.57	27.53	30.11	34.51	15.85	19.25	23.77	30.67
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Habitat	17.25	12.71	16.82	22.01	13.96	11.76	15.94	12.92
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Rare habitat	17.73	15.63	17.97	10.05	18.30	12.96	22.57	19.60
	(0.04)	(0.04)	(0.03)	(0.03)	(0.05)	(0.04)	(0.04)	(0.03)
Wildlife food plot	23.77	23.07	12.69	14.24	25.66	18.65	14.45	16.24
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
Pollinator habitat	16.72	10.81	14.12	18.00	22.04	18.68	18.27	17.40
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
Log σ_κ	2.70	2.86	2.81	2.76	2.85	2.84	2.83	2.81
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)

Notes: Table presents presents parameter estimates for 8 cells of $\mathbf{z}_i$. Standard errors calculated using the inverse of the negative Hessian, calculated numerically. Standard errors do not account for simulation error or the estimation error in the first step estimator of $G(S)$.

Figure A.15: Model Fit



Notes: Figures summarize model fit by comparing simulated choices of contracts, bids, and scores to the data.

Table A.8: Comparison Between Estimated and Administrative Cost Estimates

	Estimates	Median admin cost	Average admin cost
	(1)	(2)	(3)
Tree primary covers (rel. to grasses)	24.36	26.46	73.15
Habitat primary covers (rel. to grasses)	15.05	2.67	3.30

Notes: Table presents average revealed preference estimates of costs of aggregate primary cover categories, relative to grasses (column 1), compared to administrative data collected on the costs of these actions by the USDA (columns 2 and 3).

1.10.5 Valuing Benefits

We assume that the weights in the scoring rule $B_j(\mathbf{z}_i^s)$ reflect the relative social benefits (in dollars) across j and $\mathbf{z}_i^s$, assuming $a_{i0} = 0$ for all i. The mechanism implicitly makes trade-offs in the scoring rule that monetize relative preferences across contracts and characteristics.

Using this logic requires two assumptions. First, we require the assumption that $a_{i0} = 0$ for all i, motivated by Claassen et al. (2018), who write: *Benefit-cost indices are used to rank applications for acceptance in all major USDA conservation programs... Existing indices, however, implicitly assume full additionality*. Second, we require that the weights in the scoring rule are not distorted to reduce expenditures (Che, 1993). There is no evidence to support this behavior (Ribaudo et al., 2001), and moreover, the USDA values transfers to agricultural landowners. We assume that the USDA maximizes social welfare by announcing its preferences in the scoring rule.

However, the USDA revealed-preferred values of $B_j(\mathbf{z}_i^s)$ may not necessarily align with the true environmental benefits for a variety of reasons, e.g. political concerns (Ribaudo et al., 2001). We choose to take this USDA-revealed-preferred approach, versus calibrating $B_j(\mathbf{z}_i^s)$ from an external integrated assessments model,[53] to focus on additionality as the primary source of social welfare losses.

To calculate these scoring-rule implied relative valuations, we note that scoring rule is separable in the actions incentivized by the heterogeneous contracts and the bid ($) rental rate

$$s(\mathbf{b}_i, \mathbf{z}_i^s) = \underbrace{s_a(\mathbf{x}_i, \mathbf{z}_i^s)}_{\text{action points}} + \underbrace{s_r(r_i)}_{\text{bid rental rate points}} \tag{1.18}$$

and construct a quasi-linear approximation to the scoring rule to obtain relative willingness to pay. The scoring rule departs from quasi-linearity because of kinked incentives in points bidders receive as a percentage of their bidcap. We "quasi-linearize" the scoring rule by taking the average of $s_r'(r_i)$ in the region without the added percentage points bonus and the region with the percentage point bonus (at the median bidcap value).[54]

Using our "quasi-linearized" approximation to the scoring rule, we know how the USDA trades off higher costs with heterogeneous environmental benefits across contracts j and observable characteristics $\mathbf{z}_i^s$. However, the scoring rule is not directly informative about the level of social benefits. We obtain this using estimates of the value of the CRP from the literature. We assume that all impacts of the CRP accrue only over the contract period.

We use following values of the CRP from the literature. Our baseline estimates take the average across these three studies.

[53] See `https://naturalcapitalproject.stanford.edu/software/invest` for an example.

[54] The fact that different bidders face different scoring rules based on their bidcap does not reflect differential valuation of environmental benefits across bidders.

1. Our first estimate sums the recreational,[55] public works,[56] and air quality benefits[57] from Feather et al. (1999) and adds estimates of the value of greenhouse gas reductions from sequestered CO_2 (over the 10-year contract) and reduced fuel and fertilizer use[58] monetized at $43 per metric ton. This leads to an overall estimated value of the CRP of $98.34 per acre, per year. This is likely to be an under-estimate because biodiversity is only valued insofar as it provides recreational benefits, and this estimate does not include water quality improvements from reduced run-off.

2. Our second estimate takes the valuation of the CRP from Hansen (2007a), and adds estimates of the value of greenhouse gas reductions from sequestered CO_2 (over the 10-year contract) and reduced fuel and fertilizer use, which then equals $255.70, per acre, per year.

3. Our third and fourth estimate take a conservative and generous value of the non-GHG CRP benefits from Johnson et al. (2016a) and adds estimates of the value of greenhouse gas reductions from sequestered CO_2 (over the 10-year contract) and reduced fuel and fertilizer use. This leads to estimates of $367.96 and $456.04, per acre, per year. These may be an over-estimate because benefits are estimated in only one geographic area, which may have more environmentally sensitive land.

The description above highlights the difficulties of monetizing the value of the all of the environmental benefits of the CRP, both in terms of quantifying all of the potential environmental benefits. We emphasize that our focus is not on obtaining estimates of $B_j(\mathbf{z}_i^s)$, but rather on $\tau(\mathbf{z}_i, c_i, \kappa_i)$; results can be recalculated for any alternative valuation of $B_j(\mathbf{z}_i^s)$. Quantifying the environmental value of ecosystem services is an important complementary area of research.

1.10.6 Additional Counterfactuals

Cost of Public Funds Figure A.16 considers the same auctions presented in Figure 1.9, but evaluates social welfare with a cost of funds $\lambda = 0.15$. Under this framework, 15% of all USDA spending is considered deadweight loss, motivated by the social costs of financing expenditures via distortionary taxation. With a cost of funds, the status quo auction reduces social welfare. How-

[55] Includes sport-fishing, small-game hunting, noncompetitive viewing, and waterfowl hunting.

[56] Includes cost savings associated with reduced maintenance of roadside ditches, navigation channels, water treatment facilities, municipal water uses, flood damage, and water storage.

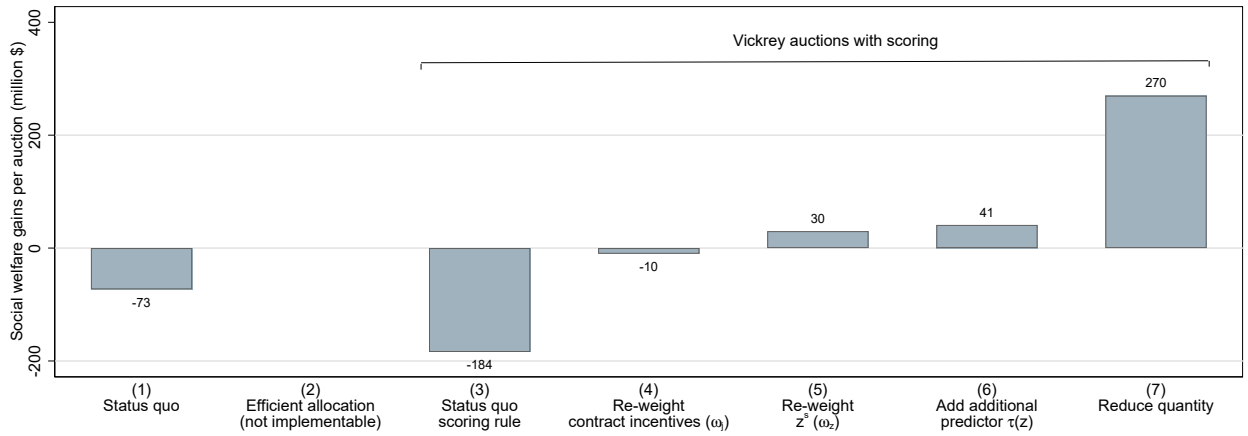
[57] Includes reduced health risks and cleaning costs associated with blowing dust.

[58] See https://www.fsa.usda.gov/Assets/USDA-FSA-Public/usdafiles/EPAS/natural-resouces-analysis/nra-landing-index/2017-files/Environmental_Benefits_of_the_US_CRP_2017_draft.pdf.

ever, social welfare gains become positive once the auction is designed to consider additionality. In bar (7), social welfare gains are large at $270 million per auction.

Figure A.16 evaluates auctions using the same scoring rules as in Figure 1.9. With weights ω_j and $\omega_{\mathbf{z}}$ re-optimized to reduce government spending, social welfare gains would be higher.

Figure A.16: Social Welfare Under Alternative Auctions: Cost of Funds



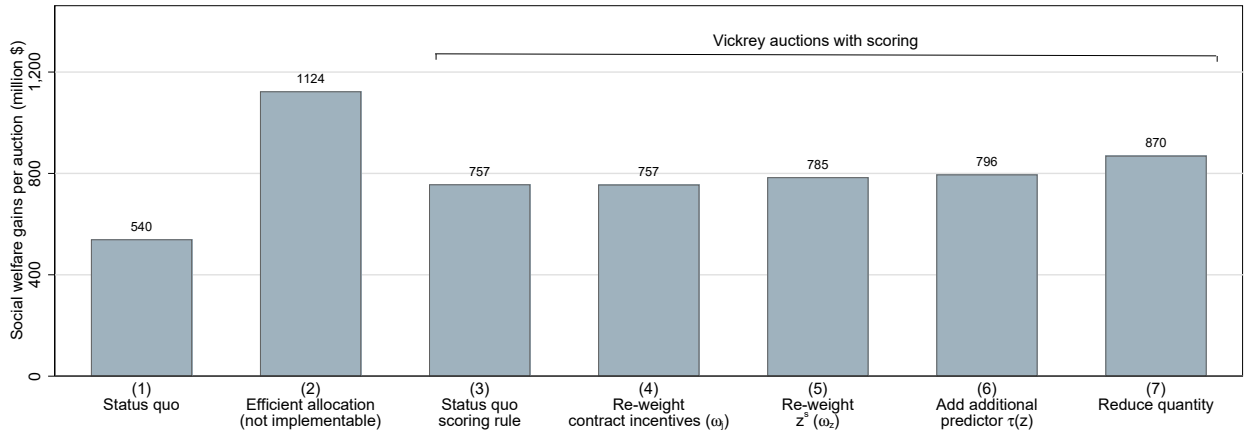
Notes: Figure presents estimates of the social welfare gains (defined in equation (1.14) with a cost of funds $\lambda = 0.15$) under status quo and alternative auctions. Results reported in million dollars per auction. All auctions impose that each landowner obtains at most one contract and that total contracts awarded cannot exceed the status quo. Bar (1) simulates the status quo. Bar (2) calculates the social welfare gains under an efficient allocation that allocates contracts using all $\mathbf{z}_i$ and (c_i, κ_i) to maximize equation (1.14). Due to adverse selection, this allocation may not be implementable. Bars (3)-(7) calculate social welfare under alternative Vickrey auctions with scoring (see Section 1.6.2 for more details). Bars (3)-(6) hold quantity (the number of landowners allocated contracts) constant at the status quo and change the scoring rule $s_j(\mathbf{z}_i)$ defined in equation (1.15). Bar (3) uses the existing scoring rule $s_j(\mathbf{z}_i) = B_j(\mathbf{z}_i^s)$. Bar (4) uses a scoring rule with the social-surplus maximizing incentives across contracts (ω_j). Bar (5) uses a scoring rule with the social-surplus maximizing asymmetry across bidders using characteristics already in the scoring rule $(\mathbf{z}_i^s)$. Bar (6) adds an additional characteristic to the scoring rule, a prediction of $\tau(\mathbf{z}_i, c_i, \kappa_i)$ based on immutable characteristics of landowners already collected by the USDA (deciles of soil productivity and wind and water erosion). Bar (7) uses the same scoring rule as bar (6) but reduces the number of contracts allocated to landowners: only landowners with positive scoring-rule-implied social surplus $\max_j s_j(\mathbf{z}_i) - c_i - \kappa_{ij} \geq 0$ are awarded contracts. See each bar's corresponding column in Table 1.5 for more details.

Top-Actions not Affected by Additionality For all analyses beyond those restricted to the base contract, we require an assumption about how additionality impacts the social value derived from the top-up actions that differentiate the contracts in the mechanism. This is due to fundamental data limitations (see Section 1.3.2). Our primary specification defines the social benefit of contracting as $B_j(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i)$. In this specification, no social benefits are generated when a landowner is not additional. This could be either because, as with land retirement, top-up actions (or close substitutes) would have occurred even absent a CRP contract. It could also be motivated by an assumption that land retirement and the top-up actions are complements in the USDA's valuation of contracting. We view this assumption as reasonable for many of the important actions being incentivized, e.g. grassland and tree maintenance.

In this section, we consider an alternative assumption in which the incremental actions incentivized

by the contracts are always additional. Specifically, we consider an alternative valuation of contracts equal to $B_0(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i) + B^j(\mathbf{z}_i^s)$, where $B_0(\mathbf{z}_i^s)$ is the social benefit of the base action and $B^j(\mathbf{z}_i^s)$ is the incremental social benefit of the top-up action beyond the base action. This could be motivated by a scenario in which contracting impacted the specific species mix, which we assume the USDA values at $B^j(\mathbf{z}_i^s)$, even if it did not impact land retirement. Under this assumption, over one third of the total social surplus at stake is not impacted by additionality at all. This makes correctly incentivizing the top-up actions — whose relative valuations are derived solely from monetizing the scoring rule — matter substantially to the performance of the mechanism. This is another reason to favor our baseline assumption over this alternative.

Figure A.17: Social Welfare Under Alternative Auctions: Alternative Top-Up Assumption



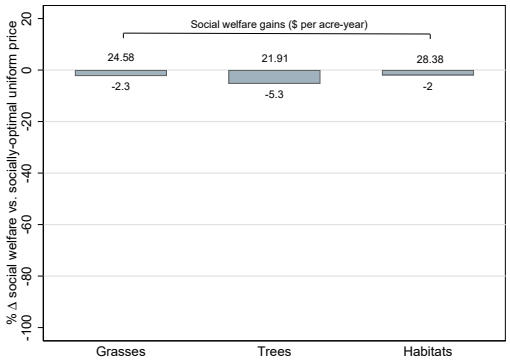
Notes: Figure presents estimates of the social welfare gains (defined in equation (1.14), but replacing $B_j(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i)$ with $B_0(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i) + B^j(\mathbf{z}_i^s)$ where $B_0(\mathbf{z}_i^s)$ is the social benefit of the base action and $B^j(\mathbf{z}_i^s)$ is the incremental value of the top-up action) under status quo and alternative auctions. Results reported in million dollars per auction. All auctions impose that each landowner obtains at most one contract and that total contracts awarded cannot exceed the status quo. Bar (1) simulates the status quo. Bar (2) calculates the social welfare gains under an efficient allocation that allocates contracts using all $\mathbf{z_i}$ and (c_i, κ_i) to maximize equation (1.14). Due to adverse selection, this allocation may not be implementable. Bars (3)-(7) calculate social welfare under alternative Vickrey auctions with scoring (see Section 1.6.2 for more details). Bars (3)-(6) hold quantity (the number of landowners allocated contracts) constant at the status quo and change the scoring rule $s_j(\mathbf{z}_i)$ defined in equation (1.15). Bar (3) uses the existing scoring rule $s_j(\mathbf{z}_i) = B_j(\mathbf{z}_i^s)$. Bar (4) uses a scoring rule with the social-surplus maximizing incentives across contracts $(\boldsymbol{\omega}_j)$. Bar (5) uses a scoring rule with the social-surplus maximizing asymmetry across bidders using characteristics already in the scoring rule $(\mathbf{z}_i^s)$. Bar (6) adds an additional characteristic to the scoring rule, a prediction of $\tau(\mathbf{z}_i, c_i, \kappa_i)$ based on immutable characteristics of landowners already collected by the USDA (deciles of soil productivity and wind and water erosion). Bar (7) uses the same scoring rule as bar (6) but reduces the number of contracts allocated to landowners: only landowners with positive scoring-rule-implied social surplus $\max_j s_j(\mathbf{z}_i) - c_i - \kappa_{ij} \geq 0$ are awarded contracts.

Figure A.17 re-creates Figure 1.9 under this alternative assumption. Social welfare under the status quo is higher, but the status quo still achieves less than half of the social welfare gains under the efficient allocation. The biggest difference between Figures 1.9 and A.17 is the large social welfare gains from the switch to the VCG mechanism, holding the scoring rule constant (bar (3)). This is because the Vickrey auction with scoring using the status quo rule now efficiently incentivizes choices across contracts. Because the top-up actions represent a large share of social welfare gains at stake in the mechanism, incentivizing them efficiently is important for the auction's performance.

Adjusting the mechanism for heterogeneity in additionality is still quantitatively important: moving from bar (3) to bar (6) increases social welfare by 7% of the status quo, or 15% of the gains of the improvement between bar (1) and bar (6). Also as in our baseline estimates, quantity procured is higher than is socially optimal. Reducing quantity to reflect the many landowners who are not additional increases social welfare by a further 13% of the status quo social welfare gains.

We emphasize that the exercise of Figure A.17 is not to document that results are quantitatively the same under this alternative assumption versus our baseline assumption. The assumptions are very different, so naturally lead to some different quantitative implications. Instead, we highlight that the insights from our baseline assumption are quantitatively relevant even when a large share of the surplus at stake in the mechanism ($B^j(\mathbf{z}_i^s)$) is not impacted by additionality.

Figure A.18: Alternative Assumption: Offset Market Design Across Contracts



Notes: Figure plots the percent reduction in social welfare under a stylized competitive market equilibrium relative to socially-optimal prices under three different hypothetical markets, each with only one contract traded at a uniform price. The numbers above the bars in panels (a) and (b) tabulate total social welfare (per acre-year) in each competitive market. The competitive market equilibrium is calculated under the assumption that buyers have the same full-information preferences as the USDA and form expectations over the value of any contract given the equilibrium price(s). In this figure, we assume that $B_0(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i) + B^j(\mathbf{z}_i^s)$, where $B_0(\mathbf{z}_i^s)$ is the social benefit of the base action and $B^j(\mathbf{z}_i^s)$ is the incremental value of the top-up action, instead of our baseline assumption of $B_j(\mathbf{z}_i^s) \cdot \tau(\mathbf{z}_i, c_i, \kappa_i)$.

We also examine how this alternative assumption impacts our analysis of competitive offset market design. Most of the analysis in the main text is focused on the base contract (e.g. Figures 1.7 and 1.11a), so is unaffected by our assumptions about top-up actions. However, our analysis in Figure 1.11b is affected. Figure A.18 re-creates Figure 1.11b under this section's alternative assumption about top-up actions. Figure A.18 documents social welfare losses from adverse selection, but markets for tree planting and maintenance contracts no longer unravel. This occurs because the social value from $B^j(\mathbf{z}_i^s)$ is generated regardless of additionality, propping up the market. While we think that our baseline assumption likely holds, this exercise is informative of an alternative lever for market design. Offset markets can bundle additional benefits (often termed "co-benefits") into the contract that are unaffected by additionality. These benefits not only provide additional social value, but can also prevent market unravelling due to adverse selection.

Chapter 2

Waiting or Paying for Healthcare: Evidence from the Veterans Health Administration

2.1 Introduction

Many goods and services are allocated at below-market-clearing prices, sacrificing efficiency in the interest of equity (Tobin, 1970; Weitzman, 1977). This trade-off is particularly acute among healthcare systems worldwide, which were designed with explicit equity considerations that preclude rationing by price (Reinhardt, 1997; Cutler, 2002a).[1] However, in the presence of scarcity, due to budget or capacity constraints, alternative rationing mechanisms must emerge to determine access to care. In healthcare, wait times for care often serve as this substitute (Cutler, 2002a; Reinhardt, 2007, 2019) and represent a major barrier to care in most OECD countries (OECD, 2020). Many countries have thus made the choice, either explicitly or implicitly, that allocating healthcare via wait times is preferable to doing so with a price mechanism.[2]

The choice of rationing mechanism — in practice, waiting versus paying on the margin — is a

[1] For example, the Beveridge Commission, which led to the creation of the National Health Service (NHS) in the UK stated:

"From the standpoint of Social Security, a health service providing full preventive and curative treatment of every kind to every citizen without exceptions, without remuneration limits and without an economic barrier at any point to delay recourse to it, is the ideal plan."

Similar ideas were included in the Canada Health Act of 1984:

"It is hereby declared that the primary objective of Canadian health care policy is to protect, promote, and restore the physical and mental well-being of residents of Canda and to facilitate reasonable access to health services without financial or other barriers."

[2] This choice is particularly explicit in a recent (2022) case before the Supreme Court of British Columbia in Canada, in which the court ruled that even though the "public system had failed to provide timely medical treatment" paying for healthcare to avoid a queue is not in "accordance with principles of fundamental justice." See https://www.bccourts.ca/jdb-txt/ca/22/02/2022BCCA0245.htm for more details.

fundamental question of international healthcare market design. However, empirical research on the implications of alternative rationing mechanisms has been limited by both scarce data and the challenge of conducting welfare analysis in settings where consumers are not paying on the margin.

In this paper, I investigate the efficiency and distributional implications of rationing access to care via wait times versus prices. I also examine the performance of second-best policy instruments that attempt to manage rationing costs while still imposing strong restrictions on the use of price to ration access to care. I focus on the setting of waiting for outpatient care at the Veterans Health Administration (VA) in the United States. The VA combines unusually rich data on wait times with a major policy intervention designed to address rationing costs, which induced variation in both wait times and prices.

The VA offers an ideal empirical setting for four reasons. First, in addition to being of substantial policy interest in the US, providing care to over nine million veteran enrollees, the VA context is similar to the international comparisons that motivate the debate over rationing mechanisms in healthcare. Copayments are regulated at low — often zero — rates, and the market for outpatient care instead clears on wait times. Second, the Veterans Access, Choice, and Accountability Act of 2014 (the "Choice Act"), which made certain veterans eligible to receive subsidies for non-VA providers, provides a large-scale policy change that shifts prices, and in equilibrium, wait times. In addition to providing rare variation in both the efficient (prices) and inefficient (wait times) market-clearing mechanisms, this class of policy intervention is a common form of "managed rationing," with similar interventions across the globe.[3] Third, VA administrative data is derived from electronic health records that integrate with scheduling systems to document wait times for care. Fourth, I am able to assemble a comprehensive dataset linking utilization across VA and non-VA providers to analyze choices — and thus, conduct revealed preference welfare analysis in a rationed goods context — in response to the Choice Act variation.

I begin by leveraging the eligibility conditions for the Choice Act in two sets of of difference-in-difference designs. First, I examine the direct effects of the policy, which made veterans eligible to obtain care at non-VA, or "community" providers, at (the lower) VA copayment rates, if they lived more than forty miles from their closest VA clinic, lived in a state without a VA hospital, or needed to wait over thirty days for care. I focus on the forty-mile threshold and document, consistent with prior work (Rose et al., 2021; Saruya et al., 2023), that veterans increase community outpatient utilization and overall outpatient spending in response to the policy. Approximately 50% of the increase in community utilization is driven by substitution from VA care, and 50% is an increase in overall utilization. This analysis highlights that (1) veterans are sensitive to prices,

[3] See OECD (2020) for Portugal and Denmark, Propper et al. (2008) for the UK, Ringard et al. (2016) for Norway, and the website of the agency that manages such a program in Chile.

as choices change in response to the increased subsidy for community care, (2) the policy achieved its intended effect of increasing access to care among the directly eligible, but at a cost of increased spending, and (3) substitution away from the VA could, in theory, lead to positive externalities on other veterans via reduced wait times.

I next examine whether this policy-induced substitution away from the VA achieved the policy's second goal of alleviating capacity constraints on congested clinics and reducing wait times in equilibrium. I test for this using clinic- and market-level exposure designs, with exposure determined by the share of pre-period patients who would have been eligible for the Choice Act subsidies. I document evidence of equilibrium effects on wait times: moving from the 10th to the 90th percentile of clinic exposure reduces wait times by between 5 and 13 days. These equilibrium effects are critical for policy analysis, as all veterans benefit from the policy via reductions in wait time, regardless of eligibility. Moreover, the equilibrium effects on wait times allow me to use the policy to provide exogenous variation in both of the rationing mechanisms of interest: within a market, changes in eligibility isolate changes in price, while conditional on any given consumer's eligibility status, the cross-product share of *others* who are eligible isolates variation in wait times.

I use these two sources of variation to describe the screening properties of the two rationing mechanisms. This is a key descriptive exercise as the welfare and allocative effects of the two regimes depend on heterogeneity in willingness-to-pay and willingness-to-wait. I first document evidence consistent with the equity motive to relinquish a price mechanism: lower income and sicker veterans are substantially more responsive to the Choice Act-induced price change than higher income veterans. This presents the tension for a planner designing a healthcare system with preferences for redistributing to lower income, sicker veterans. But while the choice to relinquish a price mechanism is intentional, its substitute — wait times — arises endogenously instead of through careful design (Cutler, 2002a), with unknown screening properties. I document qualitatively similar patterns of screening along the two rationing mechanisms: lower income and sicker veterans are also more likely to be screened out at higher wait times. These similar qualitative patterns of screening place a bound on the extent to which rationing via wait time can be a useful redistributive tool for a planner, but are consistent with both a meaningful equity-efficiency trade-off and a scenario in which status quo rationing regimes are adverse to both efficiency and distributional goals.

To quantify the allocative effects, efficiency costs, and any redistribution of surplus across the two rationing mechanisms, I move beyond descriptive analyses to estimate the joint distribution of willingness-to-wait and willingness-to-pay. In the second half of the paper, I develop and estimate a model of clinic choice and queuing for primary care providers in order to obtain this joint distribution. In the model, wait times arise endogenously in response to veteran preferences and capacity constraints. Veterans make decisions over if and where to receive care, across VA and

community options, trading off observed and unobserved clinic characteristics, wait times, and prices. I assume VA clinics are capacity constrained with wait times generated via a First-Come-First-Served queuing protocol, while community providers are uncapacity constrained. I use the Choice Act policy variation to account for endogeneity in both wait times and prices with similar empirical strategies and under similar assumptions as in the reduced-form policy analysis.

I estimate that veterans are responsive to both prices and wait times, with wait time elasticities approximately three times as large as price elasticities. The average veteran has a cost of delay — the amount a veteran would be willing to pay to move an appointment one day earlier — of approximately $2.50 per day with substantial dispersion around this average ($0.93 at the 10th percentile to $3.64 at the 90th percentile). This is a key parameter of interest that could not be obtained from descriptive analyses alone: it governs both the magnitude of the efficiency costs of waiting and any redistribution of surplus. I document that, despite the qualitatively similar patterns of screening, higher income, healthier veterans have the highest costs of delay (in dollars). Put differently, wait times implicitly discriminate in favor of lower socio-economic status veterans.

I use the estimated preferences to examine counterfactual allocations under alternative rationing regimes. I focus on comparing the polar cases of the status quo rationed regime, where wait times clear the market at very low prices, versus a counterfactual in which prices adjust flexibly to achieve zero wait times, holding total VA capacity fixed. Under the counterfactual price-based regime, veterans would be required to pay an average of $78 (approximately 20% of costs) per primary care visit, versus less than $5 under the status quo. The binding price controls in the status quo have a substantial impact on allocations: 16% of veterans who would have used VA care under the rationed regime would not under the price regime, and these veterans are substantially lower income, sicker, and older.

I use a revealed preference approach to quantify the efficiency and distributional effects of these descriptive patterns. I define an efficient benchmark, subject to available capacity, as a regime that maximizes money-metric surplus. Relative to this benchmark, rationing imposes substantial efficiency costs equivalent to 24% of achievable surplus, due to both deadweight loss from "money burnt" in costly screening via waiting (70%) and allocative distortions away from the highest willingness-to-pay consumers (30%). However, despite the wait-based regime's large costs on average, over 50% of veterans, who are lower income, sicker, and older than the overall population, prefer status quo rationing by waiting.

My estimates allow me to qualitatively evaluate the trade-off between efficiency and redistribution facing a planner. Though a planner's redistributive preferences are inherently unknown, I show that rationing is an extremely inefficient form of redistribution, destroying over $5 of surplus for every $1 gain for the winners under the status quo. This is inefficient both relative to the tax schedule

(Hendren, 2020), and in absolute terms: switching to a price mechanism can achieve close to a Pareto improvement even with the blunt instrument of uniform transfers. This finding is driven by the basic descriptive patterns of screening: because the two instruments screen on qualitatively similar dimensions, the rationing regime imposes large deadweight losses from the costly screen of wait times for any limited socially desirable redistribution of surplus.

Finally, I compare the gains from changes in the allocation mechanism to "managed rationing" policies observed in practice, including the Choice Act. Not only does the Choice Act not approach the gains of a change in rationing mechanism (reducing wait times by increasing prices), it reduces overall welfare, as consumer surplus gains do not outweigh the increase in costs. This occurs because (1) veterans are not willing to pay the full cost of care, and (2) the policy is poorly targeted, as the policy provides a uniform subsidy for veterans with heterogeneous externalities depending on their substitution patterns (Diamond, 1973). By contrast, a small, feasible targeted copayment increase at the VA, an improvement on both dimensions, dominates the Choice Act. It raises consumer surplus — concentrated among the lower income veterans who are not targeted by the copay increase — despite increasing prices, and generates revenue, instead of increasing costs. Though this policy falls far short of the gains from eliminating price controls altogether, its performance underscores the primary conclusion of this paper: relaxing price controls can lead to large welfare gains, even when considering the distributional concerns that motivate them.

Related Literature This paper contributes an empirical application to two related theory literatures in market design and public finance. In market design, I draw on both (1) theoretical work investigating the efficiency (Bulow and Klemperer, 2012; Che et al., 2013) and redistributive (Weitzman, 1977; Dworczak et al., 2021a; Akbarpour et al., 2022) implications of price controls and (2), a distinct literature on money-burning mechanisms generally (Hartline and Roughgarden, 2008; Condorelli, 2012; Yang, 2022; Dworczak, 2022) and equilibria in wait times, specifically (Leshno, 2022). Second, I contribute to the classic public finance theory on the use of ordeals to achieve targeting or redistributive goals (Nichols et al., 1971; Nichols and Zeckhauser, 1982; Besley and Coate, 1992) and benchmark the performance of the ordeal against a price mechanism with transfers (Zeckhauser, 2021).

My empirical analysis, which focuses on allocative effects of alternative rationing mechanisms, combines and builds on both a large, recent literature examining the screening properties of ordeals (Alatas et al., 2016; Dupas et al., 2016; Deshpande and Li, 2019; Finkelstein and Notowidigdo, 2019; Brot-Goldberg et al., 2023) and a more limited literature analyzing the allocative effects of price controls (Glaeser and Luttmer, 2003; Davis and Kilian, 2011; Ryan and Sudarshan, 2022). In my setting, ordeals arise endogenously to capacity constraints, and I quantify the allocative ef-

fects, redistributive properties, and deadweight loss jointly via revealed preference, as in Waldinger (2021) and Lieber and Lockwood (2019). Relative to these papers, I investigate a distinct but central trade-off: the use of an efficient versus an inefficient rationing mechanism to target a transfer.

My preference specification also connects to a literature on ride-hail that estimates preferences over efficient (prices) and inefficient (wait times) market clearing mechanisms jointly (Buchholz et al., 2020; Castillo, 2022; Fréchette et al., 2019). Similar to these papers, the key welfare object of interest is a willingness-to-pay to reduce wait times, which can only be obtained by observing and quantifying delay and dollar trade-offs.

Most specifically, my focus on wait times at the VA and the Choice Act policy contribute to literatures on demand responses to wait times for healthcare (Besley et al., 1999; Martin and Smith, 1999; Pizer and Prentice, 2011a,b; Yee et al., 2022b), quality of care at the VA (Chan et al., 2022a), the Choice Act specifically (Rose et al., 2021; Saruya et al., 2023), and the impact of subsidies for private care on public sector hospitals, more generally (Propper et al., 2008; Cooper et al., 2018).

2.2 Conceptual Framework

In this section, I present a stylized equilibrium framework to illustrate the welfare consequences of alternative rationing mechanisms in healthcare. The framework also illustrates the core empirical questions and challenges that will motivate the structure of my empirical analysis.

Demand There exists a unit mass of consumers, indexed by i, each obtaining utility from the separable consumption of healthcare and all other goods (c). Utility from healthcare is given by $h_i - \gamma_i w$, where h_i indicates an individual's value for healthcare services today, and γ_i captures i's costs of waiting w days for a visit. γ_i incorporates the myriad reasons why a consumer may dislike waiting for care: physical discomfort, anxiety, or a reduced ability to work, further reductions in health capital, or a preference to be seen immediately. In addition to healthcare, consumers value the consumption of all other goods with an increasing and concave function $u(c)$. I normalize the price of all other consumption to one.

Given price p and wait time w for care, consumer i solves:

$$\max_{x\in\{0,1\},c} (h_i - \gamma_i w)\cdot x + u(c) \quad \text{s.t. } x\cdot p + c = y_i \tag{2.1}$$

where $x \in \{0,1\}$ denotes whether to forgo or consume care, and y_i is i's income. Define $x_i(p,w)$ as i's (unit) demand for care and $v_i(x_i(p,w),p,w)$ as i's indirect utility at any (p,w) combination.

Total demand in the market is given by $D(p,w) = \int x_i(p,w) f(i) di$. For each individual i, define the dollar-denominated cost of waiting $C_i(w)$ as the price at which i is indifferent between waiting w and paying $C_i(w)$ for $x = 1$.[4]

Supply and Equilibrium As in my empirical application, I assume care is provided by a strictly capacity- (or budget-) constrained social planner, with supply $\kappa < 1$. Waiting times w are generated anonymously via a First-Come-First-Served queuing mechanism. Also consistent with my empirical application, I assume that κ is sufficiently low such that $D(0,0) > \kappa$. In equilibrium $D(p,w) = \kappa$, with a continuum of (p,w) combinations each determining an equilibrium. I focus attention on the polar case used in practice, $(0,w^*)$, and a flexible pricing benchmark, $(p^*,0)$.

Social welfare Fairness and redistribution are often the motivation for price controls and the subsequent use of an alternative mechanism to ration scarce capacity. To capture this, I define social welfare at any (p,w) equilibrium:

$$SW(p,w) = \int g_i v_i(x_i(p,w),p,w) f(i) di + g^G p \kappa \tag{2.2}$$

where g_i denote welfare weights, or the value that a social planner assigns to the welfare of individual i. The term g^G modulates the value of any revenue collected. I take capacity κ as given (consistent with the assumption of strict capacity constraints) and focus on the allocation problem facing a social planner.

In the spirit of Saez and Stantcheva (2016), g_i could theoretically encompass any societal concern for fairness that motivates relinquishing a price mechanism. In my empirical application, I will focus in particular on the distributional implications of the choice of rationing mechanism along income and health status.

The change in welfare between the $(0,w^*)$ and the $(p^*,0)$ regimes can be decomposed as follows:

$$SW(0,w^*) - SW(p^*,0) = \int g_i \left(\underbrace{v_i(x_i(0,w^*),0,w^*) - v_i(x_i(p^*,0),0,w^*)}_{\text{change in allocation}} + \underbrace{v_i(x_i(p^*,0),0,w^*) - v_i(x_i(p^*,0),p^*,0)}_{\text{change in payoff | allocation}} \right) f(i) di - \underbrace{g^G p^* \kappa}_{\text{revenue}} \tag{2.3}$$

[4] Specifically, $C_i(w)$ is implicitly defined by $v_i(x_i(C_i(w),0),C_i(w),0) = v_i(x_i(0,w),0,w)$.

The first term highlights that, because the two allocation mechanisms yield different demand curves, changing the mechanism will shift the set of individuals who obtain versus forego care. The second arises because, even conditional on an allocation, the different demand curves will lead to differences in consumer surplus. The final term illustrates that a price mechanism generates transfers between consumers and the government (who collects revenue), whereas waiting is pure social waste. With a complete set of transfers available, allocating care to the highest willingness-to-pay individuals maximizes social surplus, regardless of g_i (Kaldor, 1939; Hicks, 1939). This allocation, which is achieved by the price mechanism, is the efficient benchmark.

Figure 2.1 illustrates both the efficiency and distributional consequences of the two regimes by plotting (example) demand curves and cost curves induced by the different mechanisms. The line AC represents the demand (WTP) curve. By contrast, the line AB plots the demand curve induced by a waiting time mechanism in money-metric (WTP) space: it plots the *willingness-to-pay* among all consumers with *willingness-to-wait* above the market-clearing waiting time, w^*. The line AB must lie (weakly) below the line AC, and the area between these curves represents the allocative efficiency loss from rationing by waiting. Second, while the price mechanism merely involves splitting surplus between consumers (the area above p^*) and the government (the area below p^*), the waiting time mechanism induces a cost $C(w^*)$ for those who wait. The area underneath this curve is pure deadweight loss.

Figure 2.1 also illustrates the distributional consequences of the two instruments. Example consumer j obtains care, and positive surplus, under the waiting time mechanism where he would not have under the price mechanism. Example consumer i obtains care under both regimes, but obtains higher consumer surplus under the waiting regime versus the price regime. Others lose: the entire region between the curve AC and the curve AB, represent consumers who are displaced from care under the waiting time regime.

Although the highest money-metric surplus achievable occurs with a price mechanism, the alternative rationing regime redistributes consumption and surplus across consumers, potentially toward high g_i types. Thus, a planner (potentially) faces a trade-off between maximizing overall (money-metric) efficiency and redistributing surplus to high g_i types (a notion of equity). The existence and slope of this trade-off depend on the magnitude of efficiency losses and the extent of redistribution across types. This in turn depends on the joint distribution of willingness-to-pay, willingness-to-wait, and characteristics that may influence social welfare weights g_i, such as income or health status. This joint distribution will be my key empirical object of interest.

Estimating this joint distribution is typically hindered by three challenges. First, almost by definition, consumers are not paying on the margin for goods when they are rationed without prices. Second, money-burning activities — in this case, waiting — are infrequently recorded in standard

datasets, making it difficult to quantify the surplus dissipated to arrive at a given allocation. Finally, as wait times are determined in equilibrium, even if observed, they are subject to the same simultaneity concerns that present the core challenges to demand estimation in standard markets. In the next section, I discuss the VA setting and data, with particular emphasis on the unique features of the setting that allow me to overcome these three challenges.

2.3 Setting and Data

2.3.1 The Veterans Health Administration

The Veterans Health Administration (VHA, or VA) of the Department of Veterans Affairs is the largest integrated healthcare system in the United States. The system serves over 9 million veterans with a budget of over $80 billion (Department of Veterans Affairs, 2023). The VA has historically provided the vast majority of inpatient and outpatient care at 170 VA hospitals and over 1,000 VA community-based outpatient clinics across the United States. However, in recent years, the VA has also financed an increasing amount of care for VA enrollees at non-VA, or "community," providers. I focus specifically on outpatient care, with a particular emphasis on primary care, as these are settings where wait times and access to care are of particular concern at the VA (Yee et al., 2022a,c; 113th Congress, 2014; 115th Congress, 2018) and other settings (Mark, 2023).

Eligible veterans[5] pay no premiums for access to VA care, but may be obligated to pay copayments. Three quarters of veterans pay no copayments at all for outpatient care. Approximately 25% of veterans pay $15 and $50 for primary and specialty outpatient care, respectively. Whether or not a veteran pays copayments depends on their assigned priority group, which depends on a veteran's service-connected disabilities, service history, and income. Many veterans use the VA in combination with other sources of coverage. Approximately half of veterans are dually enrolled in Medicare.

Due to the combination of budget and capacity constraints and regulated copayments, the market for outpatient care at the VA clears on wait times to access care (Yee et al., 2022a). Wait times, and their impact on veterans' access to care, has been the subject of concern and policy activity at the VA for over two decades (U. S. Government Accountability Office, 2001, 2012, 2016, 2019; 113th Congress, 2014; 115th Congress, 2018).

[5] Veterans are eligible to enroll in the VHA if they served in the active military, naval, or air service and were not dishonorably discharged.

The Veterans Choice and Accountability Act of 2014 I focus on the policy context of the Veterans Access, Choice, and Accountability Act of 2014 (the "Choice Act"). Motivated by concerns that veterans were unable to obtain care in a timely manner, the Choice Act dramatically expanded subsidized access to outpatient care at non-VA, or "community," providers. Eligible veterans could obtain care at community providers paying VA copayment rates (zero, or $15-$50, depending on priority group). These providers were then paid Medicare rates by the VA. Veterans were eligible if they lived over 40 miles from their nearest clinic, lived in a state without a VA hospital (Alaska, Hawaii, or New Hampshire), or needed to wait over thirty days to obtain an outpatient appointment at a nearby clinic.

A key goal of the Choice Act was to alleviate capacity constraints at the VA via two channels. The first channel was the direct expansion in access for eligible veterans, who experienced a reduction in the out-of-pocket price for community providers. The second channel was the equilibrium effect of reduced wait times for *everybody* as eligible veterans substituted away from the VA.

In theory, the Choice Act policy therefore provides both variation in prices and wait times, exactly the type of variation necessary to estimate the joint distribution of willingness-to-wait and willingness-to-pay, the key objects in the framework in Section 2.2. Beyond its instrumental use, the Choice Act provides a rich laboratory to analyze second-best policy. At the VA, this policy prompted a a major shift in the delivery of care: in 2018, eligibility was expanded under the MISSION Act, and today, approximately 20% of the VA budget is dedicated to financing non-VA care (Department of Veterans Affairs, 2023). Beyond the VA, the broad class of policies offering (targeted) subsidies for private utilization is a common second-best policy tool to alleviate congestion among public sector healthcare providers.

2.3.2 Data

I assemble a comprehensive dataset describing VA enrollees and their VA and non-VA utilization. Throughout the paper, I will refer to utilization at non-VA providers as community utilization (the VA terminology). My primary analysis sample includes all enrollees from fiscal years 2011-2017 from the ADUSH (Assistant Deputy Under Secretary for Health) file. This file includes demographics such as age, priority group, income measured via VA means tests, date of death, and summary measures of utilization, as well as veterans' exact residential address and whether it grants them distance-eligibility under Choice for all VA-enrolled veterans.

I measure VA utilization, as well as additional enrollee-level characteristics, using all data recorded in electronic health records (EHR) at clinical encounters at all VA hospitals and clinics.[6] A crucial

[6]This data is recorded in the VA's Corporate Data Warehouse (CDW).

advantage of this data is that the EHR integrates with the appointment scheduling system, allowing me to measure wait times. Specifically, I measure wait times as the number of days between the date of an appointment request and the appointment itself (Yee et al., 2022a; Chartock, 2023). These data are unavailable in commonly used claims datasets, which only record the date of service and not the date of request.

Despite the existence of records documenting the number of days between an appointment request and the appointment itself, constructing the appropriate measure of wait times presents several challenges. First, wait times are only recorded if an individual makes an appointment. I construct the menu of wait times facing each patient based on average clinic wait times facing all patients who obtained an appointment at that clinic within a given time period. Second, this aggregation presents its own challenge, as the distribution of observed wait times may be selected. I address this with a supply-side queuing model, discussed in more detail in Section 2.5.4. A final threat to measurement occurs if appointments are scheduled far in advance of when the appointment is actually desired. In my sample period, scheduling appointments over 90 days in advance was prohibited, reducing dramatically the extent of follow-up appointments that erroneously appear as long wait times. I also follow previous work (Yee et al., 2022a; Pizer and Prentice, 2011b) and assess the robustness of my results to wait times constructed using only new patients.

I supplement the data on VA utilization with multiple sources describing community utilization. First, I use all authorizations (internal documentation) and claims (payments to providers) for care financed by the VA at community providers. Second, for the 36% of veterans who are dually enrolled in Traditional Medicare (TM), I link the universe of Medicare claims to VA utilization. For these veterans, I am able to characterize the complete extent of healthcare consumption across VA and community providers.

While less crucial than the datasets above, I also use estimates of average VA costs produced by the Health Economics Resource Center (HERC).[7] Together with the claims data, these cost estimates allow me to characterize spending across VA and community care.

Summary statistics (enrollees) Table 2.1 presents summary statistics of enrollee characteristics and utilization for the entire enrollee population (column (1)) the population that is dually eligible for Traditional Medicare (TM), where I observe the universe of community utilization (column (2)), and all veterans, split by whether veterans are distance-eligible for Choice (column (3)) versus ineligible (column (4)). I compare the TM sample to the overall sample because the TMs offer a unique advantage in data completeness that will make them a useful subset in a number of

[7]These data calculate encounter-level VA utilization by using Medicare relative value weights to distribute aggregate clinic-by-category VA costs.

analyses. I include the split by distance eligibility because comparing the utilization patterns of these two groups will play a key role in my empirical strategy, described in more detail in Section 2.4.

Table 2.1 documents that the VA population is atypical in some dimensions and more representative in others. The VA population skews male (over 90%) and old. The average enrollee income, which is measured using means tests conducted throughout a veteran's enrollment at the VA, is lower than the U.S. median over this time period (approximately $56,000 as of 2015 for the US a whole versus $30,204 for VA enrollees), with substantial dispersion in income in the population. The variation in socioeconomic status within the VA population makes it a relevant setting to study efficiency and distributional trade-offs in healthcare design. The majority of veterans pay no copayments for outpatient care.

Many veterans have other sources of coverage. This presents both a challenge and an opportunity. The challenge is that, for some veterans with supplemental coverage, I do not observe their universe of healthcare consumption. The opportunity is that for a substantial subsample — those dually enrolled in Traditional Medicare (column (2)) — I can observe veterans making choices both within the VA and across VA and community care as prices and wait times vary.

Spending per veteran at the VA is $5,148, with the average veteran obtaining 7.5 outpatient visits and one primary care visit annually at the VA. These averages mask substantial heterogeneity: 61.5% of enrolled veterans in a given year do not engage with the VA at all. Distance-eligible veterans engage with the VA less than their non-distance-eligible counterparts, but still a substantial amount, as veterans are willing to travel long distances for care.

In the TM sample (column (2)), VA utilization is slightly higher than the overall population, while total utilization, incorporating both VA and community care, is substantially higher. This in part reflects differences in data completeness: outside of the TM sample, I only observe VA-financed community utilization. High levels of utilization, particularly in the complete TM sample, reflect the fact that the VA population has high rates of disability and therefore high medical needs. Many veterans use both VA and non-VA care: 24% of all TM dual-eligibles saw both a VA and a non-VA primary care provider in the same year, and 50% saw both a VA and non-VA primary care provider at one point in the sample period.

Summary statistics (clinics) Figure 2.2 presents histograms to summarize key clinic-level variables for the 1,128 clinics in my sample: wait times and Choice exposure, for all specialties and for primary care. I focus on primary care specifically because these wait time measures focus on a more uniform "product" and have been used and validated in prior work at the VA (Yee et al., 2022a). Primary care care will thus also be the focus of this paper's equilibrium analysis.

Figures 2.2a and 2.2b present histograms of clinic-level wait times, calculated as the average wait time among requested appointments in a given clinic and quarter, for all specialties and for primary care.[8] Most wait times are between two to six weeks, with a right tail.

Figures 2.2c and 2.2d plot a measure of the distribution of each clinic's "exposure" to the Choice Act policy, or the extent to which the policy could impact equilibrium congestion, for all specialties and for primary care. Specifically, Figures 2.2c and 2.2d plot the distribution of the share of visits that satisfied the Choice Act eligibility requirements at each clinic for each specialty preceding the Act's enactment. Figures 2.2c and 2.2d documents dispersion in exposure, with some clinics substantially "exposed" to the Choice policy.[9]

2.4 Choice Act Policy Analysis: Shifts in p and w

2.4.1 Empirical Strategies: Estimating Direct (p) and Equilibrium (w) Effects

My empirical strategies exploit the Choice Act eligibility requirements in series of difference-in-difference designs.

Direct effects To analyze the direct effect of the policy, I compare eligible veterans, who experienced a reduction in price for community providers to ineligible veterans, around the policy's introduction. My main empirical strategy leverages the distance eligibility requirement in the following enrollee-level event study specification:

$$y_{it} = \sum_{\tau \neq 2014} \beta_t \cdot 1\{t = \tau\} \cdot 40Miles_i + 40Miles_i + \varepsilon_{it} \tag{2.4}$$

where $40Miles_i$ is an indicator for whether enrollee i lives more than 40 miles from her closest clinic and β_t are the event study coefficients of interest on years (t) relative to the introduction of

[8] I discuss this aggregation, the potential for selection in this measure if appointments are rejected at high waiting times, and how I address this selection issue to avoid bias in my quantitive estimates of preferences over waiting times in Section 2.5.4. For the purposes of simple, descriptive analyses of waiting times, in this section I present simple averages, noting this possibility for selection.

[9] It is worth noting that the measurement of waiting times in Figures 2.2a and 2.2b and Figures 2.2c and 2.2d do not correspond exactly. This is because Choice Act wait-time eligibility is based upon a patient's "desired date," which I do not use to characterize wait times except for measuring Choice eligibility, as it has been demonstrated to be subject to manipulation by clinics (U. S. Government Accountability Office, 2012). This choice follows prior research analyzing wait times at the VA (Yee et al., 2022b,a; Pizer and Prentice, 2011a).

Choice. The policy was enacted starting in Fiscal Year (FY) 2015, and I normalize the FY 2014 coefficient to be equal to zero. Outcomes y_{it} include a range of utilization and clinic choice-related outcomes. My main analysis focuses on veterans within a 10 mile window around the 40 mile threshold. In Appendix Table B.2, I also present results that include interaction terms for closest clinic (for which eligibility is determined) and year controls, $\theta_{j(i)t}$, to isolate the change in direct eligibility, or change in p, from market-level effects on wait times (discussed in more detail below). I cluster standard errors at the enrollee-level.

In this section, I focus on the reduced form effect of the policy, rather than instrumenting for price directly because veterans may face heterogeneous prices in the absence of the Choice policy depending on their other sources of coverage. When I want to interpret the Choice variation more precisely as a change in price of a specific magnitude to quantify enrollee price-responsiveness, I will restrict the sample to directly address this heterogeneity.

I also summarize results in a pooled difference-in-difference specification:

$$y_{it} = \beta \cdot 1\{t > 2014\} \cdot 40Miles_i + 40Miles_i + \varepsilon_{it} \tag{2.5}$$

I focus on the distance eligibility condition, as this is a permanent policy change that applies to all types of outpatient care. The wait-time eligibility condition is time varying and market-specific, based on equilibrium wait times and the category of care sought. In Appendix 2.10.1, I discuss how I use the wait time eligibility conditions in a similar difference-in-difference analysis using variation in eligibility across markets and over time.

Equilibrium effects To examine the potential for equilibrium effects of the policy on wait times, I leverage variation in the distribution of eligible veterans across clinics and markets. Specifically I use the variation in pre-Choice exposure, illustrated in Figures 2.2c and 2.2d, in a market- and clinic-level exposure event study design:

$$w_{jt} = \sum_{\tau \neq 2014} \eta_t \cdot 1\{t = \tau\} \cdot ShareEligPre_j + \phi_j + \chi_t + \varepsilon_{jt} \tag{2.6}$$

where w_{jt} denotes the wait time for a given market (HRR by specialty) or clinic j in year t, $ShareEligPre_j$ is the pre-period share of visits that would have been eligible under Choice, and ϕ_j and χ_t are market or clinic fixed effects (depending on the specification) and time fixed effects, respectively. The coefficients η_t capture the relative time paths of wait times for more exposed

versus less exposed clinics. I cluster standard errors at the market or clinic level, depending on the specification.

The intuition for Equation 2.6 is that clinics with more Choice-eligible potential consumers may face a larger decrease in demand than clinics with fewer choice-eligible potential consumers, and thus may have shorter wait times in equilibrium as a result of the policy.[10] Whether or not this is indeed the case depends on the substitution patterns of veterans in response to the policy, which I will investigate in the analysis of the directly eligible. I focus only on equilibrium effects on VA clinics. The VA population is small relative to the market as a whole, making equilibrium effects outside the VA unlikely.

As with the direct effects, I also summarize the impact of the policy's equilibrium effects on wait times in a difference-in-difference design

$$w_{jt} = \eta \cdot 1\{t > 2014\} \cdot ShareEligPre_j + \phi_j + \chi_t + \varepsilon_{jt} \tag{2.7}$$

pooling together all pre- and post-policy years.

2.4.2 Results: Direct Effects

Effects on choices, utilization, and spending Figure 2.3 plots event study coefficients from Equation 2.4 and Table 2.2 reports coefficient estimates from the pooled difference-in-difference specification in Equation 2.5. For clarity, I will continue to refer to all utilization at non-VA providers as "community" utilization (to use the VA terminology) and all utilization at VA providers as VA utilization.

Figure 2.3a plots the increase in total visits at community providers per year for eligible veterans, relative to ineligible veterans in the same market. The policy did indeed influence choices: eligible veterans, who experienced a decrease in out of pocket payments at community providers, obtain more care at community providers.

One complication with the interpretation of Figure 2.3a is that absent data on non-VA financed care, the patterns in Figure 2.3a could be consistent with either a behavioral response to prices or simply a change in financing for the same choices. Figure 2.3b rules out a simple change in financing: I plot total community utilization for the TM sample, where I observe the universe of utilization.

[10] Using cross-product or cross-market variation in treatment exposure is a common strategy to test for equilibrium effects (Crépon et al., 2013; Egger et al., 2022).

Figure 2.3a and Figure 2.3b show similar patterns. These two results demonstrate that the Choice-induced reduction in prices for community providers increased consumption of community care. After a ramp-up period, by 2017, this increase in visits represents a 3-6% increase as a share of total pre-period community utilization.

Figure 2.3c demonstrates that approximately 50% of the increase in community utilization is driven by substitution away from the VA. The remaining 50% represents an increase in overall utilization (Figure 2.3d) corresponding to an increase in total VA spending of approximately $28 per eligible enrollee per year (Figure 2.3e).

Together these results demonstrate three facts. First, veterans are responsive to prices, increasing their utilization of community providers when faced with a decrease in their price. Second, this translates to an overall increases in utilization, achieving one of the policy's goals of increasing access to care for the directly eligible, but at an increased cost to the VA. And third, some veterans substituted away from the VA, potentially exerting positive externalities on other veterans at the congested VA facilities. I test this directly in Section 2.4.3.

Additional results Table 2.2 presents coefficient estimates of Equation 2.5, for both the whole sample and the TM sample.[11] Table 2.2 shows similar patterns regardless of the method of measuring utilization: visits, relative value units (RVUs), or spending. Table 2.2 also presents additional results on the characteristics of chosen clinics. Veterans substitute toward less congested and closer clinics, decreasing their total wait and and travel times for care.

Appendix Table B.1 investigates the impacts of the policy — both directly and via equilibrium reductions in congestion — on veteran health. I find no detectable impact on mortality or inpatient admissions. This motivates my revealed preference approach, as the welfare effects of changes to outpatient utilization is better reflected in measures of consumer surplus than in coarse measurements of health like mortality. Table B.2 demonstrates the robustness of the results to including closest clinic by time fixed effects, which ensures that that I am comparing the same enrollees on either side of the 40-mile threshold. Results are similar.

Appendix 2.10.1 also includes results exploiting the wait-time eligibility conditions. These results document similar increases in private utilization at the clinic level, but no decrease in overall VA utilization, consistent with the assumption that wait times arise due to capacity constraints at the VA.

[11] The full set of event study figures, including Figure 2.3 replicated for the TM sample, is presented in Appendix **??**.

Heterogeneity To examine heterogeneous responses to p and, later, to w, I zoom in on the market for primary care. This yields the benefits of (1) making comparisons conditioning on a uniform set of products and prices, and (2) focusing on a domain where wait times are particularly well-measured. Appendix Figure B.3 plots the event study figures of Figure 2.3, focusing on primary care specifically, and documents similar patterns as in Figure 2.3.[12]

To investigate heterogeneity in responsiveness to p, I estimate Equation 2.4, split by income (above versus below median recorded income) and health status (above versus below median prior utilization). I focus on community primary care visits as the outcome of interest; the goal is to test whether utilization of community care responds to the change in price differently across demographic sub-groups.[13] These estimates are presented in Figure 2.4.

Figure 2.4 demonstrates that poorer, sicker veterans increase community utilization more in response to the change in price induced by the policy. This is the the primary equity or redistributive concern when allocating healthcare with a price mechanism: lower income consumers are often the most responsive to prices. This pattern is common across product markets, but presents a dilemma to a social planner with redistributive or fairness concerns in healthcare markets, as it highlights that raising prices will screen out lower socio-economic status consumers (Weitzman, 1977).

2.4.3 Results: Equilibrium Effects on w

Next, I turn to analyzing the effects of the policy on equilibrium wait times. Figure 2.3c demonstrates that a substantial share of eligible veterans substitute away from VA care when eligible for Choice. This presents the possibility that the policy could benefit *all* veterans, including those who are ineligible or inframarginal, via reductions in equilibrium clinic wait times.

Figure 2.5 plots coefficient estimates of Equation 2.6 to test this idea directly. I analyze the effect of Choice on wait times across all specialties and geographies (HRRs) (Figure 2.5a), and for primary care specifically at the HRR-level (Figure 2.5b) and the clinic level (Figure 2.5c). Coefficients in Figure 2.5 are scaled to represent a move from the 10th to the 90th percentile of the clinic exposure distribution. Figure 2.5 shows that the policy did indeed achieve its second goal of reducing congestion at the VA: wait times decreased by 13 days across all specialties and by 5 days in primary care, at the most exposed (90th percentile), relative to the least exposed (10th percentile), markets

[12] Substitution out of VA care is lower than overall, which is consistent with the smaller effects on wait times documented in Figure 2.5.

[13] In service of this aim, I also restrict the sample to be a population of veterans facing the same Choice-induced price change to interpret the effects as heterogeneity in price responsiveness, not heterogeneity in price changes. Specifically, I focus on the sample of Traditional Medicare veterans without a Medigap plan. I observe veterans' alternative insurance coverage because it is statutorily required for veterans to report this information to the VA.

and clinics. This is a 15-30% reduction in wait times based on the pre-period average of 34 days (primary care) and 43 days (overall). Appendix Figure B.4 also shows reductions in new patient wait times, a commonly used measure in prior work (Yee et al., 2022a; Pizer and Prentice, 2011b).

Heterogeneity While the motivation for price controls in healthcare is clear from policy discussion and substantiated by Figure 2.4, the substitute allocation mechanism is often determined endogenously rather than explicitly designed (Cutler, 2002a). It is unclear whether veterans' costs of waiting, γ_i, provide screening properties that are relatively advantageous to a social planner.

I investigate this by examining differential demand responses to the changes in wait time induced by the Choice Act. This is slightly more complicated than the analysis of heterogeneous effects on price because markets will experience different changes in wait times depending on their demographic composition. To address this, I estimate the following clinic-level two-stage-least-squares estimator:

$$w_{jt} = \theta \cdot \mathbf{z_{jt}} + \phi_j^w + \chi_t^w + \varepsilon_{jt} \tag{2.8}$$

$$\ln(s_{jt}^d) = \beta \cdot w_{jt} + \phi_j^s + \chi_t^s + \nu_{jt} \tag{2.9}$$

where $\mathbf{z_{jt}}$ are instruments for Choice exposure, $\left(\phi_j^w, \phi_j^s\right)$ and $\left(\chi_j^w, \chi_j^s\right)$ are market or clinic fixed effects and time fixed effects, and $\ln(s_{jt}^d)$ is the log market share of clinic j among demographic group d of *non-Choice eligible* veterans. I again zoom in on only the market for primary care.

The first stage equation, Equation 2.8, is a more general version of the difference-in-difference estimator discussed in Section 2.4.1 with a first stage illustrated in Figure 2.5. However, instead of parameterizing exposure as a continuous interaction, $\mathbf{z_{jt}} = 1\{t > 2014\} \cdot ShareEligPre_j$, which was amenable to event-study plots in Figure 2.5, I parameterize $\mathbf{z_{jt}}$ more flexibly as deciles of Choice eligibility exposure interacted with a post-2014 indicator. These instruments increase power for the more demanding task of examining heterogeneous demand responses to the equilibrium effects of the policy on wait times, while still isolating only the policy variation summarized in Figure 2.5.

Table 2.3 presents results. Perhaps surprisingly, Table 2.3 documents *similar* patterns of screening along income and health status as the price mechanism. Lower income veterans are more likely to reduce their utilization at a given clinic j due to high wait times at j than their higher income counterparts. Table 2.3 documents similar (adverse screening) patterns along measures of health capital, as proxied by lagged utilization.

Appendix Figure B.5 documents that these same screening patterns are replicated when using all panel variation in wait times in the sample, further supporting the conclusion of qualitatively similar patterns of screening across the two instruments.

2.4.4 Discussion

The reduced form analysis of the Choice Act yields insights both for the analysis of the two rationing mechanisms of interest and for understanding the performance of the Choice Act policy. First, veteran choices are responsive to both wait times and prices with qualitatively similar screening patterns. Second, the Choice Act benefited veterans both directly, through reduced prices, and indirectly, through reduced congestion, at an increased cost to the VA.

However, the analysis also leaves key questions unanswered. Making any inferences about allocations under alternative regimes or the efficiency or distributional consequences of waiting versus prices is complicated by the fact that I am not measuring and comparing willingness-to-pay and willingness-to-wait directly. Moreover, the welfare effects of the the Choice Act, and its performance relative to alternative policies, is ambiguous as it depends on veterans' willingness-to-pay to reduce wait times.

To this aim, the combined results in Section 2.4.2 and Figure 2.5 also demonstrate why the Choice Act provides an ideal policy setting to answer precisely these fundamental welfare questions about the choice of rationing mechanisms in healthcare. Obtaining the joint distribution of willingness-to-wait and willingness-to-pay requires a setting with exogenous variation in both prices and wait times. As the results above demonstrate, the Choice Act provides precisely this context. Within a given market, differences in eligibility isolate variation in prices. Across markets, conditional on eligibility, the share of others who are eligible shifts wait times.

In the next section, I will develop a model that uses the Choice Act variation to quantify the allocative, efficiency, and distributional effects of the two mechanisms, investigate the (ambiguous) welfare effects of the Choice Act, and compare it to alternative policy counterfactuals.

2.5 Clinic Choice Model

In this section, I develop a model of consumer demand for primary care in the presence of capacity constraints. Consumers make decisions over if and where to receive care, across VA and community options, as a function of observable characteristics, including travel time, wait time, and out of pocket prices, and unobserved shocks. VA clinics are subject to capacity constraints, and wait times arise endogenously to excess demand. I describe how I use the Choice Act policy variation to estimate heterogeneous preferences over characteristics of clinics, wait times, and out of pocket prices.

This model captures the consumer choice problem over a menu of options that vary along both prices and wait times, allowing me to simultaneously make use of all dimensions of the Choice

Act variation. The key outputs will be the joint distribution of willingness-to-wait and willingness-to-pay, which as discussed in the context of the stylized model in Section 2.2, is the key structural object for welfare and policy counterfactuals.

2.5.1 Demand

The utility for consumer i in market m choosing primary care clinic j in quarter t is given by the following random utility specification:

$$u_{ijmt} = \underbrace{\beta_{it}'\mathbf{X}_j}_{\text{prefs over clinics}} + \underbrace{\theta_i d_{ij}}_{\text{travel costs}} + \underbrace{\gamma_i w_{jt}}_{\text{waiting costs}} + \underbrace{\alpha_i p_{jmt}}_{\text{price sensitivity}} + \underbrace{\xi_{jmt}}_{\text{demand shock}} + \underbrace{\varepsilon_{ijmt}}_{\text{i.i.d. pref shock}} \tag{2.10}$$

with the value of the outside option to obtain no primary care normalized to zero. The key aspect of this demand specification is that consumers have preferences over both the efficient (p) and inefficient (w) market-clearing mechanisms.[14]

I define the geographic extent of markets as Hospital Referral Regions (HRRs) and define a market m as a subset of an HRR in which all veterans face the same vector of prices across clinics over time. I use a relatively large geographic healthcare market — there are 306 HRRs in the US — to capture the sparsity of VA clinics and the fact that veterans travel long distances for care.[15] Specifically, a market m is defined by a geography g (HRR) and an out-of-pocket payment class o. Because some veterans pay copayments and some do not, and distance-eligible veterans faced changes in the prices of community care under the Choice Act, while non-distance eligibles in the same HRR did not, there is within-HRR variation in prices that I segment into distinct markets.[16]

As in the stylized set-up in Section 2.2, consumers have preferences over waiting times parameterized by γ_i. The term $\alpha_i p_{jmt}$ represents a local approximation to the more general utility function in Section 2.2.

[14]This model abstracts away from consumer search, a phenomenon that may be important if consumers need to learn about the vector of wait times for clinics in their choice set. A few institutional features ameliorate this concern. First, when veterans call to schedule an appointment, they work with a scheduler, who can give them information about the options available to them. Second, a website is available for veterans to observe average wait times at each clinic: see https://www.accesstocare.va.gov/PWT/SearchWaitTimes.

[15]I define geographies as HRRs, as opposed to the more fine market definition of Hospital Service Area, HSA, of which there are 3,436, but only 1,128 VA clinics in my sample. See https://data.dartmouthatlas.org/downloads/methods/geogappdx.pdf for more information about the construction of these commonly used healthcare market definitions.

[16]This is why prices are indexed by jmt and wait times just by jt. There is cross-market variation in prices for the same clinic due to different out-of-pocket classes, o. There is no cross-market variation in wait times because all consumers face the same vector of w_{jt} within a geography, and each product is unique to a geography.

I parameterize heterogeneous preferences over care characteristics with $\beta_{it}'\mathbf{X}_j$ and $\theta_i d_{ij}$, which captures distaste for traveling to obtain care. I also allow preferences to depend on two unobservables: ξ_{jmt}, an unobserved demand shock common to all i for a given clinic-market-quarter, and ε_{ijmt}, a Type 1 Extreme Value idiosyncratic preference shock.

Consumers, or veteran enrollees, choose among all VA clinics and community providers in their HRRs, as well as the outside option of no care. VA clinics are well-defined, but the universe of potential community primary care providers is vast. I instead aggregate all community providers into 3,436 Hospital Service Area (HSA)-level — a much finer geographic market which typically contains only a single hospital — community providers to tractably capture consumer choice among geographically differentiated VA and community providers.

2.5.2 Supply

I treat VA providers and community providers asymmetrically.

VA providers For all VA clinics, I assume strict capacity constraints at observed levels of utilization κ_{jt}, measured as the number of primary care visits per quarter. The assumption of strict capacity constraints has only a very limited impact on my estimation strategy or counterfactual results: it simply disciplines the set of counterfactuals to consider only changes that hold the total number of VA visits constant.

I assume that wait times are generated via a First-Come-First-Served (FCFS) protocol, or that all potential consumers are treated identically in the queueing mechanism.[17] This is a much more substantive assumption, as it limits the extent to which wait times may be tailored to different consumer types.

I provide two pieces of evidence to substantiate this assumption. First, I have intentionally chosen a setting — primary care — where scope for prioritization is less likely,[18] and conversations with VA medical professionals support the assumption as an approximation to reality.[19] Second, the conclusions of these conversations are supported in the data. I show in Appendix Figure B.5 that the covariance between the clinic-level wait times of chosen clinics and patient characteristics

[17] Specifically, within each quarter, I assume consumers randomly arrive to the market, observe the menu of waiting times at each clinic, and enter their most preferred queue.

[18] Other healthcare settings offer rich environments to study prioritization in queues, e.g. the system of triage in emergency departments, where triage scores are explicitly recorded. Examining the efficiency of these prioritization rules in this much richer setting is an exciting area of future work to shed light on these questions of prioritization.

[19] Patients are encouraged to seek other care (at other locations) if they face particularly urgent needs, rather than jump the queue.

is robust to flexibly "risk-adjusting" clinic wait times based on the composition of patients that generate the underlying patient-level wait times.[20] This indicates that patterns of heterogeneity are dominated by demand-side choices — what I will attribute these patterns to — rather than supply-side differentiation, which is absorbed in the risk-adjustment step. Further, both of these patterns are consistent with screening patterns based on quasi-experimental shifts in wait times, presented in Table 2.3.

Community providers I assume community providers are un-capacity-constrained and provide care at constant marginal cost equal to Medicare Fee-For-Service rates. Community providers therefore can (1) provide care at zero wait times, and (2) absorb any changes in patient demand for community care as a result of VA policy changes. The first assumption is motivated primarily by data constraints: data on wait times at the VA are excellent, but extremely limited elsewhere. The second assumption is reasonable as the VA population is small relative to the overall population in an HRR and VA policy will not have substantive equilibrium effects on the whole market.

2.5.3 Equilibrium

Prices are regulated and the vector of VA wait times adjusts so that the following equilibrium condition holds:

$$\kappa_{jt} = \sum_i \frac{\exp\left(\beta_{it}'\mathbf{X}_j + \theta_i d_{ij} + \gamma_i w_{jt} + \alpha_i p_{jmt} + \xi_{jmt}\right)}{1 + \Sigma_{j' \in \mathscr{J}_m} \exp\left(\beta_{it}'\mathbf{X}_j + \theta_i d_{ij} + \gamma_i w_{jt} + \alpha_i p_{jmt} + \xi_{jmt}\right)} \tag{2.11}$$

for $j \in \mathscr{J}_m^{VA}$, or the set of VA clinics in each market.

2.5.4 Estimation

Endogeneity and identification I face two sources of endogeneity when estimating the preferences in Equation 2.10. First, $\mathbb{E}\left[w_{jt}\xi_{jmt}\right] \neq 0$. This form of endogeneity follows directly from Equation 2.11, as wait times are determined in equilibrium in response to demand shocks, ξ_{jmt}. This is the classic simultaneity problem faced in all markets. It impacts w_{jt} instead of p_{jmt} because in this setting, wait times flexibly adjust in equilibrium, while prices are regulated.

[20] I note that this is an overstatement of supply-side prioritization, because in the presence of stochastic fluctuations in wait times, the "risk-adjustment" step will purge variation from both any supply-side prioritization and selection due to demand-side preferences.

Despite the fact that prices are regulated, I still face the concern that $\mathbb{E}\left[p_{jmt}\xi_{jmt}\right] \neq 0$ for the variation in prices that exists in my setting. This is because, absent the Choice Act, all of the variation in prices is derived from cross-sectional variation across markets — veterans who pay copayments versus those who do not, either due to priority group or Medigap policies[21]— and across products — VA versus community providers. If there are unobservable differences in preferences across markets or products, this will lead to bias in α_i.

I use the Choice Act policy variation to address both sources of endogeneity. First, the direct effects of the policy, presented in Section 2.4.2, provides exogenous variation in p_{jmt} for community providers. Specifically, letting superscripts VA and C denote VA and community care, respectively, I parameterize preferences over clinic characteristics $\mathbf{X_j}$ as

$$\beta_{it}'\mathbf{X}_j = \beta_i^{VA} + \beta_i^C + \psi_o^{VA} + \psi_o^C + \chi_t^{VA} + \chi_t^C + \phi_j \tag{2.12}$$

where $\left(\beta_i^{VA}, \beta_i^C\right)$ capture heterogeneous preferences over VA and community care, $\left(\psi_o^{VA}, \psi_o^C\right)$ capture differences in preferences across out-of-pocket payment classes for VA and community care (in theory nested within i but included separately for clarity), $\left(\chi_t^{VA}, \chi_t^C\right)$ capture time effects in preferences over VA and community care, and ϕ_j are clinic fixed effects. I impose that $\mathbb{E}\left[p_{jmt}\xi_{jmt} | \psi_o^{VA}, \psi_o^C, \chi_t^{VA}, \chi_t^C, \phi_j\right] = 0$, isolating variation in prices coming only from the Choice Act policy change, as controlling for $\left(\psi_o^{VA}, \psi_o^C, \chi_t^{VA}, \chi_t^C, \phi_j\right)$ shuts down all other (cross-sectional) sources of price variation. This is a similar difference-in-difference identification strategy as in the related reduced form analysis in Section 2.4. Figure 2.6 plots the variation in the price of community care across the out-of-pocket payment classes o. This corresponds to the first stage of the reduced form event study effects of the policy in Section 2.4 for the product (primary care) and consumers (TM dual-eligibles) in my estimation sample. In my baseline specification, I use variation in prices from both wait time eligibility and distance eligibility conditions.

Second, I use the cross-product exposure to the Choice Act among VA clinics, presented in Section 2.4.3, to instrument for wait times at VA clinics.[22] Specifically, I impose the moment condition $\mathbb{E}\left[\mathbf{z_{jt}}\xi_{jmt}\right] = 0$, where $\mathbf{z_{jt}}$ are Choice eligibility instruments as in Equation 2.8. The idea of this instrument is similar to a Waldfogel (2003) IV, where the demand of other consumers in a market impact the wait times facing each individual consumer in equilibrium. In contrast to a standard Waldfogel (2003) IV, I use *policy-induced* changes in the choices of other consumers as my instrument.

Individual-level variation in distance to each of the clinics in a market provides additional variation

[21] Medigap policies cover the 20% coinsurance rate that Medicare beneficiaries must pay out of pocket for outpatient visits.

[22] Wait times at all community clinics are zero and therefore not endogenous.

to identify heterogeneity in consumer preferences over prices and wait times. Depending on where consumers are located relative to the menu of options, they face different trade-offs between prices, wait times, and other clinic characteristics. Under the additional assumption that $\mathbb{E}\left[d_{ij}\xi_{jmt}\right] = 0$, I use distance as an additional instrument to "trace-out" preferences over characteristics, w_{jt}, and p_{jmt} without requiring functional form extrapolations beyond the support of changes induced by the Choice Act (Berry and Haile, 2022). The assumption of $\mathbb{E}\left[d_{ij}\xi_{jmt}\right] = 0$ is commonly made in demand estimation exercises in healthcare. In fact, many estimates of healthcare demand invoke this assumption and measure willingness-to-pay in "travel" instead of "dollar" units (e.g. Einav et al. (2016)).

Selection In Equation 2.10, I treat wait times like "prices" at the product-by-quarter level. In reality, wait times are a stochastic process with daily fluctuations due to the random arrival of patients to the market within each quarter (Leshno, 2022; Ashlagi et al., 2022). Incorporating these day-to-day fluctuations is outside the scope of my model, however, the inherent stochasticity of wait times presents a potential selection problem in aggregation. Specifically, if consumers are more likely to decline an appointment when arriving (randomly, within t) to the market on days with high wait times, my measurement of w_{jt} will be selected and biased downward.

I address this issue by using my supply-side model to generate an unselected distribution of wait times. Specifically, if on any given day $\tilde{t}$ at any given provider, I observe no appointments being made, I assume that appointments available (and made) at date $\tilde{t}+1$ were available at date $\tilde{t}$. Appendix 2.10.2 provides more information. In particular, Appendix Figure B.6 recreates Figure 2.5, plotting the equilibrium effects of the policy on both the raw mean and the unselected measure of wait times. Results are similar.

Estimation procedure I parameterize heterogeneity in $\left(\beta_i^{VA}, \beta_i^C, \theta_i, \gamma_i, \alpha_i\right)$ as a function of veteran age bins, income bins (based on quartiles of the income distribution), VA priority group, and lagged utilization bins (based on quartiles of total VA spending in the prior year). I include an indicator for past use of the VA to capture cross-system inertia. Define $\left(\beta_0^{VA}, \beta_0^C, \theta_0, \gamma_0, \alpha_0\right)$ as the parameters corresponding to the (arbitrarily normalized) base group, and let $\left(\beta_b^{VA}, \beta_b^C, \theta_b, \gamma_b, \alpha_b\right)$ denote the parameter vector, relative to the base group, for all other bins b of observable heterogeneity. I subdivide the non-idiosyncratic component of utility into:

$$\underbrace{\beta_{it}'\mathbf{X}_j + \phi_j + \gamma_0 w_{jt} + \alpha_0 p_{jmt} + \xi_{jmt}}_{\delta_{jmt}\text{: common to all in } jmt} + \sum_b \underbrace{\beta_b^{VA} + \beta_b^C + \gamma_b w_{jt} + \alpha_b p_{jmt} + (\theta_0 + \theta_b)\, d_{ij}}_{\lambda_{ijmt}\text{: individual specific parameters}} \tag{2.13}$$

I then use a slightly-modified version of the two-step estimation approach of Goolsbee and Petrin (2004) to estimate the parameters in the random utility specification in 2.10. In a first step, I estimate the heterogeneity parameters in λ_{ijmt} via maximum likelihood including fixed effects for δ_{jmt}. To ease computational constraints, I first estimate λ_{ijmt} on a random sample of markets and time periods, estimating λ_{ijmt} and δ_{jmt} jointly via maximum likelihood.[23] Then, given $\hat{\lambda}_{ijmt}$, I loop over markets to estimate $\hat{\delta}_{jmt}$ by matching market shares exactly. In the second step, I estimate the parameters in δ_{jmt} via IV, using instruments for Choice exposure (i.e. imposing the moment condition $\mathbb{E}\left[\mathbf{z}_{jt}\xi_{jmt}\right] = 0$) and variation in prices from Choice conditional on $\left(\psi_o^{VA}, \psi_o^C, \chi_t^{VA}, \chi_t^C, \phi_j\right)$. I estimate parameters on one pre-period year (2013) and one post-period year (2017), intended to capture the effect of Choice after the ramp-up period documented in Figure 2.3. I employ this estimation strategy on the sample of veterans dually enrolled in Traditional Medicare. In this sample, I observe the universe of consumption decisions in all time periods across VA and community care.

2.5.5 Estimates

No heterogeneity I first present results with no preference heterogeneity beyond the out-of-pocket payment class heterogeneity in δ_{jmt}. This allows me to present interpretable parameter results — average preferences, rather than preferences for a normalized group — from a simple regression, with the outcome of $\delta_{jmt} = \ln\left(s_{jmt}\right) - \ln\left(s_{0mt}\right)$ (Berry, 1994). The structure of the discrete Choice set-up thus allows me to estimate preferences over both prices and wait times jointly using the Choice Act variation in a simple and interpretable regression. I use this simplified model to illustrate these responses and the performance of the Choice Act instrument in Table 2.4.

Column (1) presents results from a panel regression, estimating α from Choice Act variation in prices but γ from all within-clinic variation, residualized of overall time trends in VA and community care. Column (1) documents a negative coefficient on price that implies an average elasticity of demand to a given clinic of -0.30. The average estimated elasticity to a given clinic is slightly higher than both estimates of price elasticities for individual hospitals for inpatient care (Prager, 2020) and the price elasticity for total outpatient care estimated in Chandra et al. (2010a). These comparisons appear reasonable given the differences in setting (outpatient versus inpatient) and elasticities (overall versus product-level).

The coefficient on wait times (γ) in column (1) however, is positive and insignificant. This may

[23] I take this random sample to avoid needing to jointly search over the thousands of δ_{jmt} for each parameter guess given limited computational power. Sampling at the market level captures heterogeneity parameters without introducing bias from zero market shares from small samples. I am currently using a 4% random sample (which, for 9,000,000 enrolled veterans and 8 quarters still includes a very large number of choice instances). I plan to increase this sample until computation burdens become too severe.

lead a researcher to (incorrectly) infer that waiting is not costly to consumers and therefore imposes no efficiency costs.

In column (2), I instrument for w_{jt} using the same $\mathbf{z}_{jt}$ as in the 2SLS exercise in Section 2.4, using deciles of Choice exposure. When instrumented, γ becomes negative – consistent with Table 2.3 – and moreover, I estimate that (at prices and wait times in my sample) veterans are three times as elastic to wait times as to prices. This underscores the highly endogenous nature of wait times, as they adjust extremely flexibly to demand shocks, and the need for instruments. In column (3), I use a continuous interaction between exposure share and post-Choice in a just-identified specification. I find very similar results, though slightly less precise, providing reassurance that no one parameterization of Choice exposure is driving results.

Table 2.4 also presents an estimate of the ($) cost of delay, $\frac{\gamma}{\alpha}$, of approximately $2.50. This means that the average veteran would be willing to pay approximately $2.50 to move an appointment at the same clinic one day sooner. This object, denoted by $C(w)$ in the more general conceptual framework in Section 2.2, is the key welfare object of interest. Its magnitude governs the welfare costs of rationing, and its distribution, the redistribution of surplus.

Adding heterogeneity Figure 2.7a plots the distribution of $\frac{\gamma_i}{\alpha_i}$ after incorporating heterogeneity and documents substantial dispersion.[24] The status quo waiting-based rationing regime is implicitly price discriminating in favor of those with low costs of waiting $\frac{\gamma_i}{\alpha_i}$ and against those with high $\frac{\gamma_i}{\alpha_i}$. Appendix Table B.4 provides a detailed description of the heterogeneity parameter estimates underlying $\frac{\gamma_i}{\alpha_i}$ using the two-step estimation procedure described in the previous section.

Figure 2.7b shows that the costs of delay are increasing in income, implying that rationing "price-discriminates" in favor of lower income consumers. The pattern on health status is slightly more mixed: while on average, waiting costs are lower for sicker veterans (proxied by utilization), the relationship is U-shaped, reflecting the fact that the sickest veterans find waiting particularly costly.

The patterns in Figure 2.7 relate to the raw patterns of screening in Section 2.4. Section 2.4 documented qualitatively similar screening patterns on the two instruments, and indeed, that is reflected in a positive correlation between α_i and γ_i, shown in Appendix Figure B.7. However, these descriptive patterns were limited by the fact that I was not comparing willingness-to-wait and

[24] I note that for the approximately 38% of veterans who have no prior history using the VA (at all, i.e. not just for primary care), I estimate that they are essentially completely inelastic (with zero, or small positive coefficients) on price. This is unsurprising, as VA care and VA policy changes may simply not be under consideration for these veterans. Rather than imposing that all veterans have negative α_i and including these veterans at negative α_i very close to zero, I simply exclude them from estimates of $\frac{\gamma_i}{\alpha_i}$ and counterfactuals. Essentially, I assume that these veterans are inert to any VA counterfactuals, which seems to be a reasonable assumption. After excluding these veterans, 100% of the consumer types in my sample have a negative cost of waiting, and less than 2% have a non-negative α_i.

willingness-to-pay directly. The results in the previous section were confounded by preferences for care and the fact that veterans were potentially substituting on different margins of both the outside option and to other care. The model presented and estimated in this section allows me to account for those multiple margins explicitly. In doing so, and in the process, incorporating additional variation from the geographic distribution of veterans and clinics, I am able to quantify how the distribution of willingness-to-pay and willingness-to-wait diverge, illustrated in Figure 2.7.

2.6 Alternative Rationing Mechanisms, Welfare, and Policy Counterfactuals

In this Section, I quantify the efficiency and distributional properties of the two rationing mechanisms in equilibrium and assess where the use of rationing via wait times places a planner on a efficiency-redistributive frontier (if on the frontier at all). I also return to the Choice Act policy context and use my estimates to quantify the (ambiguous) welfare effects of the Choice Act and alternative policies. A key feature of my analysis is to compare the effects of commonly used policies to manage rationing costs to shifts in the allocation mechanism altogether.

2.6.1 Alternative Rationing Mechanisms

I examine the allocative, efficiency, and distributional effects of the status quo rationing regime, where care is allocated via wait times at approximately zero prices, relative to a benchmark in which care can be obtained immediately with no waits, subject to a price (the efficient benchmark). In these counterfactual exercises, I hold total utilization at the VA constant (i.e. I impose the assumption of strict capacity constraints), and simply change the allocation mechanism. I thus focus on allocative effects alone, without taking a stand on whether total VA capacity is too high or too low. Specifically, for any given vector of VA wait times (prices), I search for a vector of prices (wait times) such that the equilibrium condition in Equation 2.11 holds. There exists an infinite number of equilibria along combinations of (p, w). In counterfactuals, I fix one instrument and search for another, yielding a J_{gt}^{VA} -dimensional system of equations with a J_{gt}^{VA}-dimensional vector of unknowns for each geography g and time t.

I simulate counterfactual outcomes over the entire analysis period (pre- and post-Choice). My results therefore capture the average effects of the two rationing mechanisms under the two policy regimes. I analyze the contribution of the Choice policy to consumer welfare in Section 2.6.2.

My counterfactual analyses only focus on the sub-set of Traditional Medicare dual-eligibles that I use to estimate preferences in Equation 2.10. These analyses should thus be interpreted as exercises in which this population is the only population of interest, or that choices in this population are identical to choices in the overall population. The focus on this (substantial) sub-set of consumers is unlikely to change the qualitative conclusions, as over three-quarters of veterans also have alternative sources of insurance coverage and both demographics and VA utilization are generally similar across these two populations (see Table 2.1).

Positive analysis I begin by presenting positive effects on equilibrium prices and allocations without (yet) imposing a normative assumption that choices reveal value. These counterfactuals only require that my estimates accurately describe the choices that veterans make in response to prices and wait times. Table 2.5 presents results. Column (1) presents wait times, prices, and allocations under the status quo waiting-based regime, and column (2) presents these same outcomes while imposing zero waits and instead allowing prices to flexibly adjust in response to demand and capacity constraints κ_{jt}.

Table 2.5 illustrates that prices would have to be almost $80 on average to achieve zero wait times, about 20% of estimates of the VA cost of service. This value is larger than any outpatient copayments the VA currently charges, with a maximum of $50 for specialty care for certain veterans. Though the waiting costs imposed on veterans are substantial, these costs are not being collected as revenue, as would the $78 per visit under a price regime.

Table 2.5 also demonstrates substantial allocative effects across the two regimes. Slightly fewer veterans obtain any care at all under the price regime. This is flexible to adjust, even in the presence of strict capacity constraints at the VA, because of the opportunity to substitute to community care as well as the outside option of no care. As a share of the total number of veterans receiving care under the status quo regime, 4% of all veterans are displaced out of any care, and 16% are displaced out of VA care — with an equivalent, different set of 16% of veterans substituting in. These results highlight that the two rationing regimes may induce substantially different distributions of access to care and consumer surplus.

Panel C of Table 2.5 presents descriptive evidence of the distributional effects of the alternative rationing regimes. The average consumer receiving VA care under the wait-time rationing regime is over 4% lower income, 4% sicker (as proxied by lagged costs), and 1% older than those who would obtain VA care under the price regime. These patterns still exist, but are less strong, for veterans receiving any care at all, after accounting for substitution to community care. Even without any normative framework, these results present the core trade-off facing a planner, perhaps offering an explanation as to the varied use of these two regimes around the world. The price regime

allocates a scarce commodity to those who value it, while generating revenue. The waiting time regime, though strictly money burning, changes the landscape of access to care in favor of lower socio-economic status, and thus likely higher g_i, individuals.

Quantifying welfare effects Table 2.6 quantifies the efficiency and redistributive effects of the two regimes described in Table 2.5. Column (1) presents the decomposition of the welfare effects of status quo rationing, relative to a price regime from Equation 2.3 in Section 2.2. I conduct this decomposition using a standard money-metric welfare framework that puts equal weight on consumer surplus and revenue collected. Put differently, column (1) evaluates the welfare costs of rationing under a Kaldor (1939)-Hicks (1939) efficiency criterion.

Consumer surplus is $57.25 lower (per veteran, per year) under the rationed regime versus a price regime. This is driven by a combination of allocative effects (60%), as care is not being allocated to those who value it the most, and by a price discrimination effect, as waiting is more costly for inframarginals than for the marginal consumer (30%). This price discrimination effect comes from the fact that waiting is differentially costly for sicker veterans who are more likely to have very high value for the VA (i.e. the $C(w^*)$ curve in Figure 2.1 is upwards- not downwards-sloping).

An additional $62.90 per veteran, per year is further lost from a lack of recouped revenue. This is due to the fact that waiting is a costly screen, and therefore pure social waste. Together, the combined efficiency cost of the status quo rationing regime is $113, or 24% of achievable surplus. This is a substantial efficiency cost that must be weighed against any redistributive motives for imposing rationing.

Columns (2) and (3) of Table 2.6 examine the redistribution of surplus directly. I find that although consumer surplus is on average over $57 lower under a price regime, more than half of veterans (55%) prefer the status quo wait time regime. This finding underscores why the choice of rationing regimes in healthcare is so controversial in practice: the choice of one over another generates substantial winners and losers. Moreover, those who prefer the status quo rationing regime are starkly different than those who do not: they are substantially poorer, sicker, and older. Table 2.6 thus quantifies the core equity-efficiency trade-off in the choice of rationing mechanism: waiting destroys 24% of surplus, but in the process redistributes substantially — both in terms of welfare, and in terms of access to care (Table 2.5) — to seemingly high g_i individuals.

Characterizing an efficiency-redistributive trade-off How should one characterize or benchmark the magnitude of this trade-off? Knowledge of a planner's preference for redistribution in this specific context is inherently unknown. In this sub-section, I offer two methods of quantifying the magnitude of this trade-off.

Inverse optimum weights from the tax schedule First, I compare the cost of transferring \$1 from an individual who prefers the price regime (column (3) of Table 2.6) to an individual who prefers the rationed regime (column (2) of Table 2.6) to the costs of transferring \$1 from the richest to the poorest person in society via the current tax schedule. This "inverse optimum" approach takes values of g_i inferred from the tax schedule, or alternatively, compares the efficiency of redistribution occurring in this context to redistribution that occurs via the tax schedule. This approach modifies the standard Kaldor (1939)-Hicks (1939) notion of efficiency to incorporate the feasibility of transfers, i.e. that it is costly to transfer from the rich to the poor. Note that this comparison is already conservative in favor of the wait time regime as it considers only changes in consumer surplus. This is relaxed in the following exercise, which incorporates revenue.

Hendren (2020) calculates these inverse optimum welfare weights from the status quo US tax schedule and finds that the differences in these weights is bounded above by two. Put differently, it always costs less than \$2 to transfer \$1 from a person at the top of the income distribution to a person at the bottom. In comparison, the redistribution that occurs via the status quo waiting-based rationing mechanism, relative to a price-mechanism, destroys over \$5 of consumer surplus when transferring \$1 from the losers (column (3)) to the winners (column (2)). This suggests that, even without accounting for revenue, the social planner must place substantially more weight on either (1) equality in healthcare consumption than equality in income, or (2) redistribution between these two groups than between the richest and poorest person in society, for the use of rationing by wait times to be preferred.

Welfare after revenue redistribution The analysis using inverse optimum weights ignored differences in revenue generated by the mechanism. A second way to characterize the trade-off is to consider what share of the population is better off under the status quo rationing regime, relative to a price mechanism, after redistributing revenue. In this setting, as in many others, a government entity is providing the healthcare services — and thus would collect the revenue — so it is feasible to assume that this revenue could in theory be redistributed back into the population.

In Table 2.7, I calculate the share of veterans better off under the status quo, relative to a price mechanism, after redistributing revenue, modulating the value of government revenue between zero and one. At worst, all revenue is burned, and at best, I considers a scenario in which the government can only redistribute with the blunt instrument of uniform transfers. Table 2.7 documents that even in a regime where 50% of revenue is lost upon collection, only 10% of veterans prefer the wait time regime after redistributing this revenue. This number drops to 5% and 3%, respectively, under 75% and 100% redistribution of revenue. This exercise highlights that even with only blunt instruments available to a planner — at best uniform transfers — the lost revenue is so substantial

relative to the magnitude of expected welfare differences that the planner can achieve close to a Pareto improvement.

This exercise underscores the importance of considering alternative revenue-generating mechanisms among the class of screening instruments available to a social planner. The classic idea of a Nichols and Zeckhauser (1982) ordeal-screen is that the screen can potentially better target a transfer to individuals who value that transfer. Of course, this begs the question not only of whether the ordeal screen achieves more favorable targeting than an alternative mechanism, for example, a consumption tax on the in-kind transfer, but whether this favorable targeting outweighs any lost revenue due to the imposition of a screen that is pure social waste. While this point is not new (Olken, 2016), few papers have quantified the lost surplus directly in order to weigh it against alternative instruments to target a transfer.

Discussion The two exercises above lead to the same qualitative conclusion. Though there is a meaningful trade-off between efficiency and distributional concerns in the planner's choice of rationing mechanism, the efficiency losses from rationing are so large that they likely outweigh any potential redistributive benefits. Of course, a planner may still choose to use a wait-time-based rationing mechanism if the planner places very high value on the distribution of surplus that is achieved by this rationing mechanism in this specific dimension of consumption. The purpose of this section — and a primary contribution of this paper — is simply to document that the slope of the trade-off facing a planner is steep: efficiency losses are very large, relative to redistributive benefits.

Why is this case? The key pattern in the data that drives this conclusion is that prices and wait times screen on qualitatively similar dimensions. Because of this, wait times generate substantial deadweight loss without a sufficiently large enough reallocation of surplus. Moreover, waiting costs for inframarginals are high, such that for any (socially desirable) change in allocation that the waiting-based rationing mechanism yields, it induces large deadweight losses among those who are not changing their behavior.

Due to the large welfare losses from rationing, a natural question to ask is: given the ubiquity of rationing in practice, can welfare costs can be managed by effective "managed rationing" policies, such as the Choice Act?

2.6.2 Policy analysis: Choice versus Alternative Policies

In addition to providing the variation necessary to answer the key welfare questions of this paper, the Choice Act provides a useful laboratory to examine common second-best policies to manage

rationing costs. The welfare effects of the Choice Act are ambiguous: Figure 2.3e documented that total spending increased, while Figure 2.5 documented that wait times decreased. The preference estimates obtained using the Choice Act variation now provide an opportunity to trade-off these opposing forces and examine the performance of this large change to VA policy.

Table 2.8 examines the welfare effects of the Choice Act and alternative policies, and benchmarks the effects of these policy changes against the effects of eliminating wait times altogether by rationing by a price mechanism instead. I tabulate changes in consumer surplus, government revenue, and their sum.

Column (1) tabulates the gains from switching to a price mechanism, reproducing the results from the previous section. Column (2) examines the welfare effects of the Choice Act. While the Choice Act unambiguously increased consumer surplus — it made both directly eligible and ineligible veterans better off — this increase is not enough to outweigh the increase in costs. The Choice Act was thus welfare-decreasing overall. Expansions in Choice-style policy, either via later factual expansions, including the MISSION Act, and counterfactual expansions to everybody, have the same effect — increases in costs outweigh the increase in consumer surplus.

The Choice Act and similar expansions of subsidized community care increase costs more than the gains in consumer surplus for two reasons. First, as the positive counterfactuals in Table 2.5 illustrated, the marginal consumer at the VA was only willing to pay 20% of the cost of care, so increases in "effective" capacity via further subsidizing the outside option, brought "prices" further from, instead of closer to, social marginal cost. Put differently, the market failure driving rationing costs at the VA, which motivated the policy change, was not insufficient capacity of VA care, but rather an inefficient mechanism to allocate that capacity.

Second, the Choice Act policy is poorly targeted: all veterans receive the same subsidy for community care, regardless of their externality on others using the VA (Diamond, 1973). Veterans are treated identically by the policy regardless of whether they substitute away from the VA, and generate a positive externality from reducing congestion, or from no care, which has no such positive externality. As Figure 2.3 documents, for every person that substitutes away from the VA and alleviates congestion, one person substitutes from the outside option, increasing spending by more than their value of the service.

While the specific conclusion that the Choice Act policy increased costs more than benefits is unique to this setting, the challenge of designing policies with heterogeneous consumption externalities is a broader challenge to health policy designers implementing similar programs to reduce public sector queues. Subsidizing private care offers flexibility, relative to expanding public capacity, but heterogeneity in consumers' substitution patterns can complicate the design of the private subsidy. Policies that either (1) target subsidies based on (expected) substitution patterns, or (2)

increase (targeted) prices directly on the congested good, can increase welfare relative to a uniform or poorly targeted subsidy for private care.

Column (5) investigates the performance of this second class of policies in the form of a small, targeted copay increase at the VA. In 2001, copayments for primary care were decreased from slightly over $50 to $15. Column (5) investigates the effect of switching copayments to $50 — approximately 2001 levels — for only the 25% of veterans who already pay copayments. This group of veterans is on average substantially higher income than the veterans who are not obligated to pay copayments (median income of $44,127 versus $23,120). This copay increase increases consumer surplus by substantially more than any of the Choice Act policies despite *increasing prices* for some veterans. This is because levying the price increase on the congested good is much more effective at reducing congestion externalities and wait times in equilibrium. And because copayments are targeted at veterans who are on average higher socio-economic status than those who are not obligated to pay copayments, this policy disproportionately benefits low-income veterans. Finally, this policy generates revenue, rather than increasing costs.

Small, targeted copayment increases can thus increase welfare substantially, while still maintaining distributional objectives. Relaxing a planner's desire to relinquish a price mechanism entirely — in a feasible manner — can thus generate much larger gains than other second-best policies that manage rationing costs without relaxing this constraint. However, Table 2.8 also documents that the gains from even this most effective policy are dwarfed by changing the allocation mechanism altogether, which yields welfare gains that are almost an order of magnitude larger. Thus, policy designers must think carefully about whether price controls, in the presence of endogenously arising queues for care, achieve equity goals that outweigh their efficiency costs. The analysis in this paper suggests that they likely do not. Of course, more sophisticated non-price allocation mechanisms than the simple queuing used in practice could change this calculus.

2.7 Conclusion

This paper studies the efficiency and distributional consequences of the choice of healthcare rationing mechanism and analyzes policy in the presence of price controls and capacity constraints. I focus specifically on the two most commonly used and debated demand-side rationing mechanisms: prices and wait times. I leverage the rich data and policy environment of the VA and the Choice Act to make progress on questions that have previously been severely limited by data constraints and the challenge of conducting welfare analysis for rationed goods. I combine quasi-experimental variation and careful policy analysis with an equilibrium model to examine welfare trade-offs and policy counterfactuals.

I begin by analyzing the effects of the Choice Act, and document that the policy increased access to care via both reductions in prices for eligible veterans and wait times in equilibrium. I show that veterans are responsive to both rationing mechanisms, and that these two rationing mechanisms screen on qualitatively similar dimensions.

To evaluate the efficiency and distributional consequences of the rationing mechanisms of interest, I use this variation to estimate a model of clinic choice and queuing for primary care. I conduct counterfactuals that hold VA capacity constant and change the method of allocating care. The two regimes lead to meaningful differences in allocations. I find that a planner may face a trade-off between efficiency and allocating healthcare to lower income, sicker consumers, but that this trade-off is steep, destroying substantial surplus for any (potentially) socially desirable reallocation. Because of this, I document that feasible copayment increases substantially improve upon the status quo Choice Act policy.

Many healthcare systems across the globe have elected to eliminate financial barriers to care. However, in the presence of capacity constraints, other barriers, typically wait times, have emerged. Carefully evaluating the distinct consequences of these two mechanisms is essential to understanding whether the choice to ration access to care via wait times is advantageous or detrimental to a policymaker's goals. The results in this paper suggest that rationing via wait times imposes efficiency costs that likely substantially hinder a policymaker's objectives. Investigating the performance of other domains of non-market allocation in healthcare presents an exciting avenue for future research.

2.8 References

113th Congress, "Veterans Access, Choice, and Accountability Act of 2014," 2014.

115th Congress, "Maintaining Internal Systems and Strengthening Integrated Outside Networks (MISSION) Act of 2018," 2018.

Akbarpour, Mohammad, Piotr Dworczak, and Scott Duke Kominers, "Redistributive Allocation Mechanisms," June 2022.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, Ririn Purnamasari, and Matthew Wai-Poi, "Self-Targeting: Evidence from a Field Experiment in Indonesia," *Journal of Political Economy*, April 2016, *124* (2), 371–427. Publisher: The University of Chicago Press.

Ashlagi, Itai, Jacob Leshno, Pengyu Qian, and Amin Saberi, "Price Discovery in Waiting Lists," August 2022.

Berry, Steven T., "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 1994, *25* (2), 242–262. Publisher: [RAND Corporation, Wiley].

_ **and** _ , "Nonparametric Identification of Differentiated Products Demand Using Micro Data," April 2022. arXiv:2204.06637 [econ].

Besley, Timothy and Stephen Coate, "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs," *The American Economic Review*, 1992, *82* (1), 249–261. Publisher: American Economic Association.

_ **, John Hall, and Ian Preston**, "The demand for private health insurance: do waiting lists matter?," *Journal of Public Economics*, 1999, *72* (2), 155–181. Publisher: Elsevier.

_ **, Samantha Burn, Timothy Layton, and Boris Vabson**, "Rationing Medicine Through Bureaucracy: Authorization Restrictions in Medicare," January 2023.

Buchholz, Nicholas, Laura Doval, Jakub Kastl, Filip Matějka, and Tobias Salz, "The Value of Time: Evidence From Auctioned Cab Rides," May 2020.

Bulow, Jeremy and Paul Klemperer, "Regulated Prices, Rent Seeking, and Consumer Surplus," *Journal of Political Economy*, 2012, *120* (1), 160–186. Publisher: The University of Chicago Press.

C., Jr Chan David, David Card, and Lowell Taylor, "Is There a VA Advantage? Evidence from Dually Eligible Veterans," February 2022.

Castillo, Juan Camilo, "Who Benefits from Surge Pricing?," August 2022.

_ **, Jonathan Gruber, and Robin McKnight**, "Patient Cost-Sharing and Hospitalization Offsets in the Elderly," *American Economic Review*, March 2010, *100* (1), 193–213.

Chartock, Benjamin L., "Quality Disclosure, Demand, and Congestion: Evidence from Physician Ratings," February 2023.

_ **, Ian Gale, and Jinwoo Kim**, "Assigning Resources to Budget-Constrained Agents," *The Review of Economic Studies*, January 2013, *80* (1), 73–107.

Condorelli, Daniele, "What money can't buy: Efficient mechanism design with costly signals," *Games and Economic Behavior*, July 2012, *75* (2), 613–624.

Cooper, Zack, Stephen Gibbons, and Matthew Skellern, "Does competition from private surgical centres improve public hospitals' performance? Evidence from the English National Health Service," *Journal of Public Economics*, October 2018, *166*, 63–80.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora, "Do Labor Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment *," *The Quarterly Journal of Economics*, May 2013, *128* (2), 531–580.

Cutler, David M., "Equality, Efficiency, and Market Fundamentals: The Dynamics of International Medical-Care Reform," *Journal of Economic Literature*, September 2002, *40* (3), 881–906.

Davis, Lucas W. and Lutz Kilian, "The Allocative Cost of Price Ceilings in the U.S. Residential Market for Natural Gas," *Journal of Political Economy*, April 2011, *119* (2), 212–241. Publisher: The University of Chicago Press.

Department of Veterans Affairs, "FY 2024 Budget Submission: Budget in Brief," Technical Report 2023.

Deshpande, Manasi and Yue Li, "Who Is Screened Out? Application Costs and the Targeting of Disability Programs," *American Economic Journal: Economic Policy*, November 2019, *11* (4), 213–248.

Diamond, Peter, "Consumption Externalities and Imperfect Corrective Pricing," *Bell Journal of Economics*, 1973, *4* (2), 526–538. Publisher: The RAND Corporation.

Dupas, Pascaline, Vivian Hoffmann, Michael Kremer, and Alix Peterson Zwane, "Targeting health subsidies through a nonprice mechanism: A randomized controlled trial in Kenya," *Science*, August 2016, *353* (6302), 889–895. Publisher: American Association for the Advancement of Science.

Dworczak, Piotr, "Equity-efficiency trade-off in quasi-linear environments," *GRAPE Working Papers*, 2022. Number: 70 Publisher: GRAPE Group for Research in Applied Economics.

_ **, Scott Duke Kominers, and Mohammad Akbarpour**, "Redistribution Through Markets," *Econometrica*, 2021, *89* (4), 1665–1698.

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael Walker, "General Equilibrium Effects of Cash Transfers: Experimental Evidence From Kenya," *Econometrica*, 2022, *90* (6), 2603–2643. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA17945.

Einav, Liran, Amy Finkelstein, and Heidi Williams, "Paying on the margin for medical care: Evidence from breast cancer treatments," *American Economic Journal. Economic Policy*, February 2016, *8* (1), 52–79.

_ **and Matthew J Notowidigdo**, "Take-Up and Targeting: Experimental Evidence from SNAP*," *The Quarterly Journal of Economics*, August 2019, *134* (3), 1505–1556.

Fréchette, Guillaume R., Alessandro Lizzeri, and Tobias Salz, "Frictions in a Competitive, Regulated Market: Evidence from Taxis," *American Economic Review*, August 2019, *109* (8), 2954–2992.

Glaeser, Edward L. and Erzo F. P. Luttmer, "The Misallocation of Housing Under Rent Control," *American Economic Review*, September 2003, *93* (4), 1027–1046.

Goolsbee, Austan and Amil Petrin, "The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV," *Econometrica*, 2004, *72* (2), 351–381. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0262.2004.00494.x.

Hartline, Jason D. and Tim Roughgarden, "Optimal mechanism design and money burning," in "Proceedings of the fortieth annual ACM symposium on Theory of computing" STOC '08 Association for Computing Machinery New York, NY, USA May 2008, pp. 75–84.

Hendren, Nathaniel, "Measuring economic efficiency using inverse-optimum weights," *Journal of Public Economics*, July 2020, *187*, 104198.

Hicks, J. R., "The Foundations of Welfare Economics," *The Economic Journal*, 1939, *49* (196), 696–712. Publisher: [Royal Economic Society, Wiley].

Kaldor, Nicholas, "Welfare Propositions of Economics and Interpersonal Comparisons of Utility," *The Economic Journal*, 1939, *49* (195), 549–552. Publisher: [Royal Economic Society, Wiley].

Leshno, Jacob D., "Dynamic Matching in Overloaded Waiting Lists," *American Economic Review*, December 2022, *112* (12), 3876–3910.

_ **and Lee M. Lockwood**, "Targeting with In-Kind Transfers: Evidence from Medicaid Home Care," *American Economic Review*, April 2019, *109* (4), 1461–1485.

Mark, Nathaniel, "Access to Care in Equilibrium," 2023.

Martin, Stephen and Peter C. Smith, "Rationing by waiting lists: an empirical investigation," *Journal of Public Economics*, January 1999, *71* (1), 141–164.

Nichols, Albert L. and Richard J. Zeckhauser, "Targeting Transfers through Restrictions on Recipients," *The American Economic Review*, 1982, *72* (2), 372–377. Publisher: American Economic Association.

Nichols, D., E. Smolensky, and T. N. Tideman, "Discrimination by Waiting Time in Merit Goods," *The American Economic Review*, 1971, *61* (3), 312–323. Publisher: American Economic Association.

OECD, "Waiting Times for Health Services: Next in Line," Technical Report May 2020.

Olken, Benjamin A., "Hassles versus prices," *Science (New York, N.Y.)*, August 2016, *353* (6302), 864–865.

Pizer, Steven D. and Julia C. Prentice, "Time is money: outpatient waiting times and health insurance choices of elderly veterans in the United States," *Journal of Health Economics*, 2011, *30* (4), 626–636. Publisher: Elsevier.

_ **and** _ , "What are the consequences of waiting for health care in the veteran population?," *Journal of General Internal Medicine*, 2011, *26* (2), 676–682. Publisher: Springer.

Prager, Elena, "Healthcare Demand under Simple Prices: Evidence from Tiered Hospital Networks," *American Economic Journal: Applied Economics*, October 2020, *12* (4), 196–223.

Propper, Carol, Simon Burgess, and Denise Gossage, "Competition and Quality: Evidence from the NHS Internal Market 1991–9*," *The Economic Journal*, 2008, *118* (525), 138–170. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0297.2007.02107.x.

Reinhardt, Uwe, *Priced Out: The Economic and Ethical Costs of American Health Care*, Princeton University Press, 2019.

Reinhardt, Uwe E., "Wanted: A Clearly Articulated Social Ethic for American Health Care," *JAMA*, November 1997, *278* (17), 1446–1447.

_ , "Keeping Health Care Afloat: The United States versus Canda," *Milken Institute Review*, 2007, *Second Quarter.*

Ringard, Ånen, Ingrid Sperre Saunes, and Anna Sagan, "The 2015 hospital treatment choice reform in Norway: Continuity or change?," *Health Policy*, April 2016, *120* (4), 350–355.

Rose, Liam, Marion Aouad, Laura Graham, Lena Schoemaker, and Todd Wagner, "Association of Expanded Health Care Networks With Utilization Among Veterans Affairs Enrollees," *JAMA Network Open*, October 2021, *4* (10), e2131141.

Ryan, Nicholas and Anant Sudarshan, "Rationing the Commons," *Journal of Political Economy*, January 2022, *130* (1), 210–257. Publisher: The University of Chicago Press.

Saez, Emmanuel and Stefanie Stantcheva, "Generalized Social Marginal Welfare Weights for Optimal Tax Theory," *American Economic Review*, January 2016, *106* (1), 24–45.

Saruya, Hiroki, Todd Wagner, and Diana Zhu, "Complementing Public Care with Private: Evidence from Veterans Choice Act," January 2023.

Tobin, James, "On Limiting the Domain of Inequality," *The Journal of Law & Economics*, 1970, *13* (2), 263–277. Publisher: [University of Chicago Press, Booth School of Business, University of Chicago, University of Chicago Law School].

Train, Kenneth E., *Discrete Choice Methods with Simulation*, 2nd edition ed., Cambridge ; New York: Cambridge University Press, June 2009.

U. S. Government Accountability Office, "VA Health Care: More National Action Needed to Reduce Waiting Times, but Some Clinics Have Made Progress," Technical Report August 2001.

_ , "VA Health Care: Reliability of Reported Outpatient Medical Appointment Wait Times and Scheduling Oversight Need Improvement," Technical Report December 2012.

_ , "VA Health Care: Actions Needed to Improve Newly Enrolled Veterans' Access to Primary Care," Technical Report March 2016.

_ , "Veterans Health Care: Opportunities Remain to Improve Appointment Scheduling within VA and through Community Care," Technical Report July 2019.

Waldfogel, Joel, "Preference Externalities: An Empirical Study of Who Benefits Whom in Differentiated-Product Markets," *The RAND Journal of Economics*, 2003, *34* (3), 557–568. Publisher: [RAND Corporation, Wiley].

Waldinger, Daniel, "Targeting In-Kind Transfers through Market Design: A Revealed Preference Analysis of Public Housing Allocation," *American Economic Review*, August 2021, *111* (8), 2660–2696.

Weitzman, Martin L., "Is the Price System or Rationing More Effective in Getting a Commodity to Those Who Need it Most?," *The Bell Journal of Economics*, 1977, *8* (2), 517–524. Publisher: [RAND Corporation, Wiley].

Yang, Frank, "Costly Multidimensional Screening," August 2022.

Yee, Christine A., Kyle Barr, Taeko Minegishi, Austin Frakt, and Steven D. Pizer, "Provider supply and access to primary care," *Health Economics*, 2022. Publisher: Wiley Online Library.

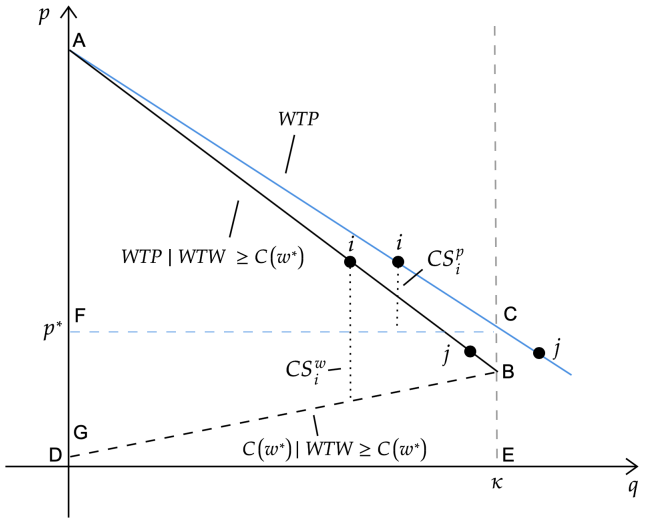
_ **, Yevgeniy Feyman, and Steven D. Pizer**, "Dually-enrolled patients choose providers with lower wait times: Budgetary implications for the VHA," *Health Services Research*, 2022. Publisher: Wiley Online Library.

Yee, Christine, Sivagaminathan Palani, Kyle Barr, and Steven D. Pizer, "Provider Supply and Access to Specialty Care," December 2022.

Zeckhauser, Richard, "Strategic sorting: the role of ordeals in health care," *Economics & Philosophy*, March 2021, *37* (1), 64–81. Publisher: Cambridge University Press.

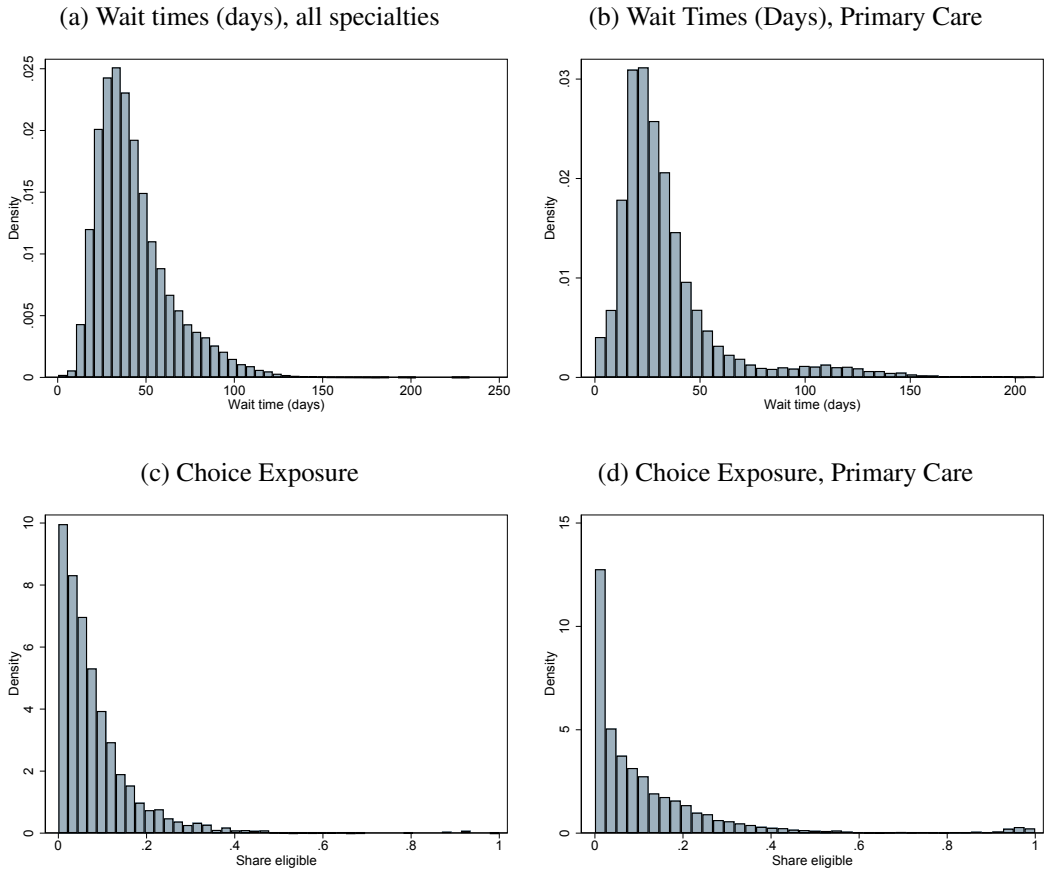
2.9 Figures and Tables

Figure 2.1: Conceptual framework: graphical illustration



Notes: Figure presents a graphical illustration of the efficiency effects and opportunities for redistribution of surplus under a price-based mechanism and a wait-time-based mechanism. The area between curves AC and AB represents the allocative efficiency loss from rationing by waiting. The curve GB calculates the average cost of waiting at each point along the curve AB; the area underneath the curve GB represents the deadweight loss. Example individual j switches from obtaining no care under the price regime to obtaining care under the wait time regime. Example individual i obtains care under both but achieves different levels of consumer surplus. κ represents available capacity, taken as given. These curves are simply examples for illustration: the curve $C(w^*)$ could be upward sloping, flat, or non-monotonic, with alternative distributions of preferences, and the curves AC and AB could lie closer or further from each other.

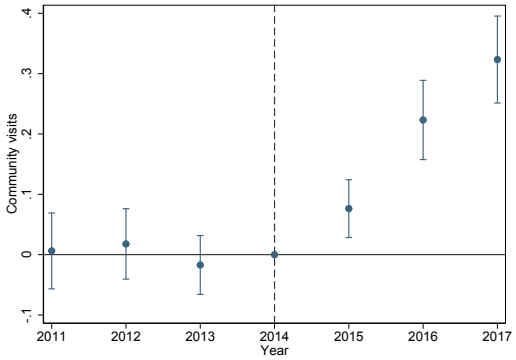
Figure 2.2: Clinic-level Summary Statistics



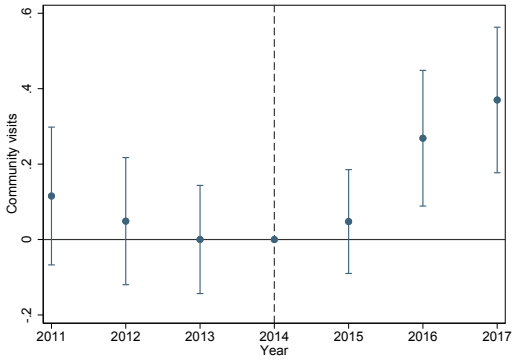
Notes: Figure presents histograms of average wait times and Choice exposure for all outpatient specialties and primary care, both calculated in the pre-Choice (2010-2014) period. Wait times are calculated as the average time between the request date and visit date among all appointments (completed or cancelled) in a clinic in a quarter. Exposure is calculated as the share of visits in a clinic and specialty for which the patient is distance eligible (lives over 40 miles from their closest clinic or in a state without a VA hospital) or wait-time eligible (has a wait time, based on the patient's desired or clinically indicated date, of over 30 days). Sample includes 1,128 clinics.

Figure 2.3: Effect of Choice Eligibility on Utilization

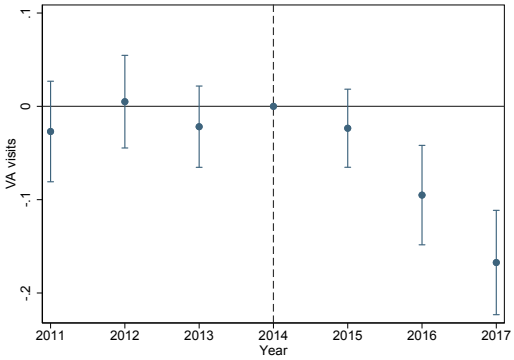
(a) VA-Financed Community Visits



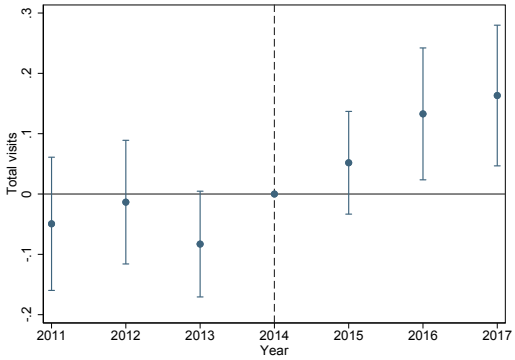
(b) All Community Visits



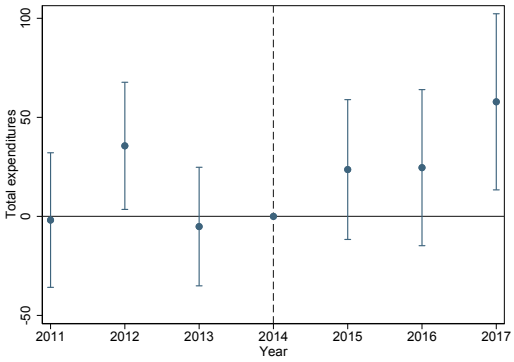
(c) VA Visits



(d) Total Visits (VA + Community)



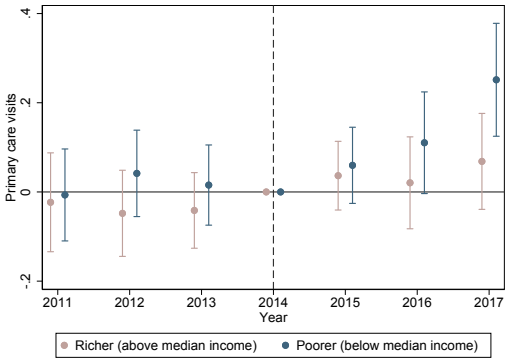
(e) Total VA Spending



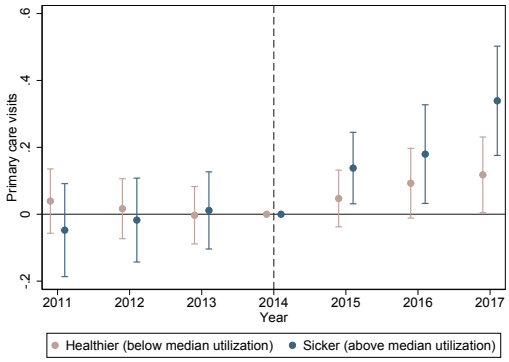
Notes: Figure presents event study coefficient estimates from Equation 2.4. In sub-figure (a) the outcome is all community outpatient visits per year, estimated on the whole sample. In sub-figure (b), I restrict to the population of veterans dually eligible for TM, where I observe the complete universe of utilization. Sub-figure (c), (d), and (e) plot VA visits, total visits (VA + community), and total VA spending in the whole sample, respectively. Total VA spending is calculated based on per-visit cost estimates at the VA from HERC and claims for VA-financed community care. Estimates restricted to a sample living 10 miles from the 40 mile eligibility threshold.

Figure 2.4: Heterogeneity on Responses to p

(a) Community PC visits, by income



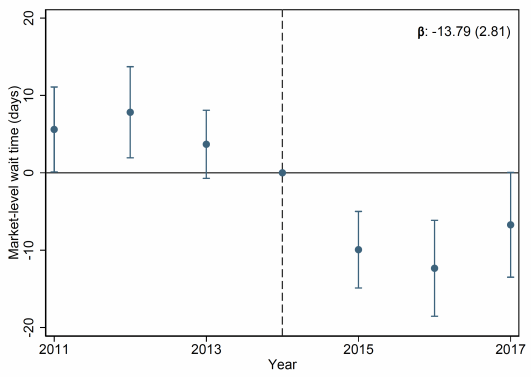
(b) Community PC visits, by health status



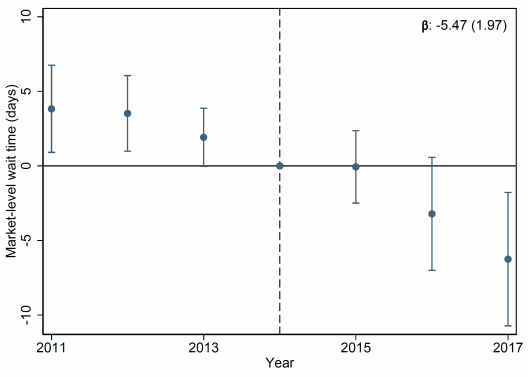
Notes: Figure presents event study coefficient estimates from Equation 2.4 with the outcome equal to community primary care visits, split by whether a veteran is above versus below median income (a) and health status (b), where health status is measured by the total VA spending in the prior year. The sample is restricted to a sub-set of veterans who face the same prices: TMs without a Medigap plan choosing primary care. Sample restricted to enrollees living within a 10 mile window of the 40 mile threshold.

Figure 2.5: Effect of Choice Exposure on Wait Times

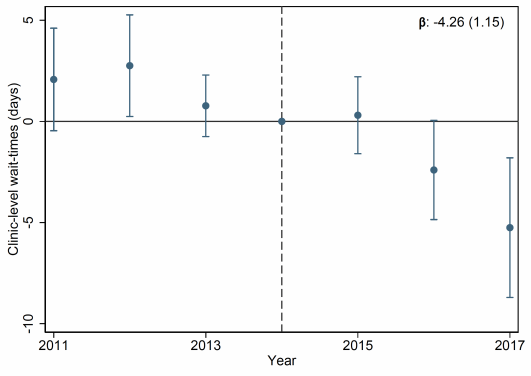
(a) HRR-by-specialty level



(b) HRR-level, primary care

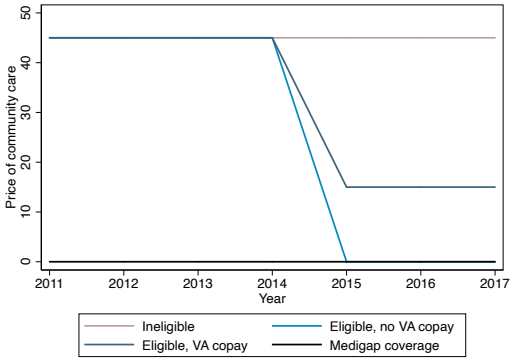


(c) Clinic-level, primary care



Notes: Figure presents event study coefficient estimates from Equation 2.6 with pooled difference in-difference estimates (from Equation 2.7) presented on the graph with standard errors clustered at the market-by-specialty (a), market (b), or clinic (c) level). Sub-figure (a) calculates effects on wait times across all markets and specialties, sub-figure (b) calculates effects on wait times at the geographic market level (HRR) for primary care only, and sub-figure (c) calculates effects on wait times at the clinic-level. Coefficients are scaled to represent a move from the 10th to the 90th percentile in the pre-period share eligible distribution.

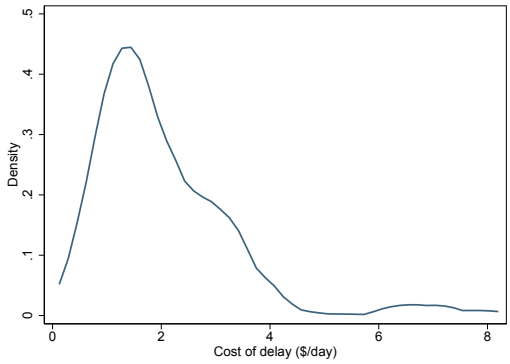
Figure 2.6: Variation in p



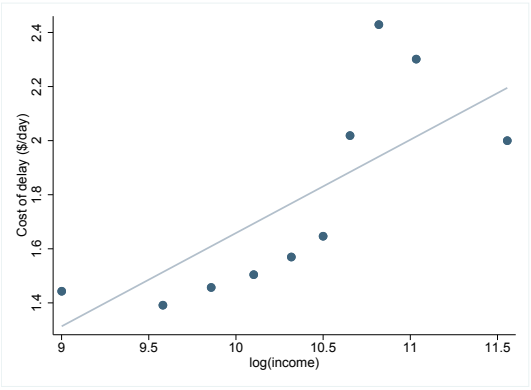
Notes: Figure presents variation in prices for community care among the different out of pocket payment classes for the TM sample. The out of pocket price for TMs without Medigap is calculated as the median out of pocket payment for primary care. Medigap prices are set to zero because the vast majority of Medigap policies held by VA enrollees include essentially no outpatient cost-sharing. I observe veterans' Medigap status because they report this information to the VA.

Figure 2.7: Distribution of the cost of delay, $C_i(w) = \frac{\gamma_i}{\alpha_i}$

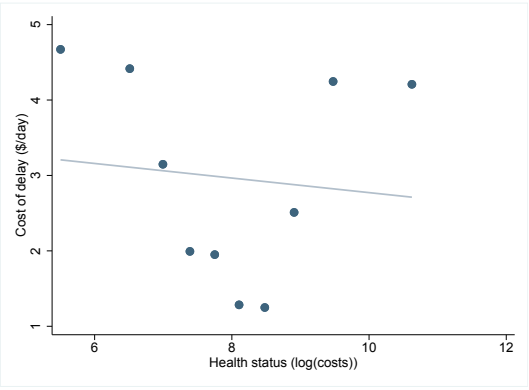
(a) Kernel density plot



(b) Correlation with income



(c) Correlation with health status (costs)



Notes: Figures summarize estimates of $\frac{\gamma_i}{\alpha_i}$ from the two-step estimation procedure described in Section 2.5.4 with variation in prices coming from the Choice Act and waiting time instrumented using deciles of Choice exposure interacted with a post Choice indicator. Sub-figure (a) plots a kernel density plot of $\frac{\gamma_i}{\alpha_i}$, or the cost of delay. Sub-figures (b) and (c) correlate $\frac{\gamma_i}{\alpha_i}$ with log income and lagged utilization (a proxy for health status) in binned scatterplots. I exclude veterans who do not engage with the VA at all in the preceding year.

Table 2.1: Enrollee-level Summary Statistics

	All	Dually-eligible for TM	Distance-eligible for Choice	Not distance-eligible for Choice
	(1)	(2)	(3)	(4)
Panel A. Veteran characteristics				
Age	61.9	73.2	63.9	61.7
	(17.1)	(10.7)	(16.2)	(17.2)
Share male	0.927	0.970	0.944	0.925
Income	36,862	39,070	35,000	37,033
	(29,638)	(29,138)	(27,294)	(29,838)
Share paying copays	0.293	0.348	0.289	0.294
Share with any supplemental insurance	0.758	1.000	0.779	0.756
Share with Traditional Medicare	0.361	1.000	0.452	0.353
Panel B. Annual utilization				
VA spending	5,148	6,207	4,714	5,188
	(18,726)	(22,007)	(16,073)	(18,945)
VA outpatient visits	7.54	8.46	6.25	7.66
	(13.96)	(14.54)	(11.12)	(14.20)
Total outpatient visits	12.15	19.17	12.03	12.16
	(24.02)	(30.82)	(22.39)	(24.17)
VA primary care visits	1.02	1.15	1.03	1.02
	(1.61)	(1.70)	(1.55)	(1.62)
Total primary care visits	2.31	4.38	2.49	2.29
	(4.89)	(6.29)	(5.12)	(4.86)
Share with both VA and non-VA primary care visit				
In same year	0.114	0.244	0.149	0.110
Ever	0.311	0.512	0.403	0.303
N enrollees	11,329,529	5,231,539	1,274,957	10,619,494
N enrollee-years	64,944,118	23,458,147	5,420,656	58,756,964

Notes: Table presents summary statistics from the veteran enrollee sample from 2011-2017. Column (1) includes all veterans, column (2) includes only veterans who are dually eligible for Traditional Medicare, for whom I observe all of their community utilization, column (3) includes only veterans whose home address is over forty miles from the closest VA clinic, and column (4) includes veterans who live closer than forty miles from their VA clinic. Income is averaged across all means-tests conducted from 2000-2019 and presented in 2015 USD. VA spending tabulates all spending by the VA per enrollee-year. Outpatient visits and primary care visits tabulate the number of outpatient and primary encounters per enrollee-year at the VA, and across VA and community care (total).

Table 2.2: Coefficient Estimates: Direct Effects

	All		TM dual eligibles sample	
	2014 mean	Coefficient estimate	2014 mean	Coefficient estimate
	(1)	(2)	(3)	(4)
Panel A. Utilization				
Community visits (all)	5.475	0.235	10.457	0.177
		(0.033)		(0.067)
Community spending (VA-financed)	302.25	69.85	346.77	67.37
		(5.52)		(9.40)
Visits at VA clinics	6.181	-0.082	6.861	-0.086
		(0.020)		(0.031)
RVUs as VA clinics	1779.26	-26.92	1981.03	-20.71
		(8.50)		(14.42)
Spending at VA clinics	2440.32	-42.03	2696.07	-32.99
		(11.51)		(18.36)
Total visits (VA + community)	11.69	0.150	17.32	0.091
		(0.041)		(0.076)
Total VA-financed spending	2742.57	27.82	3042.83	34.38
		(13.35)		(21.59)
Panel B. Clinic characteristics				
Wait time	38.53	-2.25	41.19	-2.65
		(0.10)		(0.15)
Distance (miles)	54.12	-1.28	54.36	-1.15
		(0.06)		(0.08)

Notes: Table presents coefficient estimates from Equation 2.5, for the whole sample (column (2)) and for the TM sample (column (4)) for whom the universe of community utilization is observed. Columns (1) and (3) present the year -1 (2014) mean in each sample for interpretation. Community visits (all) indicate all visits at non-VA providers across VA and Medicare financing. Community spending (VA-financed) indicates community spending that is VA (not Medicare) financed. RVUs at VA clinics is a measure of utilization in which procedures are weighted identically to Medicare. VA spending is attributed to specific visits from accounting data by HERC. Total visits (VA + all community) captures all visits at VA and non-VA providers across VA and Medicare financing. Total VA financed spending include all VA spending across VA clinics and community care. Wait time and drive time indicate the average wait time and drive time patients experience conditional on receiving any care, across VA and community options, where community wait times are calculated based on the time between authorization and visit, and VA wait times are calculated as described in the main text. Robust standard errors in parenthesis clustered at the enrollee level. All utilization outcomes are for outpatient care.

Table 2.3: Heterogeneity in Responses to w

		By income		By health status	
	All	Poorer (below medium)	Richer (above medium)	Sicker (above median utilization)	Healthier (below median utilization)
	(1)	(2)	(3)	(4)	(6)
Coefficient estimate	-0.008	-0.010	-0.008	-0.012	-0.008
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)

Notes: Table presents coefficient estimates the clinic-level 2SLS specification in Equation 2.8 for everyone, and split by below vs above income, and below versus above median prior VA utilization (as a proxy for health status). The dependent variable is the log of the market share of each clinic among Choice-ineligible veterans in a given demographic group. Market shares calculated using HRRs as market definitions. Instruments are deciles of choice exposure interacted with a post indicator. First stage F-statistic = 16. Regressions are weighted by the market size of each demographic group across HRRs. Robust standard errors in parenthesis.

Table 2.4: Parameter Estimates of γ and α: No Heterogeneity

	$\delta_{jmt} = \ln(s_{jmt}) - \ln(s_{0jmt})$		
	OLS	IV deciles exposure	IV continuous exposure
	(1)	(2)	(3)
γ : coefficient on wait time	0.0004	-0.0198	-0.0194
	(0.0003)	(0.0021)	(0.0052)
α : coefficient on price	-0.0079	-0.0079	-0.0079
	(0.0004)	(0.0004)	(0.0004)
$\frac{\gamma}{\alpha}$: cost of delay ($/day)	-0.05	2.52	2.47
	(0.037)	(0.23)	(0.66)
Clinic fixed effects (ϕ_j)	✓	✓	✓
Time fixed effects for VA and community care (χ_t^{VA}, χ_t^C)	✓	✓	✓
Average elasticity w.r.t. w	0.017	-0.932	-0.913
Average elasticity w.r.t. p	-0.296	-0.294	-0.294

Notes: Table presents parameter estimates from a simplified version of Equation 2.10, $\ln(s_{jmt}) - \ln(s_{0jmt}) = \psi_o^{VA} + \psi_o^C + \chi_t^{VA} + \chi_t^C + \phi_j + \gamma_0 w_{jt} + \alpha_0 p_{jmt} + \xi_{jmt}$, with variation in prices coming from the Choice Act. In column (1), I use all residual variation variation in wait times. In column (2), I instrument for wait time using interaction between deciles of exposure share and a post indicator (first stage F-statistic = 43). In column (3), I instead use a continuous interaction (first stage F-statistic = 69). Figure 2.5 provides a visual representation of the the first stage. Elasticities are calculated for each clinic with non-zero prices or wait times. Robust standard errors in parenthesis. Standard errors on the cost of delay calculated via the delta method.

Table 2.5: Counterfactuals: Status Quo Waiting Regime vs. Market-Clearing Prices

	Status quo w-regime (1)	Price regime (2)
Panel A. Prices and wait times		
Median clinic wait (VA)	40.84	0
Median copay charged (VA)	5.04	78.41
as a share of cost of service	0.01	0.20
Panel B. Changes in allocations		
Obtained care in quarter	0.42	0.41
Share displaced relative to s.q.		
out of VA care		0.16
out of any care		0.04
Panel C. Characteristics of veterans receiving care		
VA care		
Log income	10.31	10.35
Log health costs	8.36	8.32
Age	71.49	70.75
Any care		
Log income	10.63	10.64
Log health costs	7.91	7.90
Age	74.75	74.56

Notes: Table presents status quo (column (1)) and counterfactual (column (2)) prices, wait times, and allocations spanning the pre-Choice and post-Choice period. In column (2), I search for a vector of VA prices (uniform at each clinic in each quarter) that reach an equilibrium (given by Equation 2.11) with no wait times for the same period, keeping total VA utilization constant. Wait times and prices vary across clinics and time based on excess demand: this table reports the median. Panel C reports the average characteristics of consumers served at the VA, or at all, including VA and community care, under the two regimes.

Table 2.6: Welfare Effects of Waiting Regime vs. Market-Clearing Prices

	Δ: status quo (waiting) - price regime		
	All	$\Delta CS > 0$	$\Delta CS < 0$
	(1)	(2)	(3)
Δ consumer surplus ($ per veteran, per year)	-57.25	18.37	-150.46
From change in allocation	-34.28		
From change in payoff | allocation	-22.98		
Revenue	-62.90		
Total difference in welfare	-112.95		
Δ as a share of total achievable surplus (%)	24		
Pop share		0.55	0.45
Log income		10.23	10.58
Log health util		8.05	7.89
Age		73.77	71.67

Notes: Table presents changes in welfare per enrollee, per year from the status quo rationing regime, relative to a counterfactual equilibrium in which VA care can be obtained immediately with no wait, at a price per clinic that is uniform across veterans in a market and time period (the counterfactual presented in column (2) of Table 2.5). Column (1) presents changes in consumer surplus, decomposed into changes in allocative efficiency and changes in payoff conditional on allocation, and revenue. Columns (2) and (3) split the sample into those who prefer the status quo wait-time rationed regime (column (2)) and those who prefer a price regime (column (3)) and presents consumer surplus and characteristics among those two groups. Consumer surplus calculated using the log-sum formula (Train, 2009).

Table 2.7: Evaluating the Efficiency-Redistribution Trade-off

	Share of revenue redistributed				
	0	0.25	0.5	0.75	1
Share better off under waiting regime after transfer	0.55	0.21	0.10	0.05	0.03

Notes: Table presents the share of veterans who are better off under the status quo waiting regime after accounting for redistributed revenue, modulating the extent of redistribution between zero (all revenue is burnt) and one, where all revenue is uniformly redistributed lump-sum. Share better off determined by consumer surplus calculated using the log-sum formula (Train, 2009).

Table 2.8: The Performance of Policy Instruments

	Price mech.	Choice	Expand to MISSION	Expand to everyone	Copay ↑ from $15 to $50
	(1)	(2)	(3)	(4)	(5)
All differences are relative to the status quo					
Δ CS ($ per veteran, per year)	57.25	1.93	4.28	7.63	11.50
Δ Revenue - Costs	62.90	-3.22	-7.61	-17.59	2.52
Δ Welfare	122.95	-1.84	-4.72	-6.56	14.40

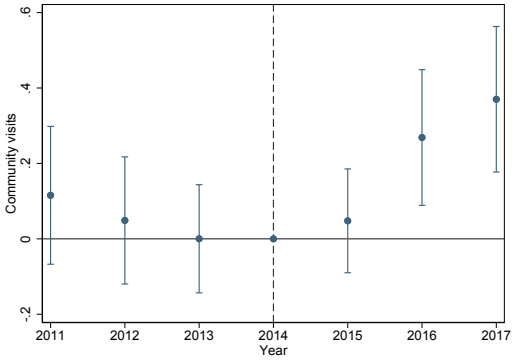
Notes: Table presents consumer surplus, government revenue net of costs, and total welfare (adding CS and revenue) under the price mechanism (column (1), re-created from Table 2.6) and factual and counterfactual policies, relative to the status quo. Column (2) presents the welfare effects of the Choice Act, column (3) considers the expansion in eligibility requirements under the MISSION Act of 2018 (115th Congress, 2018), and column (4) subsidizes everyone. Column (5) increases copayments from $15 to $50 (undoing a policy change in 2001 that reduced primary care copayments from just over $50 to $15) for the 25% of veterans who are already obligated to pay copayments. Consumer surplus calculated using the log-sum formula (Train, 2009).

2.10 Appendix

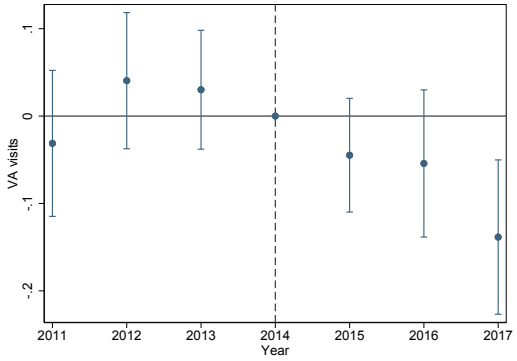
2.10.1 Additional Analyses of the Choice Act

Figure B.1: Effect of Choice Eligibility on Utilization, TM Sample

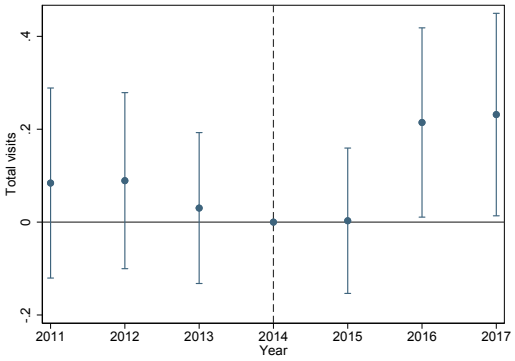
(a) All community visits



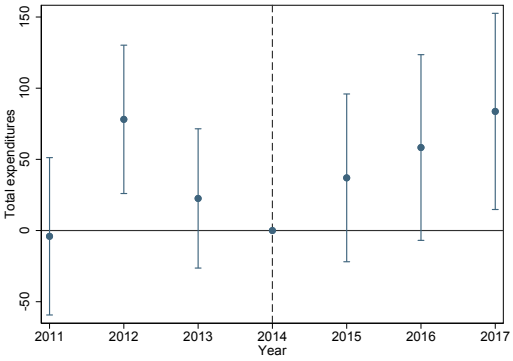
(b) VA visits



(c) Total Visits (VA + community)

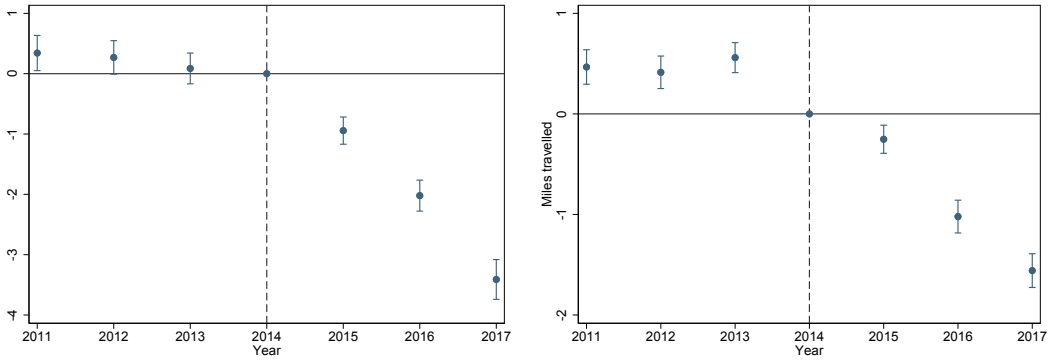


(d) Total VA spending



Notes: Notes: Figure presents event study coefficient estimates from Equation 2.4, estimated on on the TM sample. In sub-figure (a) the outcome is all community outpatient visits per year. Sub-figure (b), (c), and (d) plot VA visits, total visits (VA + community), and total VA spending. Total VA spending is calculated based on per-visit cost estimates at the VA from HERC and claims for VA-financed community care. Estimates restricted to a sample living 10 miles from the 40 mile eligibility threshold.

Figure B.2: Effects of Choice on Characteristics of Chosen Clinic

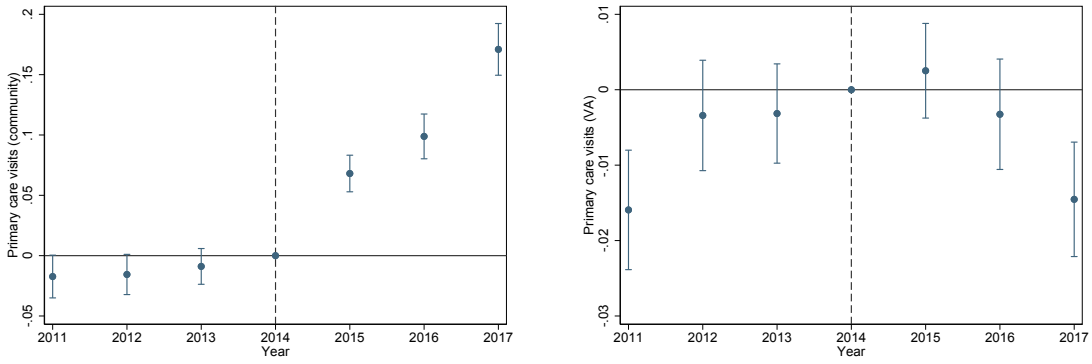


(a) Wait times of chosen clinics (b) Distance to chosen clinics

Notes: Figures present event study coefficients from Equation 2.4 with the outcomes equal to the average wait time (a) and travel distance (b) at the VA and across VA and community care. Wait times at community care are determined by the time between the authorization for community care and the date of the visit. Wait times for VA visits are calculated as described in the main text. Effects estimated on the whole sample of enrollees living 10 miles from the 40 mile eligibility threshold.

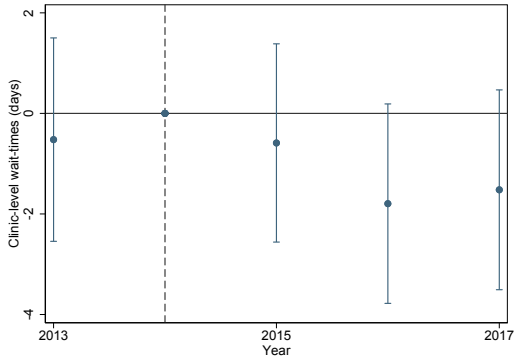
Figure B.3: Effect of Choice Eligibility on Primary Care Utilization

(a) Community visits: primary care (b) VA visits: primary care



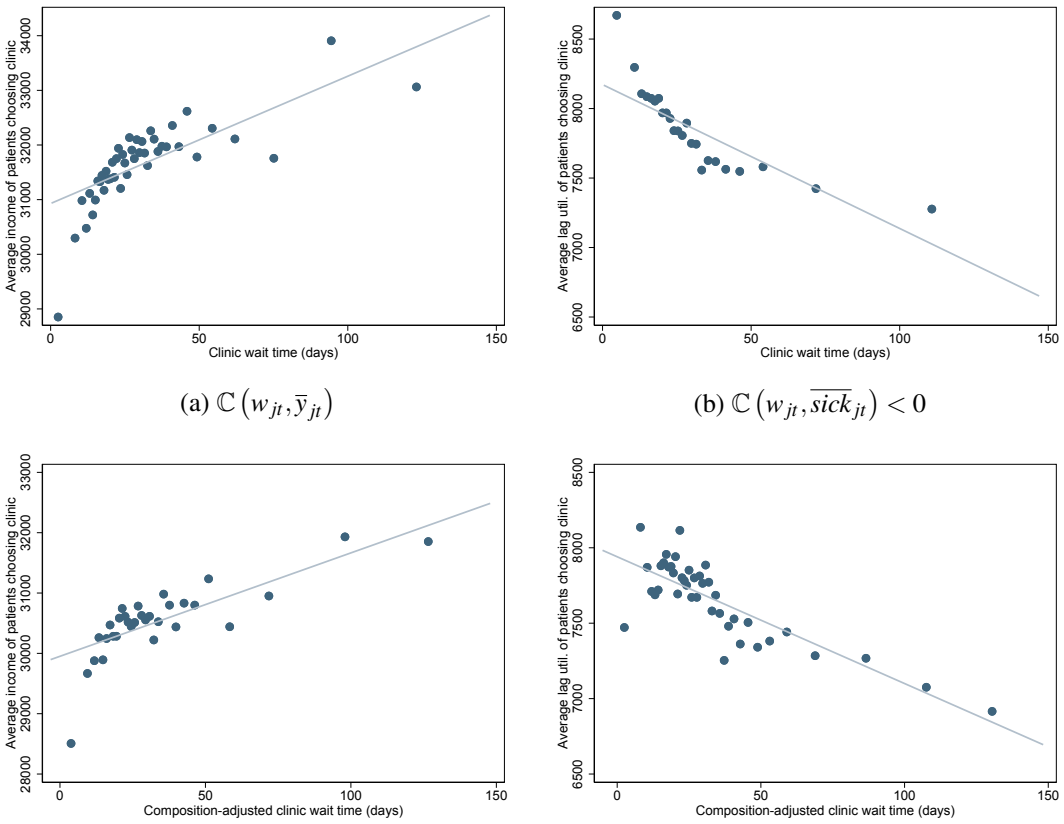
Notes: Figure presents event study coefficient estimates from Equation 2.4 for primary care specifically. In sub-figure (a) the outcome is all community outpatient visits per year, and in sub-figure (b) I plot VA visits. Effects estimated on the whole sample of enrollees living 10 miles from the 40 mile eligibility threshold.

Figure B.4: Equilibrium Effects: New Patient Wait Times



Notes: Figure replicates Figure 2.5c, calculating wait times only for new patients. The figure excludes 2011 and 2012 because patients are defined as new based on having no visits in two-year look back period.

Figure B.5: Screening Effects of w: Full panel



(a) $\mathbb{C}\left(w_{jt}, \bar{y}_{jt}\right)$

(b) $\mathbb{C}\left(w_{jt}, \overline{sick}_{jt}\right) < 0$

(c) $\mathbb{C}\left(\tilde{w}_{jt}, \bar{y}_{jt}\right) > 0$: residualized wait time

(d) $\mathbb{C}\left(\tilde{w}_{jt}, \overline{sick}_{jt}\right) < 0$: residualized wait time

Notes: Figure plots the covariance between the wait times at a clinic and the characteristics of patients seeking visits at that clinic, residualized of clinic and market-by-time fixed effects. Figures (a) and (b) use the raw average wait time, while figures (c) and (d) first residualize the wait times for each visit on the patient characteristics for a given visit and take the average of that residualized measure. The positive covariance between wait times and income indicates that at higher wait times, lower income veterans are less likely to choose a given clinic. The negative relationship between prior utilization and wait times indicates that at higher wait times, sicker (as proxied by lower VA spending) veterans are less likely to choose a given clinic. These patterns are consistent with the patterns documented using only the Choice Act variation in Section 2.4.3.

Table B.1: Effects on Mortality and Inpatient Admissions

	Direct effects 10-mile window (1)	Equilibrium effects Market-level, inelig. only (2)
Log(mortality)	-0.012	0.361
	(0.026)	(0.199)
Log(inpatient admits)	-0.031	-0.294
	(-0.030)	(0.185)
Year FEs	✓	✓
Closest PC-site FEs	✓	
Market FEs		✓

Notes: Table presents results from Equation 2.5 (column (1)), aggregating observations to the closest-VA clinic by year level. The outcome is equal to log mortality and log inpatient admissions. Column (2) estimates Equation 2.7, with the outcome equal to log mortality and log inpatient admissions, among only ineligibles to investigate equilibrium effects on health.

Table B.2: Direct effect of Choice: Robustness

	All		TMs	
	Baseline	Include closest clinic x year effects	Baseline	Include closest clinic x year effects
	(1)	(2)	(3)	(4)
Panel A. Utilization				
Community visits (all)	0.235	0.149	0.177	0.108
	(0.033)	(0.036)	(0.067)	(0.072)
Community spending (VA -financed)	69.85	55.77	67.37	50.63
	(5.52)	(5.71)	(9.40)	(9.73)
Visits at VA clinics	-0.082	-0.097	-0.086	-0.070
	(0.020)	(0.022)	(0.031)	(0.033)
RVUs as VA clinics	-26.92	-41.29	-20.71	-31.43
	(8.50)	(9.05)	(14.42)	(15.27)
Spending at VA clinics	-42.03	-47.50	-32.99	-29.02
	(11.51)	(12.48)	(18.36)	(19.88)
Total visits (VA + community)	0.150	0.074	0.091	0.037
	(0.041)	(0.045)	(0.076)	(0.081)
Total VA-financed spending	27.82	8.27	34.38	21.62
	(13.35)	(14.37)	(21.59)	(23.18)
Panel B. Clinic characteristics				
Wait time	-2.25	-0.75	-2.65	-0.81
	(0.10)	(0.09)	(0.15)	(0.15)
Distance (miles)	-1.28	-0.84	-1.15	-0.77
	(0.06)	(0.05)	(0.08)	(0.08)

Notes: Table presents coefficient estimates from Equation 2.5, for the whole sample (columns (1) and (2)) and for the TM sample (columns (3) and (4)) for whom the universe of community utilization is observed. Table presents robustness to including closest clinic by year fixed effects (columns (2) and (4)). Community visits (all) indicate all visits at non-VA providers across VA and Medicare financing. Community spending (VA-financed) indicates community spending that is VA (not Medicare) financed. RVUs at VA clinics is a measure of utilization in which procedures are weighted identically to Medicare. VA spending is attributed to specific visits from accounting data by HERC. Total visits (VA + all community) captures all visits at VA and non-VA providers across VA and Medicare financing. Total VA financed spending include all VA spending across VA clinics and community care. Wait time and drive time indicate the average wait time and drive time patients experience conditional on receiving any care, across VA and community options, where community wait times are calculated based on the time between authorization and visit, and VA wait times are calculated as described in the main text. Robust standard errors in parenthesis clustered at the enrollee level.

Wait time eligibility analysis Unlike distance eligibility, which impacts all care for all time periods in the post period, wait time eligibility is determined based on endogenous market conditions, making it harder to analyze cleanly. Despite this, I use the variation from the wait time eligibility condition under Choice with the following specification:

$$y_{gst} = \beta \mathbf{1}\{WaitElig_{gst}\} + \theta_{gs} + \tau_t + \varepsilon_{gst} \tag{2.14}$$

at the geography g (HRR), specialty s, and quarter t level. $WaitElig_{gst}$ is an indicator for whether a given geography, specialty, and time is wait time eligible, θ_{gs} capture specialty by geography fixed effects, and τ_t capture quarter fixed effects. This specification captures differences in outcomes y_{gst} in time periods and markets in which patients are wait-time eligible, relative to overall time trends and levels in the market. My two primary outcomes are interest are VA authorizations for community care, internal VA documentation that indicates that a patient is allowed to obtain community care, and VA utilization in a given specialty, market, and quarter. I focus on authorizations instead of claims because authorizations are more amenable to accurately capturing specialty categorizations, as both authorizations and VA utilization use the same method of specialty classification. This classification differs from information contained in the claims data. Authorizations understate the extent of utilization relative to claims because authorizations cover a period of time in which multiple visits may occur. Community utilization in Table B.3 is therefore not comparable to Table 2.2 in magnitudes. The primary purpose of this exercise is to examine qualitative patterns.

Table B.3 documents that community authorizations do indeed increase in wait time eligible quarters, but that this is not accompanied by a reduction in VA visits. This is consistent with the hypothesis that the VA is capacity constrained, as the ability to substitute to community care does not reduce overall utilization at VA clinics.

Table B.3: Wait Time Eligibility Results

	Coefficient estimate
Community care authorizations / quarter	0.0010
	(0.0000)
VA visits / quarter	0.0004
	(0.0002)

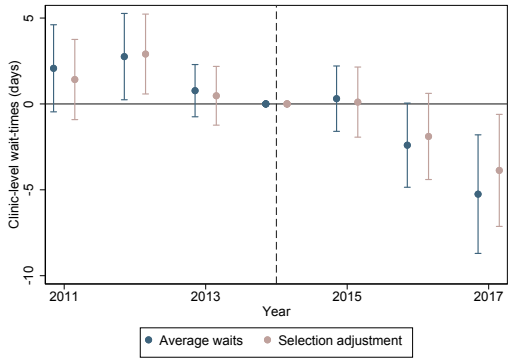
Notes: Figure presents results from Equation 2.14 at the specialty, geography, and quarter level. Community care is measured in authorizations, not visits. Authorizations generally encompass multiple visits. Robust standard errors in parenthesis.

2.10.2 Model Details

Generating an unselected distribution of w_{jt} To generate an unselected distribution of waiting times w_{jt}, I zoom in to the smallest unit of analysis to be as conservative as possible: the doctor by day level. For every day and at every physician where I do not observe an appointments made, I assume that the appointment date of the first subsequent appointment with that physician was available to any arriving patient at the initial index date. This is consistent with a First-Come-First-Served protocol, but would be violated if physicians tailored wait times to specific patients.

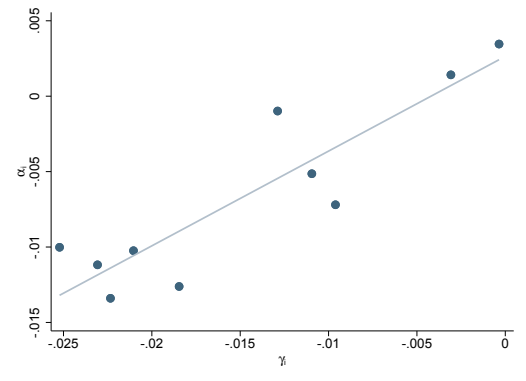
If a physician does not have any subsequent appointments for 90 days, I consider that physician unavailable. I then take the average all "augmented" wait times across physicians and days within a clinic to obtain the unselected distribution of wait times. Figure B.6 re-creates Figure 2.5c for both the raw mean and the adjusted "unselected" distribution of wait times and shows similar results.

Figure B.6: Effects of Choice exposure on the raw mean and unselected wait times



Notes: Figure plots coefficients from Equation 2.6 using both raw averages and the constructed unselected distribution of wait times, described above.

Figure B.7: Correlation Between γ_i and α_i



Notes: Table presents binned scatterplot of the estimated relationship between γ_i (cost of waiting) and α_i (price sensitivity).

Table B.4: Parameter Estimates: Heterogeneity

	Prefs for VA	Prefs for community	Distance	Wait times	Prices
	(1)	(2)	(3)	(4)	(5)
Age bin 2	0.144	0.523	-0.011	0.002	0.002
	(0.223)	(0.161)	(0.007)	(0.003)	(0.005)
Age bin 3	0.336	0.786	-0.009	0.003	-0.008
	(0.206)	(0.149)	(0.007)	(0.003)	(0.005)
Age bin 4	0.433	1.172	-0.014	0.003	-0.006
	(0.198)	(0.143)	(0.006)	(0.003)	(0.005)
Income bin 2	0.127	0.278	-0.010	0.001	0.002
	(0.145)	(0.089)	(0.005)	(0.003)	(0.003)
Income bin 3	0.026	0.489	-0.013	0.005	0.004
	(0.174)	(0.095)	(0.006)	(0.003)	(0.004)
Income bin 4	-0.065	0.407	-0.015	0.004	0.007
	(0.201)	(0.099)	(0.007)	(0.004)	(0.004)
Income bin 5	-0.339	0.351	-0.019	0.004	0.003
	(0.218)	(0.101)	(0.007)	(0.004)	(0.004)
Missing income	-0.209	0.421	-0.022	0.003	0.000
	(0.119)	(0.071)	(0.004)	(0.002)	(0.003)
Prior util. bin 2	0.367	-0.260	0.008	-0.000	-0.005
	(0.083)	(0.042)	(0.003)	(0.001)	(0.002)
Prior util. bin 3	0.430	-0.522	0.018	-0.000	-0.012
	(0.082)	(0.044)	(0.003)	(0.001)	(0.002)
Prior util. bin 4	0.499	-0.824	0.030	-0.002	-0.006
	(0.084)	(0.047)	(0.003)	(0.001)	(0.002)
No prior VA util.	-2.311	-0.028	-0.006	0.009	0.008
	(0.128)	(0.040)	(0.003)	(0.002)	(0.001)
New enrollee	-0.727	-0.216	0.006	0.008	0.012
	(0.172)	(0.079)	(0.005)	(0.003)	(0.003)
Priority 2	-0.271	-0.118	-0.000	0.002	0.001
	(0.100)	(0.051)	(0.003)	(0.001)	(0.002)
Priority 3	-0.143	-0.092	-0.007	0.001	-0.004
	(0.085)	(0.043)	(0.003)	(0.001)	(0.002)
Priority 4	-0.550	-0.157	0.002	-0.001	-0.004
	(0.128)	(0.066)	(0.004)	(0.002)	(0.003)
Priority 5	-0.402	-0.242	-0.006	0.003	-0.009
	(0.086)	(0.044)	(0.003)	(0.001)	(0.002)
Priority 6	-0.259	-0.117	-0.010	0.003	-0.009
	(0.148)	(0.069)	(0.005)	(0.002)	(0.002)

Notes: Table presents parameter estimates of heterogeneity parameters. Low age bins imply low ages. Low income bins imply low income. Low lagged utilization bins imply low utilization.

Chapter 3

The Effects of Floodplain Regulation on Housing Markets

3.1 Introduction

Floods cause an estimated $32 billion in damages per year, making them the costliest form of natural disaster in the United States (Wing et al., 2022). These losses are expected to grow by 25% by 2050, reflecting both an increase in flood hazard due to climate change and a concentration of population growth in risky locations (Wing et al., 2022). A central concern among policymakers and economists is that these trends are driven by a wedge between social and private values for flood safety (Kydland and Prescott, 1977; Coate, 1995; Ben-Shahar and Logue, 2016). Indeed, a large body of work has documented that private incentives to reduce flood risk are muted by both misperceptions of that risk and expectations of government aid (Gallagher, 2014; Kousky et al., 2018a; Davlasheridze and Miao, 2019; Mulder, 2021; Bakkensen and Barrage, 2022; Landry et al., 2021a; Wagner, 2022; Hsiao, 2023). To correct these market frictions, policymakers demarcate especially risky locations as "Special Flood Hazard Areas" (SFHAs) and regulate them more strictly. Inside the SFHA, developers are required to build elevated homes, and homeowners face a flood insurance purchase mandate and higher flood insurance prices.[1]

This paper investigates the impact of floodplain regulation on the location of new construction, housing prices, estimated flood damages, and social welfare. The effect of this coarse policy on welfare is ambiguous, as it might not reduce damages more than the costs it imposes via distortions

[1]The policy instrument of creating a binary distinction of "floodplain" or not and imposing both insurance and building requirements is not unique to the United States. EU countries and Australia also manage flood risk via the creation of flood maps that influence both flood insurance and building codes (de Moel et al., 2009; Golnaraghi et al., 2020).

in the housing market. In this paper, we study the extent to which floodplain regulation reduces flood damages via both the location and flood-safe adaptation of construction, and we weigh these benefits against the costs of that regulation. We do so by studying effects around the boundaries of the regulated area and around the date of the policy's introduction. We then embed these empirical results in a model of residential choices and construction to estimate the market-wide effects of floodplain regulation and investigate the welfare impacts of current and counterfactual policies.

To conduct our analysis, we assemble a new comprehensive and spatially-granular dataset to describe regulation, real estate development, and flood risk in Florida over a 40-year time horizon. Florida is both populous and flood-prone, with 45% of land currently designated as a high-risk flood zone. Our dataset combines maps of historic flood zone extents, granular remote-sensing-based measures of historic and current development, administrative data describing house prices and attributes, and the flood risk profile of current and counterfactual development, generated from a hydrological model. Importantly, this dataset extends to the first maps delineating regulated areas, which we digitized from archival scans. This allows us to study the policy's effect on long-run development and confirm the validity of our empirical approach. And because our dataset details both adaptation status and location decisions, as well as how flood damages vary along these margins, we are able to comprehensively measure the policy's effect on flood risk.

We use two complementary empirical strategies — a spatial regression discontinuity and an event study around the regulation's implementation — to characterize the policy's risk reduction effects along two margins: reduced construction and mandatory adaptation in risky areas. Our spatial regression discontinuity design compares current development and house prices on either side of the regulatory boundary delineated at the time of the policy's introduction in the 1970s and 1980s. Our analysis relies on the assumption that flood risk and other amenities are smooth through these initial regulatory boundaries. While unlikely to hold in modern maps, this assumption is reasonable for the original maps because mapping technologies were rudimentary and homeowners lacked the ability to influence the initial regulatory boundaries.[2] Importantly, we validate this assumption by demonstrating smoothness in pre-period land use through the historic boundaries. We document that the modern-day share of developed land is 9% lower just inside the regulated area, highlighting the potential for the policy to reduce damages by shifting construction out of risky areas. This decrease in new construction is not accompanied by a reduction in prices – if anything, prices are slightly higher just inside the regulated area – indicating that floodplain regulation imposes costs on developers, which they at least partially pass through to consumers.

We also document reduced damages on the intensive margin via building standards that impose

[2]After the initial maps were drawn, landowners could deregulate developed parts of their properties either by petitioning to correct a mistake or physically elevating land to reduce its risk. Since maps are updated over time, this behavior produces a negative correlation between development and floodplain designation in modern flood maps.

mandatory adaptation of houses built in the floodplain. Building on Wagner (2022), we exploit the sharp timing of the policy's introduction in an event study design. We document that building standards reduce average flood damages by 55%, or 3% of average home value. Though building standards generate social value via reduced damages, we also show that homeowners do not privately value this reduction in flood risk. This result is consistent with prior work documenting a large wedge between social and private valuations of flood risk, implying scope for welfare-improving intervention. It also suggests that the increase in price across the regulatory boundary is attributable to construction costs, not differences in willingness-to-pay for adapted homes.

Together, our empirical results yield four facts. First, floodplain regulation suppresses construction in high-risk areas. Second, the policy's adaptation mandate binds, reducing flood risk. Third, the cost of mandating adaptation is large enough that the price effect of the resulting inward shift of housing supply dominates at the boundary. Fourth, our setting exhibits a wedge between private and social valuations for flood safety, yielding the potential for the policy to improve social welfare. However, these results alone are insufficient to quantify the total effect of the policy on either damages or social welfare. The policy's effects depend on the location of counterfactual development: both risk and amenities vary across space. And due to the large size of the regulated area, house prices in unregulated locations may be affected, impacting relative incentives to develop across space and housing prices and consumer welfare market-wide. A complete welfare analysis of status quo or alternative policy designs requires quantifying both benefits via market-wide reductions in flood risk and costs via distortions in the housing market.

We therefore specify and estimate a model of residential choice and real estate development. In our model, individuals maximize utility when choosing census-tract-by-flood-zone locations, as a function of prices, floodplain status, and location characteristics including unobserved amenities. Developers build houses when doing so is more profitable than the outside option of land use; housing profits depend on housing prices and construction costs, which include a cost of compliance with flood-safe building codes. Our quasi-experimental results inform our model and estimation strategy. Because the event study estimates imply that homeowners are unwilling to pay for flood safety, developers of housing do not expend costs to adapt absent policy intervention. We estimate effects of floodplain designation on consumer choices and construction costs by matching the spatial discontinuity estimates around the regulatory boundaries.

We first use the model to quantify the policy's impact on expected flood damages. We find that the policy reduces expected flood damages by 62%, or approximately $3.5 billion per county. Both the extensive-margin location channel and intensive-margin adaptation channel are quantitatively important. The gains from mandatory flood-safe construction in regulated areas account for the majority, or 84%, of this reduction. The policy's incentive to build new houses in safer areas

contributes the remaining 16%. Risk reductions are driven by effects on both the demand for and supply of housing. Our parameter estimates imply that building standards increase construction costs by 24%. Consumers are willing to pay 27% more for an equivalent house to avoid living in a regulated area.

We conclude by developing a normative framework that allows us to estimate and compare social welfare under current and counterfactual policies. We define social welfare as the sum of consumer surplus, producer surplus, government revenue from each policy, and uninternalized flood damages, due either to internalities or externalities. Because the wedge between private and social valuation of flood risk could be due to misperceptions, our welfare framework allows consumers' decision and experienced utility to diverge (Allcott and Taubinsky, 2015). We contrast the *status quo* policy regime against both an unregulated benchmark and a first-best corrective tax in the spirit of Pigou (1920) equal to the social cost of flooding, which varies by location and adaptation status. This corrective tax is not only a useful theoretical benchmark, but a plausible extension of recent policy changes which have introduced more property-specific granularity in flood insurance premiums.[3]

The current policy achieves approximately the socially-efficient degree of damage reductions and improves social welfare substantially relative to an unregulated benchmark, an increase of $5,919 per newly-developed house. However, these social welfare gains are 27% lower than under a first-best corrective tax. The current policy imposes distortions in the housing market it attempts to correct, relying more on adaptation and less on relocation than is socially-optimal. In some regulated but relatively safe areas, mandating adaptation is costlier than its benefits, while in the riskiest locations, development is still inefficiently high despite strong disincentives to build. Motivated by this shortcoming, we propose a simple change to the current policy: improved targeting of regulated areas. By imposing regulation only in locations where the benefits of mandating adaptation exceed costs, fewer locations are regulated and fewer homes are adapted, but flood risk levels remain similar to under the status quo. When well-targeted, the simple binary regulation can achieve 94% of first-best social welfare gains, or $7,567 per newly-developed house.

Our work contributes to several literatures. Most directly, we contribute to a literature analyzing regulations designed to reduce damages from floods and other natural disasters, including effects on house prices (e.g. Hino and Burke (2021); see Beltrán et al. (2018) for a survey) and in-place adaptation (Mulder, 2021; Baylis and Boomhower, 2021a; Wagner, 2022). We build on this literature, first, by jointly studying the impacts of floodplain regulation on both the location and type of construction, which together determine flood damages.[4] Furthermore, we contribute by

[3] See The Congressional Research Service for an overview of this policy change and The National Flood Insurance Program for more details.

[4] This complements work studying the effects of related policies on population flows, including the introduction

developing an equilibrium framework to trade off the damage-reducing benefits and regulatory costs of current and counterfactual policy designs.

We also contribute to a literature documenting frictions in mitigating or adapting to climate risk (Annan and Schlenker, 2015; Deryugina and Kirwan, 2018; Kousky et al., 2018a; Mulder, 2021; Bakkensen and Barrage, 2022; Wagner, 2022; Baylis and Boomhower, 2023; Hsiao, 2023; Balboni, 2024). Beyond documenting frictions, we investigate the extent to which status quo second-best corrective policy can reduce resulting social welfare losses. In doing so, we relate to other work studying such policies, addressing internalities or externalities in other settings with coarse labels (e.g. Barahona et al. (2023)) and standards and attribute-based regulation (e.g. Ito and Sallee (2018); Kellogg (2020); Jacobsen et al. (2020, 2023)).

Methodologically, we relate to other work embedding boundary discontinuity designs in discrete choice frameworks (Bayer et al., 2007; Turner et al., 2014; Song, 2021; Anagol et al., 2023). Our model also incorporates recent estimates of location-specific supply elasticities from Baum-Snow and Han (2023) to characterize equilibrium changes in housing supply, as in Almagro et al. (2023).

The next section describes the institutional details of the National Flood Insurance Program, including the regulations imposed inside the SFHA and the process of generating flood maps that distinguish between SFHA and non-SFHA land. Section 3.3 describes our setting — the state of Florida — and data. In Section 3.4, we present quasi-experimental evidence of the causal effects of SFHA designation. We specify and estimate our equilibrium model of the housing market in Section 3.5. Section 3.6 simulates distributions of development and prices and discusses welfare under factual and counterfactual policies. In Section 3.7, we conclude.

3.2 Institutional Background

3.2.1 The National Flood Insurance Program and Special Flood Hazard Areas

Congress established the National Flood Insurance Program (NFIP) in 1968 in response to high flood losses and a perception that lackluster local regulation permitted excessive construction in high-risk areas (Burby, 2001). Today, the NFIP, administered by the Federal Emergency Management Authority (FEMA), remains the primary provider of flood risk protection and regulator of floodplain development in the United States. The NFIP underwrites over 90% of flood insurance

of the National Flood Insurance Program (NFIP) (Peralta and Scott, 2024) and Home Seller Disclosure Requirements (Lee, 2022).

policies, creates the most widely-used measures of flood risk through its flood mapping process, and sets construction standards for buildings in areas mapped as high risk (Kousky et al., 2018b).

After the NFIP was established in 1968, the program was rolled out to communities throughout the country in the late 1970s and 1980s.[5] When a community joined the NFIP, it obtained a Flood Insurance Rate Map (FIRM), produced by the NFIP using a hydrological study. After the FIRM was produced, developers of new buildings in specific areas had to comply with flood safety regulations, and flood insurance became available to homeowners.

In both the initial FIRM and subsequent, updated flood maps, an important distinction for both insurance policies and floodplain regulation is between areas that are determined to be high-risk, known as Special Flood Hazard Areas (SFHAs), and those that are not. In this paper, we will refer to the SFHA as the "flood zone." All new construction and substantial home improvements in the flood zone must comply with building regulations that require that a home's lowest floor lie above the Base Flood Elevation (BFE).[6] In the flood zone, some homeowners face a flood insurance mandate and all homeowners face higher flood insurance prices for otherwise-equivalent houses.[7] Approximately 50% of homeowners in the flood zone hold a flood insurance policy, compared to 2% outside of the flood zone (Bradt et al., 2021). In Florida, annual flood insurance premiums inside the flood zone cost twice as much as outside the flood zone: $820 compared to $435.

Throughout the United States, 10% of land[8] and 6% of properties are in the flood zone (First Street Foundation, 2020). Due to both climate change and population growth, the share of the US population at a level of risk that triggers SFHA classification is expected to rise from 13% to 15% by 2050 (Wing et al., 2018). This makes flood-zone-induced building requirements one of the most common forms of zoning regulation in the U.S., comparable to minimum lot area requirements, which apply to an estimated 16% of single-family homes (Song, 2021).

3.2.2 The Flood Mapping Process

Our spatial regression discontinuity approach relies on the assumption that flood zone delineation is a coarsening of a continuous measure of flood risk and does not follow the contours of true discontinuities in flood risk or other amenities. The validity of this assumption relies on the details

[5] Communities are geographic units specific to the NFIP. They are generally municipalities or unincorporated areas of a county.

[6] While popular images of elevated houses commonly show those on posts or piles, this adaptation tends to appear only in close proximity to the coast, where wave action can destroy walls. In the mostly inland areas we study, enclosed elevated foundations are more common. This approach allows garages and unfinished basements to be constructed at ground level. See Figure C.1 for an example of a house with an elevated foundation.

[7] Homeowners with federally-backed mortgages are legally required to purchase flood insurance.

[8] Authors' calculations using 2017 flood maps.

of the mapping technology. In our specific context, there is substantial scope for imprecision in the historic boundaries we exploit.

The accuracy of a flood map depends on both the accuracy of the estimates of land elevation and the accuracy of the hydraulic model which simulates the amount of excess water in a flood event (National Research Council, 2009). Historically, engineers estimated land elevation based on US Geological Service contour lines, which suffer absolute elevation error on the order of meters.[9] After floodwater heights have been mapped, the floodplain is delineated by transforming vertical flood elevation profiles into horizontal floodplain boundaries. Because the same elevation of floodwaters yields a much wider floodplain in flat than steep areas, the floodplain boundary delineation is four to five times more uncertain in flat areas, such as Florida, compared to hillier areas (National Research Council, 2009). The floodplains of inland Florida are particularly uncertain since their drainage is dominated by shallow water flow, an atypical landscape for which FEMA does not specify hydrology and hydraulics guidelines.

After the construction of the initial flood map, FEMA is required to update these maps every 5 years to account for improved mapping technology and changes in development that may impact flood risk (National Research Council, 2009). In practice they are often updated much less frequently: as of 2017, more than 50% of maps were more than 5 years old (U. S. Office of Inspector General, 2017). In between official remapping cycles, property owners can request map amendments to correct inaccuracies (National Research Council, 2009) or petition for a map carve-out if homeowners have physically changed the land elevation (e.g. by adding dirt, called "fill").[10]

3.3 Setting and Data

Our empirical context is the state of Florida, one of the most flood-prone and populous states in the United States. This makes it an ideal setting to study how floodplain regulation impacts housing markets and disaster damages. Nearly 50% of land and 19% of homes in Florida are located in the flood zone, underscoring the relevance of this form of regulation for real estate development across the state. (First Street Foundation, 2020). Florida alone accounts for 35% of the nation's NFIP policies.[11] We bring together four primary sources of data to conduct our analysis.

[9] Today, LiDAR technology has improved the accuracy of land elevation models. Powerful computing has also improved the precision of hydraulic modeling over time.

[10] According to a floodplain manager in Florida, in the early years of the program the scale of paper maps meant that fill-based carve-outs of the flood zone had to be at least 6 acres (personal communication). Because of this requirement, most houses did not find it cost-effective to pursue a carve-out. More recently, the adoption of digital maps has enabled these carve-outs at a smaller scale, and they have subsequently become more common.

[11] See Lingle and Kousky (2018) and https://nfipservices.floodsmart.gov//reports-flood-insurance-data for details.

Digitized Historic and Current Flood Maps Our analysis is organized around archival scans of early flood maps that we digitized for parts of eleven counties.[12] We aimed to collect the first Flood Insurance Rate Maps (FIRMs) ever drawn. In a few instances, constraints on the availability or formatting of these first maps made this impossible. In these cases, we were able to digitize maps that were drawn only a few years later. All but two of the 120 panels we digitized became effective between 1977 and 1984. Appendix 3.10.2 describes the sample selection process in more detail.

Figure 3.1a presents an excerpt of these digitized flood maps. Figure 3.1b shows the geographic coverage of our digitized sample. While budget constraints prohibited digitizing the entire state of Florida, we are able to obtain good coverage of most major population centers. Table 3.1 illustrates that our sample covers 10.5% of the land mass in Florida, but 14% of all homes, reflecting the fact that our digitized areas are more developed and populous than average.

We pair our newly-digitized historic flood maps with snapshots of flood maps for the whole state from 1996 and 2017. In our digitized counties, 32% of land is in the flood zone.

Satellite-Derived Land Use Data Figure 3.1a demonstrates that the floodplain distinctions are detailed, necessitating spatially-granular data on land use to study outcomes on either side of the boundary. We use two datasets to measure land use at two points in time. The first is US Geological Survey data on land use patterns contemporaneous to the time the original maps were drawn. This dataset consists of a 30x30 meter raster describing land use and land cover as belonging to one of nine mutually exclusive meta-categories, including urban/built-up land, agriculture, wetland, and water.[13] The categories were determined based on high-altitude photographs taken between 1971 and 1982 (1976 is the median and mode image date). We define "developed" land in this data as land falling into the "urban/built-up" category, which includes land used for residential, commercial, industrial, or transportation purposes. For current land use, we employ the National Land Cover Database (NLCD) from 2016, which classifies Landsat remote sensing imagery into similar categories of land cover, also in a 30x30 meter grid. Our main category of interest, "developed", indicates land that is covered by a mixture of constructed materials and mostly-lawn-grass vegetation.

Table 3.1 panel A presents land use summary statistics for the state of Florida and our digitized

[12] These archival scans were downloaded from FEMA's Map Service Center https://msc.fema.gov/portal/advanceSearch. In order to maximize power, we prioritized areas with substantial new development over the last 40 years. Our estimates on development when expressed in levels may therefore generalize less well to other settings, but this choice will not affect results expressed as a percentage of new development.

[13] Across Florida, the median number of raster grid cells per census tract is about 4900.

subsample. Commensurate with Florida's population boom between 1980 and 2020,[14] Table 3.1 illustrates that development increased substantially both statewide (2.6x) and in our sample of interest (2.5x).

Parcel Characteristics and NFIP Policies and Claims Data from the Florida Department of Revenue property tax records from 2005 to 2020 provide detailed information about structures, including sales prices and parcel outlines and location. We precisely geolocate the exact location of any buildings on each parcel using Microsoft's open-source building footprints dataset.[15] In Table 3.1 panel C we summarize average home prices statewide and in our sample of interest. We also obtain historical data on home prices from the 1980 Census (Manson et al., 2021). We use data from NFIP claims and policies from 2010 to 2018 to provide information about flood damages for insured structures.

Flood Risk Model To assess the risk profile of development across policy regimes, we draw on spatially-granular estimates of flood risk from a third-party hydrological model. This model is produced by the First Street Foundation, a nonprofit organization devoted to quantifying and communicating climate risks. First Street aims to improve on government-issued risk assessments, which have been criticized for being out-of-date and inaccurate (Wing et al., 2022).[16] First Street takes into account sources of flooding that NFIP maps ignore (e.g. rainfall), provides estimates for areas that FEMA had not been able to survey, and accounts for sea level rise due to climate change (First Street Foundation, 2020). Although First Street's model does not employ the "gold standard" of surveying that FEMA uses in the highest-risk locations, their validation exercises have achieved 80-90% flood extent similarity with historical observations and they are considered to "fus[e] the accuracy of local studies with the spatial continuity of large-scale models" (Wing et al., 2022). Nationally, First Street's model estimates that NFIP flood maps identify only 60% of areas that face a 1% chance of flooding every year (First Street Foundation, 2020). In Florida this discrepancy is smaller, but First Street and FEMA disagree about the exact location of risk. Appendix Table C.1 tabulates the discrepancies between FEMA's flood maps and the First Street model in our sample, showing that more than one-fifth of parcels are categorized differently by FEMA and First Street.

[14] Between 1980 and 2020, Florida's population more than doubled from 9.75 million to 21.5 million (US Census Bureau, 2022).

[15] This dataset is also derived from satellite imagery, mostly captured in 2019. See https://github.com/microsoft/USBuildingFootprints for more details.

[16] These criticisms regularly appear in the national media, see e.g., here and here.

3.4 Quasi-Experimental Evidence

We begin our analysis by describing the effects of flood zone designation on flood risk, which could occur both by shifting construction away from high risk areas and by mandating adaptation through building standards in those risky areas. In this section, we employ a spatial regression discontinuity design around the regulatory boundary to study the policy's impact on the location of new construction and perform an event study around the introduction of the policy to study its effect on flood damages among built structures.

3.4.1 Spatial RD Around the Boundaries of Regulation

Our spatial regression discontinuity design compares current development and house prices on either side of the historic regulatory boundary to investigate the extent to which floodplain regulations reduce construction in risky areas. We also use our boundary discontinuity to compare differences in home prices in regulated *versus* unregulated areas. In the context of our model, these two equilibrium points – prices and quantities inside and outside the regulated areas at the boundary – will later allow us to estimate the effects of floodplain regulations on consumers' residential choices and developers' construction costs.

Empirical Strategy Estimating the effect of flood zone designation on new construction presents two challenges. First, flood zone designation could be correlated with unobserved amenities, such as coastal access or views. Indeed, Table 3.1 indicates that land inside the flood zone is more likely to be water or wetlands and is closer to the coast. Second, flood zone designation may be endogenous to real estate construction, as the mapping process allows homeowners to deregulate parts of their properties by petitioning for map corrections or "filling" in dirt to elevate the land. Inside the flood zone, homeowners who are correctly mapped have an incentive to elevate their house to "escape" the flood zone. Homeowners who were incorrectly mapped have an incentive to petition FEMA to correct a mistake that overstates a home's risk. Meanwhile, owners of undeveloped land face no such incentives. Appendix Table C.2 shows direct evidence of such reverse causality: land that was developed as of 2004 is more likely to be remapped out of a floodplain in the next map revision than land that was undeveloped. This endogenous amendment process would lead to a mechanical negative correlation between development status and flood zone status that is unrelated to the causal effect of interest.

We address these two challenges with a spatial regression discontinuity design that leverages the first flood maps drawn in the late 1970s and early 1980s. The historic maps address concerns

about reverse causality, since homeowner petitions and reactive adaptations were not reflected in the original maps: amendments and endogenous adaptation happen only after the maps are drawn. The regression discontinuity addresses omitted variables bias by leveraging the coarse classification of the flood zone and the assumption that unobservable characteristics of the land evolve smoothly through the historical flood zone boundary. We probe this identifying assumption by examining land use outcomes before or contemporaneous to the drawing of these initial flood maps.

The boundary discontinuity design examines how outcomes at each 30x30m pixel vary as a function of the distance to the flood zone boundary. Specifically, we estimate

$$y_i = \beta \mathbb{1}\{d_i > 0\} + f(d_i) + \gamma_{j(i)} + \varepsilon_i, \tag{3.1}$$

where y_i is a characteristic of pixel i and d_i is the perpendicular distance from each pixel i to the nearest flood zone boundary, with positive values indicating that the pixel is located inside the flood zone. Our coefficient of interest is β, the magnitude of the discontinuity at the boundary. In our baseline specification, $f(d_i)$ are local linear functions allowed to differ on either side of the boundary.[17] Finally, since boundaries in this setting do not have natural segments, we include census tract fixed effects $\gamma_{j(i)}$ as a substitute for boundary fixed effects. We cluster standard errors at the census tract level to allow for spatial correlation in the error term.

For each outcome, we compute the MSE-optimal bandwidth proposed by Calonico et al. (2014) and estimate equation 3.1 on land within that distance of a flood zone boundary. Following previous work, we exclude boundaries that trace a body of water (Dell, 2010). Columns 4 and 5 of Table 3.1 present summary statistics for our boundary estimation sample: land close to a boundary is more developed than areas further from the boundary. Appendix Figure C.2 plots a histogram of the number of pixels in our estimation sample across distance-to-boundary bins.

Results We discuss our results in the context of an *intent-to-treat* framework: the treatment of interest is the initial flood zone designation, which may evolve over time. This is motivated by our focus on the effects of floodplain regulation on long-run adaptation to flood risk. Appendix Figure C.3 documents the evolution of the relationship between initial designation and floodplain status over time.

[17] Results are similar under alternative specifications of $f(d_i)$. Columns 3 and 4 of Table 3.2 illustrate robustness to alternative specifications. We observe similar effect sizes on both quantities developed and prices with a linear function of the running variable estimated on a fixed bandwidth or a fourth-order polynomial.

Exogeneity of Boundaries To validate our identifying assumptions, we check for smoothness in land use through the regulatory boundaries prior to flood zone designation.[18] If preexisting amenities differed discontinuously across the boundary, or if boundaries were drawn around the contours of existing development, we would observe discontinuous patterns in development around the flood zone boundary. We test this by estimating equation 3.1 with y_i equal to pre-period development. Figure 3.2a shows that pre-period development is smooth across the boundary, and the estimated coefficient, reported in Table 3.2, is small.[19] This test supports our hypothesis that the institutional details of the initial mapping process provide a compelling setting in which to conduct a boundary discontinuity analysis.

Development Falls in the Flood Zone By 2016, we see a sharp discontinuity in development at the SFHA boundary, as shown in Figure 3.2b. Table 3.2 reports the estimated level shift at the boundary ($\hat{\beta}$), which indicates that land just inside the SFHA is 3.8 percentage points less likely to be developed than land just outside the SFHA. This effect is substantial: it is 9% of the outside-SFHA mean level of development and it represents an 18% reduction as a share of total new development occurring between 1980 and 2016 just outside the SFHA.[20] The effect is driven by single family homes, which make up the majority of residences (87%) in our sample. This effect indicates both the potential for substantial reductions in flood damages via reduced building in risky areas and the possibility of costs imposed on developers or consumers that yield this behavioral response.

Prices Increase in the Flood Zone The effects of floodplain regulation on house prices are *ex-ante* ambiguous. On the one hand, by suppressing demand for regulated houses, flood zone designation could push prices down. On the other hand, by requiring developers to employ costlier construction methods, flood zone designation could push prices up. Figure 3.3, which plots the estimated coefficients for the sales price outcome, indicates that supply must shift inward substantially as a result of the policy: though we observe a large decrease in development, prices are non-decreasing through the boundary. In fact, our point estimates suggest that house prices are 6.7% higher inside the flood zone, indicating that the construction costs imposed in the floodplain

[18] Our land use data was collected via aerial photographs between 1971 and 1982, while the flood maps were drawn between 1977 and 1984. Most aerial photographs were taken during or before 1976, before any of the maps were drawn. While it is possible that some aerial photographs were taken after the maps had been drawn, we will interpret these land use outcomes as a pre-period. Land use evolves slowly and the worst-case scenario is that the photographs were taken five years after the drawing of the map.

[19] Appendix Figure C.4 and Table C.3 examine smoothness in other pre-period land use outcomes.

[20] Table 3.2 Panel B shows that these results are robust to alternative definitions of development, including the share of land covered by a building footprint.

dominate any negative demand effect driven by mandatory flood insurance, higher flood insurance prices, or any salience or risk perception effects of living in a flood zone.[21]

Our evidence suggests that this positive price effect should be interpreted as an inward shift in supply. We rule out other competing explanations: changes in willingness to pay for housing characteristics – including house adaptation — and changes in amenities. In Appendix Table C.4 and Figure C.5, we demonstrate that prices increase for houses built both before and after the introduction of the regulation, and price increases persist when we control for compositional changes via polynomials in square footage and lot area and indicators for county-by-sale-month and year built. Furthermore, Panel B and Appendix Figure C.6 show that single-family housing square footage, the main housing quality measure we observe in the property tax records, exhibits only a marginal increase across the boundary in houses built after the regulation.[22] To rule out changes in amenities across the flood zone boundary, column 5 of Table 3.2 shows that our results are robust to excluding areas close to the coast, complementing the pre-period smoothness test in Figure 3.2a.

Testing for Heterogeneity Across Risk Levels We observe our baseline pattern — sharp reductions in development without commensurate decreases in prices at the SFHA boundary — across levels of flood risk. We leverage the fact that the SFHA boundaries cut across a distribution of flood risk levels, measured using estimates from the First Street hydrological model. Appendix Figure C.7 replicates our main regression discontinuity plots, split by land in census tracts with below- *versus* above-median flood risk.[23] Beyond qualitatively similar patterns, the magnitude of the reduction as a share of new development is quantitatively similar across these two samples: 18% in lower-risk tracts *versus* 16% in higher-risk tracts. Price effects in both samples are noisy, but positive. These results are consistent with the policy's design as a coarse, binary instrument.

[21] This finding contrasts with recent work that has found flood zone designation *decreases* house prices, e.g. Hino and Burke (2021). That work exploits map updates, capturing short-term demand effects. We study the effect of flood zone designation over 40 years. This long-run setting allows supply to respond to mandatory building codes, leading to an increase in prices that offsets short-run reductions based on demand effects alone.

[22] We note that there is a positive, albeit not statistically-significant, point estimate on the outcome of log square footage in post-regulation houses. However, because controlling for square footage does not eliminate (or reduce) the point estimates of SFHA designation on sales price, we remain comfortable interpreting our results as an inward shift in supply. Ruling out compositional changes in housing characteristics is a simplification to focus on the primary housing attribute of interest, adaptation status.

[23] Flood risk is measured as the depth of the 100-year flood in a grid cell's 1980-era census tract. In our analysis sample, the median depth is 1.05 feet.

3.4.2 Event Study Around the Introduction of Building Standards

The preceding results indicate that regulations impose a cost on new home construction that shifts supply inward. In this section, we investigate whether these regulations also reduce damages via impacts on built homes, and if so, whether consumers are willing to pay for this attribute.

Empirical Strategy Following Wagner (2022), we exploit the fact that a community had to adopt flood-safe building standards at the time it enrolled in the National Flood Insurance Program, but not before. This suggests an event-study design, in which we regress our outcomes of interest against the year a house was built relative to community enrollment within the flood zone. We estimate the following specifications for all counties in Florida:

$$\begin{aligned} m_{jbt} &= \Sigma_r \beta_r \mathbb{1}\left\{r = b - e_j\right\} + \gamma_j + \varepsilon_{jbt} \\ p_i &= \Sigma_r \beta_r \mathbb{1}\left\{r = b_i - e_{j(i)}\right\} + \gamma_{j(i)} + \varepsilon_i \end{aligned} \tag{3.2}$$

where m_{jbt} measures insurance payouts per dollar of coverage in policy year t, for homes in census tract j, built in year b, p_i is the log sales price of house i, r is the construction year relative to the year of NFIP enrollment of the census tract (e_j), and γ_j are census tract fixed effects.[24] The model is estimated at the census tract j by house construction year b by policy year t level for insurance claims (line 1) and at the house i level for house prices (line 2).[25]

Our baseline specification in equation 3.2 relies on the sharp change in outcomes in the year the building standards are introduced (similar in spirit to an RD). Appendix Figure C.8 replicates the analysis and finds similar results in a difference-in-differences specification that includes controls for construction year.

We also estimate the following pooled specification, among houses built in a ten-year window around NFIP enrollment:

$$\begin{aligned} m_{jbt} &= \alpha + \beta Post_{jb} + \nu r_{jb} + \eta r_{jb} Post_{jb} + \gamma_j + \varepsilon_{jbt} \\ p_i &= \alpha + \beta Post_i + \nu r_i + \eta r_i Post_i + \gamma_{j(i)} + \varepsilon_i \end{aligned} \tag{3.3}$$

[24] Census tracts are smaller than communities.

[25] Because individual claims cannot be linked to individual policies, we aggregate from the house to the census-tract-by-flood-zone-by-relative-year-built level for outcomes related to insurance payouts or policies. Appendix 3.10.3 describes the construction of the datasets used in this analysis in more detail and Appendix Table C.5 presents summary statistics.

where *Post* indicates houses constructed in or after the year of the community's NFIP enrollment ($r \geq 0$). We cluster standard errors in all specifications at the census tract level. Under the assumption that the year of construction was not manipulated, β indicates the causal effect of building standards.[26]

Results Figure 3.4 presents the coefficient estimates on relative year from the event study specification (equation 3.2). Variable means and regression coefficients from the pooled specification (equation 3.3) are presented in Table 3.3.

Building Standards Reduce Flood Damages Figure 3.4a shows that the introduction of building standards causes insurance payouts to fall by $1.60 per $1000 of coverage (Table 3.3). This accompanies a sharp increase in reported house elevation in flood insurance policies, as mandated by the regulation (Appendix Figure C.9a). The $1.60 (per $1000 of coverage) reduction represents a 55% decrease in expected flood damages (on a pre-period mean of $2.93), highlighting that building standards generate substantial social benefits in expected damage reduction. At average coverage levels, this reduction in expected flood damages is equal to 3% of the average home value (using a 5% discount rate).

The reductions in damages following the introduction of building standards are largest in high-risk locations. We conduct the event study analysis separately by the flood risk of the census tract, either measured via the First Street hydrological model or via flood insurance payouts. Appendix Table C.6 shows that locations with high baseline risk experience larger reductions in damages (in levels).

House Prices are Unchanged Despite Lower Damages Despite their higher social value, houses that are compliant with the building standards do not command a higher price than non-adapted houses (Figure 3.4b). The point estimate on log prices is -0.007 (Table 3.3), and the upper bound of the confidence interval reflects only one third of the reduction in expected damages. This result implies that homeowners are unwilling to pay for reductions in risk, as long as all other house attributes remain constant (see Appendix 3.10.3 for a stylized model). This is plausible in our setting, as homes are typically adapted by adding fill below the house while keeping the structure constant.

We conduct a series of empirical tests to probe the conclusion that consumers are unwilling to pay for flood safety. First, we confirm that the null house price effect does not mask price declines

[26] Appendix Figure C.10 shows no bunching of house construction in the years prior to NFIP enrollment, ruling out such manipulation.

generated by changes in other house attributes. We show in Appendix Figure C.11 that the null effect on price is robust to controlling flexibly for observable housing characteristics, including polynomials in parcel size and total living area and fixed effects for county-by-month-of-sale and year of construction. We also investigate the possibility that the existence of stairs on adapted houses may be a confounding disamenity by testing whether the observed result is concentrated among consumers with a strong distaste for stairs. We proxy for this heterogeneity by splitting the sample across census tracts with an above- *versus* below-median elderly population. We find no evidence of this confounding disamenity (see Appendix Table C.7). Second, we show that the null effect on price is not driven by a failure of insurance prices to reflect changes in risk, as premiums for adapted homes incorporate at least 75% of the reduction in risk (Appendix Figure C.9b). Finally, we show similar effects on price across locations with varying flood risk (Appendix Table C.6).

This result therefore confirms existing work documenting a wedge between social and private valuations of flood risk (Gallagher, 2014; Kousky et al., 2018a; Mulder, 2021; Bakkensen and Barrage, 2022; Wagner, 2022). The fact that safe houses provide no private value to consumers indicates the presence of frictions, including behavioral frictions such as risk misperception or moral hazard from consumer expectations of government aid in case of disaster (or both). Regardless of the source, this wedge between private and social value will yield construction of inefficiently-risky houses — both in their type (adapted *versus* not) and location — absent regulation.

3.5 An Equilibrium Model of the Housing Market

The quasi-experimental results indicate that the current policy affects both flood damages and regulatory costs, and may improve social welfare. Floodplain regulation suppresses construction and requires flood-safe building in high-risk areas. The policy therefore potentially decreases damages on both extensive and intensive margins of construction. However, our results also indicate that the costs of building standards are substantial, since flood zone designation substantially decreases new construction without a commensurate fall in price. Finally, building standards reduce damages, but consumers are not willing to pay more for houses with lower flood risk, suggesting a role for policy to correct inefficient risk exposure.

While informative, these results alone are insufficient to quantify the total effect of the policy on either damages or overall welfare. The policy's effects depend on the location of counterfactual construction: if the regulation shifts construction to equally-risky areas, damages will not fall. Counterfactual risk depends on the joint distribution of risk and amenities, since more-desirable locations will attract more counterfactual construction. Additionally, the large size of the regulated

area — nearly one-third of all land in our sample is currently designated as a flood zone — may lead to equilibrium price effects in unregulated locations. Higher demand and consequently higher prices outside the flood zone could then reduce the relative incentive to build in safe, unregulated areas and also reduce consumer surplus market-wide. Finally, we require a model to quantify the regulation's costs to developers and consumers. To account for counterfactual location and spillover effects, quantify the regulation's costs, and compare welfare across current and counterfactual policies, we specify and estimate a model of residential choice and real estate development.

Our model's key parameters of interest describe how flood zone designation shifts the demand for and supply of housing in each location, capturing how consumers and developers trade off home prices and living or building in a regulated area. Both of our empirical exercises inform our model. We use our event study estimates that consumers are unwilling to pay for adaptation to simplify the model: consumers choose only among differentiated locations and developers do not endogenously supply adapted homes. We use our cross-boundary changes in prices and quantities from our RD analysis as moments to estimate the effect of flood zone designation on consumer choices and construction costs.

3.5.1 Residential Choice

Motivated by the result in Section 3.4.2 that consumers are indifferent between adapted (post-standards) and non-adapted (pre-standards) homes, we model consumers as choosing among uniform homes across differentiated locations. We further assume that consumers do not privately value differences in flood risk when deciding between houses, consistent with prior work and with the results in Section 3.4.2. Each individual i makes a discrete choice of where to live within market m, which we take to be a county.[27] Locations are differentiated goods characterized by tract j and flood zone designation status z. Census tracts are small geographic units of analysis: the average county in Florida has 63 census tracts, each containing roughly 1,600 residential structures.

Following the standard discrete choice framework of Berry et al. (1995), we write the indirect utility of individual i living in location jz as:

$$u_{ijz} = \alpha^D p_{jz} + \phi SFHA_z + X_{jz}\beta + \xi_{jz} + \varepsilon_{ijz} \tag{3.4}$$

where p_{jz} is the log price of housing in location jz,[28] $\phi SFHA_z$ indicates flood zone status, X_{jz} is

[27] In Florida, counties are large but tend to only contain one major city and commuting zone.

[28] The price p_{jz} is for the bundle of housing that a consumer purchases, which includes both the structure and the land on which the structure is built.

a vector of observed housing characteristics,[29] ξ_{jz} are unobserved amenities, and ε_{ijz} is an i.i.d. preference shock, distributed according to a Type 1 Extreme Value Distribution. Modeling the effect of the flood zone as an indicator captures the binary nature of the regulation — in *versus* out of the flood zone — and is consistent with our evidence documenting similar effects across locations with heterogeneous levels of risk.

A large body of work suggests that misperceptions likely drive consumers' failure to internalize flood risk (Mulder, 2021; Bakkensen and Barrage, 2022; Wagner, 2022). This complicates interpretation of $\phi SFHA_z$ if it debiases consumers, as SFHA status may influence choices without imposing a utility cost. Following Allcott and Taubinsky (2015), we will refer to equation 3.4 as a consumer's *decision utility,* which is not necessarily her *experienced utility*. While equation 3.4 alone is sufficient for estimation and to characterize the effects of the policy on housing market outcomes, normative evaluation requires an assumption about the welfare-relevance of $\phi SFHA_z$. We return to this issue in Section 3.6.

Individuals choose the location jz that maximizes their decision utility within a market m. The fraction of individuals choosing to live in location jz is:

$$s_{jz} = \frac{\exp(\alpha^D p_{jz} + \phi SFHA_z + X_{jz}\beta + \xi_{jz})}{\Sigma_{j' \in J_m, z' \in \{0,1\}} \exp(\alpha^D p_{j'z'} + \phi SFHA_{z'} + X_{j'z'}\beta + \xi_{j'z'})}. \tag{3.5}$$

3.5.2 Housing Supply

Our model of housing supply is designed to simply and flexibly capture heterogeneity in housing supply elasticities across locations. Each tract-zone pair is composed of L_{jz} plots, each of which could either be developed into a house or used for some outside option (e.g. agriculture). The value of the outside option for plot g is denoted c_g (for opportunity cost) and is distributed Normally with a mean and standard deviation that varies by census tract j: $c_g \sim N(\mu_j, \sigma_j^2)$. Developers make static decisions about whether to develop at two points in time: before the regulations are imposed ($t = 0$) and after they are imposed ($t = 1$).

The value of developing a house in period t depends on the (log) price p^t_{jz} for which it could sell, which varies by location and time, and the cost to build the house η^t_{jz}, which also varies by location and time and increases by a constant amount ψ when the house is adapted. A house's adaptation status is determined exogenously by building standards and varies by location (SFHA *versus* not)

[29] Observed housing characteristics include the share of residences that are single-family houses, the average age of residential buildings, the average square feet of land and living area for residential parcels, the share of buildings ranked "average," "high," or "superior" quality, and the share of parcels that are residential, commercial, industrial, agricultural, or open space.

and time (pre- *versus* post-regulations).[30]

The development decision for an undeveloped plot g in time period t is given by

$$D_g^t = 1\{p_{jz}^t \geq c_g + \psi E_{jz}^t + \eta_{jz}^t\} \tag{3.6}$$

where $D_g^t = 1$ indicates that a plot of land is developed and E_{jz}^t indicates whether or not houses built in location jz are adapted in period t. Let $\Phi(\,)$ denote the Normal Cumulative Distribution Function. The share of land developed at the end of time $t = 1$ is

$$\Phi\left(\frac{p_{jz}^1 - \psi E_{jz}^1 - \mu_j - \eta_{jz}^1}{\sigma_j}\right). \tag{3.7}$$

3.5.3 Equilibrium

In equilibrium, the quantity of housing supplied in each location equals the number of individuals choosing to live there:

$$L_{jz}\Phi\left(\frac{p_{jz}^1 - \psi E_{jz}^1 - \mu_j - \eta_{jz}^1}{\sigma_j}\right) = Q_{jz} = N_m s_{jz} \tag{3.8}$$

where L_{jz} is the total amount of land in location jz, Q_{jz} is the quantity of developed land in jz, and N_m is the number of households in the market.

3.5.4 Estimation

Demand We use the standard inversion to estimate 3.4 from observed market shares:[31]

$$ln(s_{jz}) - ln(s_{0m}) = \delta_{jz} = \alpha^D p_{jz} + \phi SFHA_z + X_{jz}\beta + \xi_{jz} \tag{3.9}$$

[30] In some locations, a majority of houses built in the flood zone before the introduction of building standards appear to be elevated above the minimum required level, as reported in NFIP policies. This is concentrated in locations where the base flood elevation is very low, indicating that these homes are likely measured as adapted despite being built under standard construction practices. To be conservative, we assume these houses are adapted even in the unregulated benchmark.

[31] We construct the empirical market shares using

$$\hat{s}_{jz} = \frac{Q_{jz}^{2016}}{\sum_{j' \in J_m, z' \in \{0,1\}} Q_{j'z'}^{2016}}$$

where Q_{jz}^{2016} is the total amount of developed land in geography jz.

where s_{0m} is the market share of an arbitrary geography within each market that we have normalized to be utility 0.

Because amenities ξ_{jz} may be correlated with both $SFHA_z$ and prices p_{jz}, simply estimating equation 3.9 via OLS would yield biased estimates. These correlations arise because flood zone status may coincide with other amenities, like coastal access (biasing ϕ). In equilibrium, higher-amenity locations will command higher prices (biasing α^D).

We address the possible correlation between amenities ξ_{jz} and $SFHA_z$ using our boundary discontinuity design, which isolates the causal effect of SFHA designation. In the boundary discontinuity analysis, we assume that as distance to the boundary approaches zero, amenities ξ_{jz} are constant on either side of the boundary, and thus independent of SFHA status. We therefore construct moments to match the cross-boundary differences we document in the RD analysis.

To address the possible correlation between amenities ξ_{jz} and prices p_{jz}, we construct instruments for p_{jz} that exploit variation in the characteristics of *other* locations within the market, excluding areas surrounding each location jz. The characteristics of distant houses in the same housing market influence prices in location jz in equilibrium, but do not directly affect the utility provided by a house in location jz.[32] Our instruments include averages of observable characteristics X_{jz} of locations in the same housing market that are located more than 3 miles away from geography jz, weighted by land area. Following Bayer et al. (2007); Calder-Wang (2021); and Almagro et al. (2023), we also construct a price vector $\tilde{p}_{jz}$ that rationalizes market shares under no unobserved amenities (i.e., setting $\xi_{jz} = 0$), using the equilibrium conditions to capture the price impact of the observable attribute space.

Discussion We estimate the floodplain effect on consumer choices (ϕ) using observations close to the boundary. We assume that ϕ is constant across regulated areas. The most salient information provided (in a flood zone *versus* not) and regulations imposed (insurance mandate) are consistent across the floodplain. Although the exact insurance premium may vary within the flood zone, flood zone (SFHA) status is an important determinant of premiums. Moreover, we observe qualitatively similar patterns across boundaries in locations with heterogeneous risk (Appendix Figure C.7). However, to the extent that the true flood zone effect on demand is larger in higher-risk areas than at the flood zone boundary, ours will be an underestimate.

We have imposed the assumption that consumers do not respond to expected flood damages. If this assumption is incorrect, any increases in flood expenses to which consumers *do* react will load onto the flood zone term if they change discretely at the boundary (e.g. through insurance premiums)

[32] Because each location is a small share of the market, we assume that prices in location jz do not in turn affect characteristics of houses constructed in other locations.

or onto the unobserved amenity terms ξ_{jz} if they do not change discretely at the boundary. Underestimating consumer responsiveness to risk would lead us to overstate the gains from regulation in our normative analysis, but would not impact any conclusions about the positive effects of the policy on construction, damages, or house prices.

Supply We assume that flood zone designation affects supply only via building standards mandating adaptation. However, estimating ψ is still challenging, as flood zone status $SFHA_z$ may be correlated with construction costs η_{jz}, like the presence of wetlands. As with demand, we address this challenge by constructing moments to match the cross-boundary differences in the regression discontinuity analysis.

To capture heterogeneity in housing supply elasticities across tracts, we match estimates of tract-specific supply elasticities from Baum-Snow and Han (2023), which are estimated in the same locations as our setting. These estimates allows us to capture realistic patterns of within-market heterogeneity in supply elasticities, e.g. that housing supply in urban, coastal areas is more inelastic than inland, suburban or rural areas.

Estimation Details Estimation proceeds in two steps. First, in a preliminary step we calibrate μ_j and σ_j based on estimates from Baum-Snow and Han (2023). Then, we estimate the remaining parameters via the Generalized Method of Moments (GMM) using two sets of moment conditions. The first set are regulatory boundary moments. These moments require that cross-boundary differences in price ($\beta^{p,2016}$) and pre-period ($\beta^{q,1980}$) and post-period ($\beta^{q,2016}$) development shares at the boundaries match the coefficient estimates in the spatial regression discontinuity analysis. We operationalize this requirement by constructing moments restricted to a window around the regulatory boundary (100 feet). Even within this narrow window, average amenities and construction costs across the boundary may diverge from the differences at the limit. Because of this, we construct moment conditions that directly target the regression discontinuity estimates from Section 3.4.1 while allowing for mean differences across the boundary in the 100-foot window. The second set of moments are based on the exogeneity of instruments $\tilde{p}_{jz}$, average characteristics $X_{j'z'}$ of distant locations in the same market, and own-location characteristics X_{jz}. Appendix 3.10.3 specifies these moment conditions, provides more details about the estimation procedure, discusses the data used for estimation, and assesses model fit.

Parameter Estimates Table 3.4 presents select parameter estimates with standard errors in parentheses.[33] We estimate that α^D (price elasticity of demand) is approximately -1.3, in line with other

[33] Appendix Table C.9 presents parameter estimates under alternative assumptions about supply elasticities and alternate values of boundary discontinuity estimates.

estimates (Song, 2021). Flood zone status is disliked by consumers (negative ϕ) and imposes costs on developers (positive ψ): consumers are willing to pay 27% more to avoid living in a floodplain, and it costs developers 24% more to build a compliant home. Intuitively, we arrive at these numbers by rationalizing the changes in quantity and price around the regulatory boundary with estimates of how consumers trade off home prices with other attributes and housing supply elasticities.

The result that consumers are willing to pay 27% more to avoid floodplain regulations is at the high end of a range of recent estimates, which find floodplain discounts ranging from 1 to 28% (Indaco et al., 2019; Gibson and Mullins, 2020; Hino and Burke, 2021; Lee, 2022).[34] This difference could be attributed to our setting: residents living in flood-prone Florida may be particularly sensitive to signals of risk as awareness of climate change grows. Nevertheless, a 27% premium on avoiding the floodplain exceeds the risk difference between the floodplain and unregulated areas, equal to 11% of average house price. Flood zone designation may cause consumers to over-update beliefs about risk, as the policy does not communicate risks in terms of expected damages. There may also be hassle costs associated with complying with floodplain regulations that could contribute to this estimated effect.

The estimated 24% increase in construction costs is also large, but within the plausible (albeit wide) range of estimates of the effect of building codes and zoning regulations on construction costs: from 5% to 42% (Listokin and Hattis, 2005; Emrath, 2021; Song, 2021). The wide variation reflects both differences in strategies to estimate regulatory costs and variation in the types of regulations imposed. Yet, our informal interviews indicated that a 24% increase in costs is reasonable. For example, the minimum elevation requirement may necessitate a stem wall, which can add $100,000 to the cost of a new build. The results in Section 3.4.2, combined with estimates of average forward-looking risk levels from our hydrological model, indicate that mandating adaptation reduces expected flood damages by 5% of home value. On average, the benefits of mandatory adaptation are less than costs. However, this comparison masks substantial heterogeneity across locations within the flood zone: mandating adaptation in the highest-risk areas will generate social value, but in safer ones, will incur net social costs (see Appendix Table C.6).

3.6 Model-Informed Estimates

We use our estimates to quantify the benefits and costs of *status quo* regulation and alternative policy designs. First, we investigate the impact of floodplain regulation by simulating an equilibrium with *versus* without the *status quo* policy. Then, we compare the performance of the observed

[34] A meta-analysis (Beltrán et al., 2018) finds an even wider range of a 75.5% discount to a 61% premium.

policy to a counterfactual first-best that imposes corrective taxes in the spirit of Pigou (1920) on consumers equal to the present discounted value of expected damages in each tract-by-flood-zone.[35] Finally, we propose and evaluate a change in the targeting of the *status quo* policy, holding the binary policy design constant while changing the location of the regulated areas. Motivated by our estimates that imposing adaptation may only be socially-optimal in the highest risk locations, this counterfactual restricts regulation to locations where the risk-reducing benefits of mandating adaptation exceed costs.

For each factual and counterfactual policy — determined by flood zone designation $SFHA_z$ and taxes τ_{jz} — we search for a price vector P_{jz}^{cf} and adaptation status E_{jz}^{2016} that equates housing supply and demand in each location:

$$L_{jz}\Phi\left(\frac{\ln\left(P_{jz}^{cf}\right)-\psi E_{jz}^{2016}-\mu_j-\eta_{jz}^{2016}}{\sigma_j}\right)= N_m\frac{\exp\left(\alpha^D\ln\left(P_{jz}^{cf}+\tau_{jz}\right)+\phi SFHA_z+X_{jz}\beta+\xi_{jz}\right)}{\Sigma_{j'\in J_m,z'\in\{0,1\}}\exp\left(\alpha^D\ln\left(P_{j'z'}^{cf}+\tau_{j'z'}\right)+\phi SFHA_{z'}+X_{j'z'}\beta+\xi_{j'z'}\right)} \tag{3.10}$$

given our estimated parameters and values of ξ_{jz} and η_{jz}.[36] Under the status quo policy, τ_{jz} is equal to 0 and houses are adapted when it is mandated in the flood zone. Under the unregulated benchmark (no policy), $SFHA_z = 0$ and $\tau_{jz} = 0$. In this scenario, there exists no requirement to adapt to a minimum standard, and consumers face no disincentives to locate in risky areas (e.g. an insurance mandate or higher insurance premiums). Under the corrective taxation policy, $SFHA_z = 0$ and taxes vary with both location and adaptation status. Homes are endogenously adapted in locations where it is socially-efficient to do so, that is, where the gains from lower tax rates on an adapted house exceed the difference in construction costs. Solving equation 3.10 generates counterfactual prices and quantities of development in each of the locations jz. Counterfactuals assume a closed city, or that the 2016 population in each county is held constant at observed levels.

We quantify the magnitude of flood-risk reduction across counterfactuals — the policy's intended goal. We also decompose the risk reduction due to building in safer areas *versus* mandated adaptation-in-place in risky areas. This risk quantification exercise relies on both hydrological-model-based estimates of flood risk by location and our estimates from Section 3.4.2 on the risk-

[35] This policy could be implemented as mandatory flood insurance with actuarially-fair rates at the property level. Over the course of this research project, the National Flood Insurance Program has moved the program in this direction — from a very coarse to a more granular pricing scheme. However, even under this recent policy change, prices likely will not increase to reflect the level of damages estimated by First Street Foundation, and adaptation will continue to be mandated in all areas demarcated as flood zones.

[36] We take the housing stock as of 1980 as given and hold it constant across counterfactuals.

reducing benefits of the imposed building standards.[37] Consistent with our results in Appendix Table C.6 showing larger damage reductions in higher-risk areas, we assume that the effect of building standards (mandated adaptation) on flood damages is proportional to the level of flood risk of each location. Appendix 3.10.4 describes how we compute expected damages in further detail.

Quantifying the risk-reducing benefits of each policy does not require a normative welfare framework. In Section 3.6.2, we will add further assumptions in order to calculate the social welfare impacts of observed and counterfactual policies.

3.6.1 Effects on Flood Damages

Table 3.5 presents results. Relative to a counterfactual without the policy, *status quo* floodplain regulation (column 2) almost triples the number of new flood-adapted houses. It also reduces the number of homes built in regulated areas by 17%, reallocating more than 180,000 houses out of the regulated floodplain. Note that this effect is (slightly) smaller than the boundary discontinuity estimate (an 18% reduction in new construction). This difference is due to two factors, both of which motivate our use of a model to extrapolate our reduced-form results to counterfactual outcomes. First, with the model we account for the market-wide distribution of amenities. At the boundary, amenities are identical, but at the market level, amenities are correlated with regulation status. This leads the boundary discontinuity analysis to over-estimate the policy's total effect. Second, because such a large share of Florida is at risk of flooding, regulations have market-wide price effects that undermine the risk-reducing effect of the policy. As Table 3.5 illustrates, prices outside of the regulated area increase as consumers substitute to unregulated areas.

The policy achieves its goal of reducing flood damages by $8,737 per newly-developed house, a 62% reduction in damages. The gains from adapted construction mandated via building standards account for 84% of this reduction, while incentivizing construction in safer areas contributes the remaining 16%. These damage reductions are substantial, both in absolute numbers (a decrease in $3.8 billion of expected damages per county), and in comparison to alternative policy instruments: the number of homes relocated in just these eleven counties exceeds the total number of houses removed from risky areas by the NFIP's home buyouts program across the entire nation (Frank,

[37] We use this external measure of flood risk rather than the FEMA flood maps for a few reasons. First, the FEMA flood maps have received extensive criticism for being out-of-date and backwards-looking and for failing to include certain important components of flood risk, e.g. pluvial (rainfall) risk. The First Street model incorporates climate change predictions as well as all major flood drivers in a novel peer-reviewed approach. Second, the First Street flood risk estimates provide granular estimates of the average annual loss for each parcel. First Street's estimates are increasingly used in economics research as an independent assessment of flood risk (Bradt et al., 2021; Mulder, 2021; Sastry, 2021).

2020a).

The damage reductions due to the policy approximate the first-best benchmark, implemented via the corrective tax policy (column 3). However, the damage reductions under the tax policy are achieved with fewer distortions in the housing market: 43% as many houses are relocated out of the designated floodplain and 35% as many houses are adapted. Relative to the *status quo* policy, a corrective tax achieves the same damage reductions by incentivizing *less* adaptation overall, but simultaneously setting stronger incentives to suppress development in the highest-risk locations (see Figure 3.5). A better-targeted counterfactual policy (column 4) approximates the first-best policy's substantial reduction in damages with fewer housing market distortions, allowing more construction in moderately-risky areas. The targeted policy regulates fewer locations and achieves a greater share of damage reductions from relocation (19%) by restricting regulations to *only* the highest-risk locations.

These findings suggest that *status quo* regulations generate the intended benefits of substantial reductions in expected flood damages. However, an important question is the extent to which the costs of housing market distortions offset the policy's benefits under each counterfactual, and whether these same benefits could be achieved with smaller distortions.

3.6.2 Effects on Social Welfare

In this section, we develop a welfare framework that allows us to quantify the policy's costs and compare these costs against the estimated benefits discussed in Section 3.6.1. Based on results — in both our setting and others — that consumers do not internalize flood risk in their housing decisions, we define social welfare as the sum of consumer surplus, producer surplus, government revenue raised,[38] and uninternalized expected flood damages. Expected flood damages are included in the calculation of social welfare despite the fact that consumers do not appear to value them because they impose social costs, either onto the government (externalities) or onto the consumers themselves (internalities).

To calculate consumer surplus, we must make an assumption about the normative interpretation of the effect of floodplain status on consumer choices. Does $\phi SFHA_z$ represent real costs or simply a change in expectations about flood risk? A large body of work suggests that misperceptions likely drive consumers' failure to internalize flood risk (Gallagher, 2014; Mulder, 2021; Bakkensen and Barrage, 2022; Wagner, 2022).[39] Therefore, we assume in our baseline calculations that the

[38] Floodplain regulations impose higher flood insurance premiums and therefore have a revenue impact that must be incorporated in social welfare calculations.

[39] See Wagner (2022) for a discussion about the likely contribution of misperceptions relative to other frictions in explaining observed consumer choices in the face of flood risk.

magnitude of ϕ in excess of the financial costs of living in a regulated area (via increased insurance prices) is *not* welfare-relevant, and is instead de-biasing as in Allcott and Taubinsky (2015). This would be incorrect if flood zone status imposed substantial hassle or psychic costs on consumers. While we think that the magnitude of these costs is likely to be small, we also report results under the assumption that ϕ is fully welfare-relevant and represents costs imposed upon consumers.

We report results in Table 3.6, comparing social welfare under the observed regulation and counterfactual policies to the unregulated benchmark.[40] In our baseline specification, *status quo* floodplain regulation achieves over $28 billion of social welfare gains in our setting, or $5,919 per newly-developed house. Because of the wedge between the private and social value of flood risk, risk would be inefficiently-high in the absence of any policy intervention. However, the social welfare effects of the *status quo* policy are theoretically ambiguous, as the risk-reducing benefits need not exceed the costs imposed via distortions in the housing market. We find positive and substantial social welfare gains, as the policy corrects inefficient exposure to flood risk. The policy would need to impose substantial deadweight loss via the demand cost ϕ — imposing hassles four times the magnitude of financial costs of living in a flood zone — to overturn the result that the status quo policy leads to a net increase in social welfare.

Despite achieving the socially-optimal degree of damage reduction, the *status quo* policy introduces distortions that reduce its total social welfare gains, which fall 27% short of the first-best outcome implemented by the corrective tax policy. This reflects the patterns documented in Table 3.5. The damage reductions achieved under the *status quo* policy impose substantially more distortions than the corrective tax: yielding *too little* construction and inefficient adaptation in moderately-risky areas, while still permitting *too much* construction in the riskiest locations.

While the corrective tax implements the first-best, it involves substantial changes to policy design that may be challenging to implement in practice. We conclude by assessing whether there are simple changes that can improve upon the *status quo:* is the 27% shortcoming, relative to the first-best, driven by the limitations of regulating continuous flood risk with a binary instrument, or can improvements be achieved with better targeting of regulated areas? In columns 4 and 6, we explore a simple "remapping," imposing regulations only in areas where the social benefits from adaptation exceed costs (corresponding to column 4 of Table 3.5). When well-targeted, the simple policy design can achieve 94% of first-best social welfare gains, equivalent to $7,567 per newly-developed house.

[40] We compute per-person consumer surplus in each market as $CS_i = -\frac{1}{\alpha^D} \ln \Sigma_{j,z} \exp(\alpha^D p_{jz} + \phi_F SFHA_z + X_{jz}\beta + \xi_{jz})$, where ϕ_F reflects our estimate that the PDV of insurance premiums amounts to about 5% of average house price (for our baseline specification). We define producer surplus as the integral of price minus costs for those who develop in the post-period. We compute compensating variation for both producer and consumer surplus to account for the fact that prices are modeled in logs. See Appendix 3.10.5 for more details.

3.7 Conclusion

For over 40 years, floodplain regulations have influenced housing markets, with limited evidence on their costs or benefits. This paper combines a spatial regression discontinuity analysis of floodplain boundaries, an event study of the introduction of flood-safe building standards, and a model of the housing market to investigate the effects of floodplain regulations on the location of construction, housing prices, expected damages, and social welfare.

Local to the regulatory boundary, floodplain regulation decreases construction and increases house prices. Flood-safe building standards reduce flood damages by 3% of house value, but these reductions are not privately valued by consumers. Using a model to interpret these results, we find that floodplain regulation in Florida reduces damages by 62% and achieves large social welfare gains.

Our results stand in contrast to common critiques of the National Flood Insurance Program, which argue that it promotes inefficiently high development in risky areas (e.g. Ben-Shahar and Logue (2016)). We find instead that the *status quo* policy leads to *approximately efficient* levels of flood risk and achieves substantial social welfare gains. Despite these gains, however, there is scope for further improvement, achieved by simple remappings of regulated areas that implement over 90% of first-best social welfare gains.

Our results come with caveats and directions for future work. Our model is static, and regulating durable construction in the face of evolving flood risk may involve additional unmodeled considerations. And while we provide evidence suggesting that consumers do not internalize flood risk, more work is needed to precisely identify the source of this friction.

Our findings have important implications for policies to promote the resilience of cities in the face of sea level rise and other climate-change-induced increases in flooding. Policy proposals (e.g. for home buyouts) focus largely on the benefit of reduced risk. However, as we highlight, equivalent risk reductions can entail substantially different costs. When designed to balance these benefits and costs, even simple, existing policy instruments can promote efficient adaptation to climate change.

3.8 References

_ **and Dmitry Taubinsky**, "Evaluating Behaviorally Motivated Policy: Experimental Evidence from the Lightbulb Market," *American Economic Review*, August 2015, *105* (8), 2501–2538.

Almagro, Milena, Eric Chyn, and Bryan Stuart, "Urban Renewal and Inequality: Evidence from Chicago's Public Housing Demolitions," NBER Working Paper No. 30838 January 2023.

_ **,** _ **, and** _ , "Estimating the Economic Value of Zoning Reform," NBER Working Paper No. 29440 March 2023.

Annan, Francis and Wolfram Schlenker, "Federal Crop Insurance and the Disincentive to Adapt to Extreme Heat," *American Economic Review*, May 2015, *105* (5), 262–266.

_ **and** _ , "Going Underwater? Flood Risk Belief Heterogeneity and Coastal Home Price Dynamics," *The Review of Financial Studies*, August 2022, *35* (8), 3666–3709.

_ , "In Harm's Way? Infrastructure Investments and the Persistence of Coastal Cities," February 2024.

_ **, Cristobal Otero, and Sebastian Otero**, "Equilibrium Effects of Food Labeling Policies," *Econometrica*, 2023, *91* (3), 839–868.

_ **and** _ , "The Microgeography of Housing Supply," *Journal of Political Economy*, October 2023, pp. 000–000.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan, "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, August 2007, *115* (4), 588–638.

Baylis, Patrick and Judson Boomhower, "Mandated vs. Voluntary Adaptation to Natural Disasters: The Case of U.S. Wildfires," Technical Report w29621, National Bureau of Economic Research, Cambridge, MA December 2021.

_ **and** _ , "The Economic Incidence of Wildfire Suppression in the United States," *American Economic Journal: Applied Economics*, January 2023, *15* (1), 442–73.

Beltrán, Allan, David Maddison, and Robert J R Elliott, "Is Flood Risk Capitalised Into Property Values?," *Ecological Economics*, April 2018, *146*, 668–685.

Ben-Shahar, Omri and Kyle D. Logue, "The Perverse Effects of Subsidized Weather Insurance," *Stanford Law Review*, 2016, *68*, 571–628.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, July 1995, *63* (4), 841.

Bradt, Jacob T., Carolyn Kousky, and Oliver E.J. Wing, "Voluntary purchases and adverse selection in the market for flood insurance," *Journal of Environmental Economics and Management*, October 2021, *110*, 102515.

Burby, Raymond J., "Flood insurance and floodplain management: the US experience," *Environmental Hazards*, January 2001, *3* (3), 111–122.

Burchfield, M., H. G. Overman, D. Puga, and M. A. Turner, "Causes of Sprawl: A Portrait from Space," *The Quarterly Journal of Economics*, May 2006, *121* (2), 587–633.

Calder-Wang, Sophie, "The Distributional Impact of the Sharing Economy: Evidence from New York City," December 2021.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs: Robust Nonparametric Confidence Intervals," *Econometrica*, November 2014, *82* (6), 2295–2326.

Coate, Stephen, "Altruism, the Samaritan's Dilemma, and Government Transfer Policy," *American Economic Review*, March 1995, *85* (1), 46–57.

Davlasheridze, Meri and Qing Miao, "Does Governmental Assistance Affect Private Decisions to Insure? An Empirical Analysis of Flood Insurance Purchases," *Land Economics*, February 2019, *95* (1), 124–145.

de Moel, H., J. van Alphen, and J. C. J. H. Aerts, "Flood maps in Europe - methods, availability and use," *Natural Hazards and Earth System Sciences*, March 2009, *9* (2), 289–301.

Dell, Melissa, "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 2010, *78* (6), 1863–1903.

Deryugina, Tatyana and Barrett Kirwan, "Does the Samaritan's Dilemma Matter? Evidence from U.s. Agriculture," *Economic Inquiry*, 2018, *56* (2), 983–1006.

Emrath, Paul, "Government Regulation in the Price of a New Home: 2021," May 2021.

First Street Foundation, "The First National Flood Risk Assessment," 2020.

_ , "The Cost of Climate: America's Growing Flood Risk," February 2021.

Frank, Thomas, "Removing 1 Million Homes from Flood Zones Could Save $1 Trillion," *E&E News*, April 2020.

Gallagher, Justin, "Learning about an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States," *American Economic Journal: Applied Economics*, July 2014, *6* (3), 206–233.

Gibson, Matthew and Jamie T. Mullins, "Climate Risk and Beliefs in New York Floodplains," *Journal of the Association of Environmental and Resource Economists*, November 2020, *7* (6), 1069–1111.

Golnaraghi, Maryam, Neil Dufty, and Andrew Dyer, "Flood Risk Management in Australia.," 2020.

Hino, Miyuki and Marshall Burke, "The effect of information about climate risk on property values," *Proceedings of the National Academy of Sciences*, April 2021, *118* (17), e2003374118.

Hsiao, Allan, "Sea Level Rise and Urban Adaptation in Jakarta," November 2023.

Indaco, Agustin, Francesc Ortega, and Suleyman Taspinar, "The Effects of Flood Insurance on Housing Markets," *Cityscape*, 2019, *21* (2), 129–156.

Ito, Koichiro and James M. Sallee, "The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel Economy Standards," *The Review of Economics and Statistics*, May 2018, *100* (2), 319–336.

Jacobsen, Mark R., Christopher R. Knittel, James M. Sallee, and Arthur A. van Benthem, "The Use of Regression Statistics to Analyze Imperfect Pricing Policies," *Journal of Political Economy*, May 2020, *128* (5), 1826–1876.

Jacobsen, Mark R, James M Sallee, Joseph S Shapiro, and Arthur A van Benthem, "Regulating Untaxable Externalities: Are Vehicle Air Pollution Standards Effective and Efficient?," *The Quarterly Journal of Economics*, August 2023, *138* (3), 1907–1976.

Kellogg, Ryan, "Output and attribute-based carbon regulation under uncertainty," *Journal of Public Economics*, October 2020, *190*, 104246.

_ , Erwann O. Michel-Kerjan, and Paul A. Raschky, "Does federal disaster assistance crowd out flood insurance?," *Journal of Environmental Economics and Management*, January 2018, *87*, 150–164.

_ , Howard Kunreuther, Brett Lingle, and Leonard Shabman, "The Emerging Private Residential Flood Insurance Market in the United States," *Risk Management and Decision Processes Center Working Paper*, July 2018, p. 53.

Kydland, Finn E. and Edward C. Prescott, "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, June 1977, *85* (3), 473–491.

Landry, Craig E., Dylan Turner, and Daniel Petrolia, "Flood Insurance Market Penetration and Expectations of Disaster Assistance," *Environmental and Resource Economics*, June 2021, *79* (2), 357–386.

Lee, Seunghoon, "Adapting to Natural Disasters through Better Information: Evidence from the Home Seller Disclosure Requirement," 2022.

Lingle, Brett and Carolyn Kousky, "Florida's Private Residential Flood Insurance Market," September 2018.

Listokin, David and David B. Hattis, "Building Codes and Housing," *Cityscape*, 2005, *8* (1), 21–67.

Manson, Steven, Schroeder, Jonathan, David Van Riper, Tracy Kugler, and Steven Ruggles, "National Historical Geographic Information System: Version 16.0," 2021.

Mulder, Philip, "Mismeasuring Risk: The Welfare Effects of Flood Risk Information," November 2021.

National Research Council, *Mapping the Zone: Improving Flood Map Accuracy*, Washington, D.C.: National Academies Press, May 2009. Pages: 12573.

_ **and Jonathan B. Scott**, "Does the National Flood Insurance Program Drive Migration to Higher Risk Areas?," *Journal of the Association of Environmental and Resource Economists*, March 2024, *11* (2), 287–318.

Pigou, A. C., *The Economics of Welfare*, London: Macmillan and Co., Limited, 1920.

Sastry, Parinitha, "Who Bears Flood Risk? Evidence from Mortgage Markets in Florida," November 2021.

Small, Kenneth A. and Harvey S. Rosen, "Applied Welfare Economics with Discrete Choice Models," *Econometrica*, January 1981, *49* (1), 105.

Song, Jaehee, "The Effects of Residential Zoning in U.S. Housing Markets," November 2021.

Sun, Liyang, "EVENTSTUDYINTERACT: Stata module to implement the interaction weighted estimator for an event study," 2021.

_ **and Sarah Abraham**, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, December 2021, *225* (2), 175–199.

Turner, Matthew, Andrew Haughwout, and Wilbert van der Klaauw, "Land Use Regulation and Welfare," *Econometrica*, 2014, *82* (4), 1341–1403.

U. S. Office of Inspector General, "FEMA Needs to Improve Management of its Flood Mapping Program," September 2017.

US Census Bureau, "Historical Population Change Data (1910-2020)," 2022. Section: Government.

Wagner, Katherine R. H., "Adaptation and Adverse Selection in Markets for Natural Disaster Insurance," *American Economic Journal: Economic Policy*, August 2022, *14* (3), 380–421.

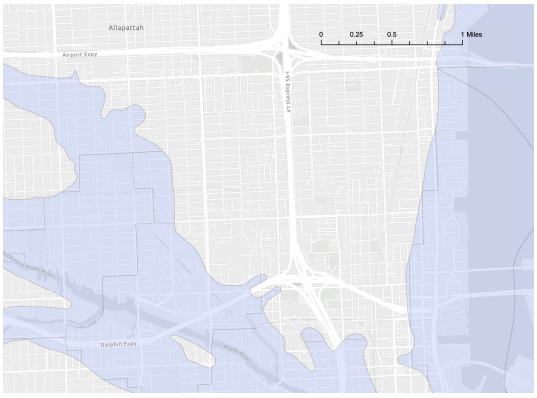
Wing, Oliver E. J., Paul D. Bates, Andrew M. Smith, Christopher C. Sampson, Kris A. Johnson, Joseph Fargione, and Philip Morefield, "Estimates of present and future flood risk in the conterminous United States," *Environmental Research Letters*, February 2018, *13* (3), 034023.

_ **, William Lehman, Paul D. Bates, Christopher C. Sampson, Niall Quinn, Andrew M. Smith, Jeffrey C. Neal, Jeremy R. Porter, and Carolyn Kousky**, "Inequitable patterns of US flood risk in the Anthropocene," *Nature Climate Change*, February 2022, *12* (2), 156–162.

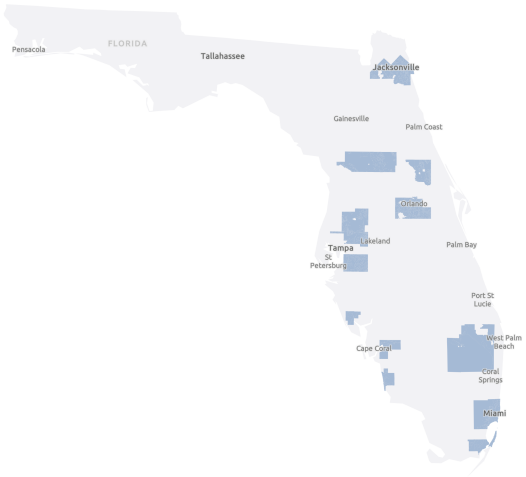
3.9 Figures and Tables

Figure 3.1: Digitized Flood Maps

(a) Digitized Map of Miami (1978)



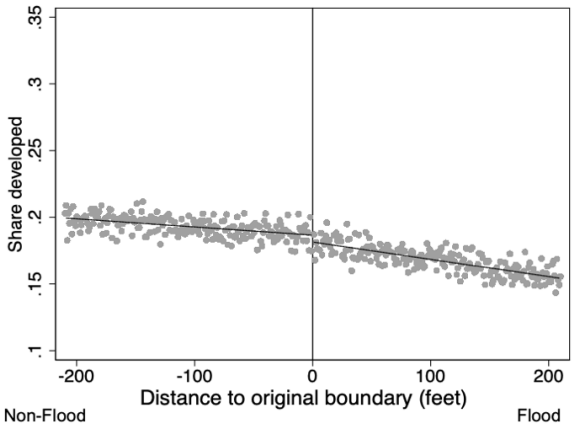
(b) Coverage of Digitized Maps



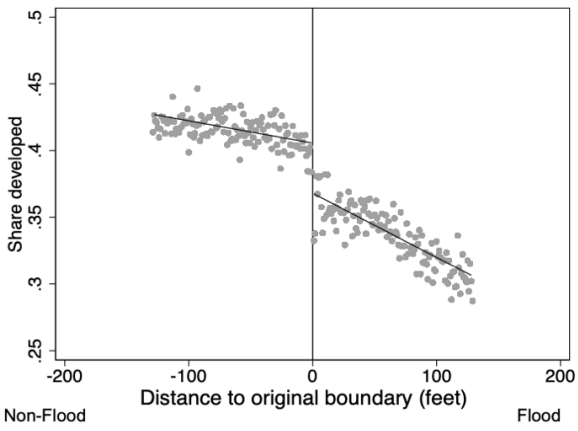
Notes: Figure (a) shows an example of a historical flood map in Miami. Flood zones are depicted in blue. Figure (b) displays the coverage of our digitized sample.

Figure 3.2: Spatial Regression Discontinuity Estimates: Development

(a) Pre-Period Share Developed (about 1980)



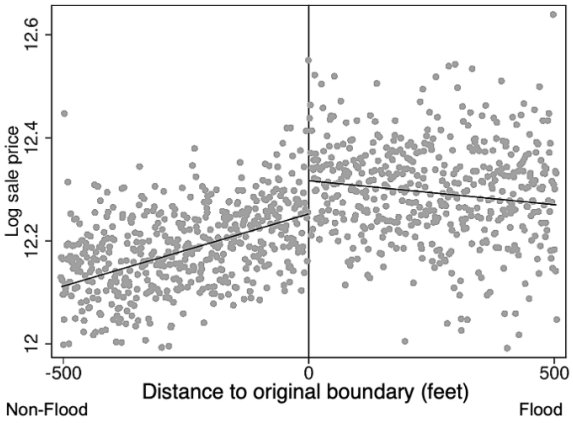
(b) Share Developed in 2016



Notes: Figures present spatial regression discontinuity plots with a local linear fit on either side of the flood zone boundary (equation 3.1). Distance to boundary is measured in feet, with positive distance indicating being inside the flood zone. Sub-figure (a) plots the share of land that was developed as of the late 1970s and early 1980s, and sub-figure (b) plots the share of land developed as of 2016. Plotted points are binned averages of grid-cell-level observations. Estimates are residualized of census tract fixed effects.

Figure 3.3: Spatial Regression Discontinuity Estimates: Prices

(a) Log Single Family Sale Price (2005-2020)

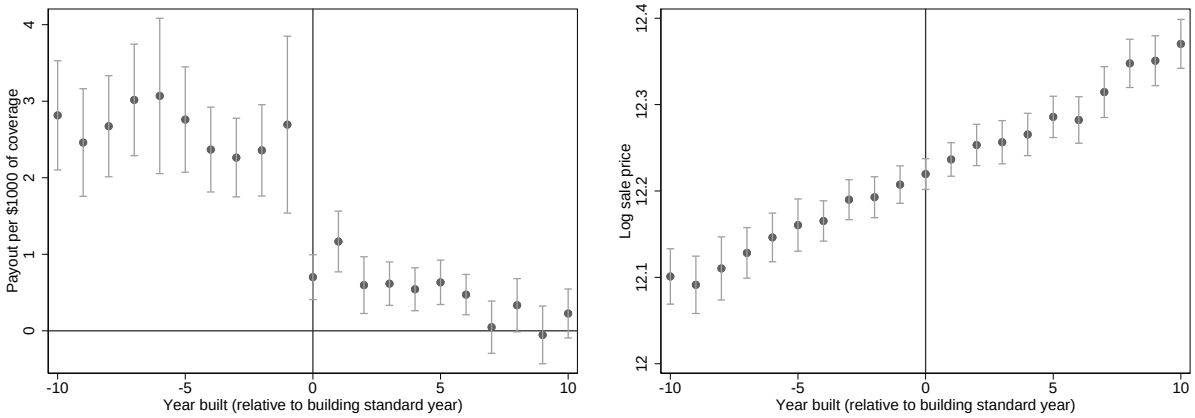


Notes: Figure presents spatial regression discontinuity plot with a local linear fit on either side of the flood zone boundary (equation 3.1). Distance to boundary is measured in feet, with positive distance indicating being inside the flood zone. Outcome is log sales prices of arms-length sales for single-family homes that sold between 2005 and 2020. Plotted points are binned averages of grid-cell-level observations. Estimates are residualized of census tract fixed effects.

Figure 3.4: Event Study Estimates: The Effect of Building Standards

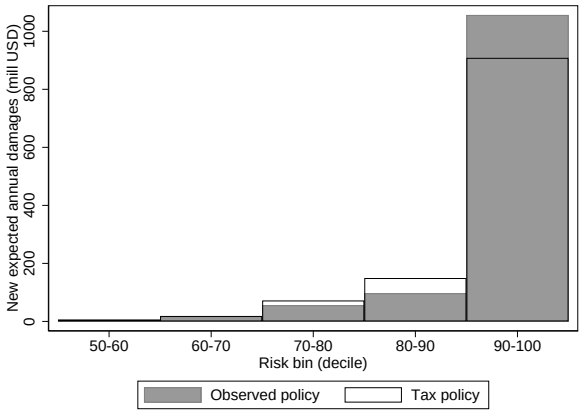
(a) Damages

(b) Log Sale Price



Notes: Figures present coefficients on bins of year built relative to the year of a community's enrollment in the National Flood Insurance Program (at which time building standards began to be imposed on newly-constructed housing), from equation 3.2. Sample is restricted to single-family residences inside the flood zone. Sub-figure (a) shows insurance payouts from 2010 to 2018 as a share of total dollars of coverage (estimated using the specification in line 1 of equation 3.2). Sub-figure (b) shows the log sale price in 2010 dollars between 2005 and 2020 (estimated using the specification in line 2 of equation 3.2). Coefficients in sub-figure (a) are estimated at the policy-year by census-tract by year-built level, weighted by the number of policies. Coefficients in sub-figure (b) are estimated at the house level, unweighted. Standard errors are clustered at the census tract level.

Figure 3.5: Expected new annual damages by risk bin: *status quo* vs. first-best tax policy



Notes: Figure plots total expected annual damages of newly-constructed houses by risk bin under the observed policy and counterfactual first-best corrective tax policy. Sample includes both flood zone and non-flood-zone land.

Table 3.1: Summary Statistics

	Florida	Digitized map sample		Boundary sample	
		Outside historic SFHA	Inside historic SFHA	Outside historic SFHA	Inside historic SFHA
	(1)	(2)	(3)	(4)	(5)
Panel A. Development					
Share developed in 1980	0.056	0.116	0.124	0.243	0.178
Share developed in 2016	0.145	0.313	0.243	0.473	0.313
Single family homes	5,175,979	552,230	191,507	159,607	80,218
Single family share of structures	0.662	0.749	0.739	0.734	0.747
Share post-FIRM	0.616	0.606	0.520	0.478	0.429
Panel B. Other characteristics					
Share wetlands	0.348	0.205	0.449	0.101	0.343
Share water	0.069	0.015	0.163	0.011	0.098
Distance to coast (miles)	10.7	7.9	6.4	7.7	7.9
Panel C. House Prices (Median)					
Price (1980)	47,459	48,198		46,912	
Price (2005-2020)	167,233	177,673	274,065	188,148	252,935
Price (single family, 2005-2020)	182,452	179,419	298,716	190,204	277,492
Panel D. Risk					
FEMA flood maps					
Land share in SFHA as of 1996	0.345	0.040	0.739	0.061	0.707
Land share in SFHA as of 2017	0.450	0.106	0.720	0.130	0.664
First Street risk measures					
Land share w. $\geq$ 1% risk of flooding	0.447	0.425	0.582	0.240	0.573
Land share w. substantial flood risk	0.153	0.065	0.225	0.061	0.223
Total area (square miles)	58,257	4,169	1,978	746	553

Notes: Table displays summary statistics for the entire state of Florida, the geographic area covered by the digitized flood maps, and a sample restricted to 2,000 feet on either side of the flood zone boundary. Median house price in 1980 is a population-weighted census tract average of 1980 Census estimates of the average value of owner-occupied single family housing, tabulated in 1980 dollars. Median house price (2005-2020) tabulates the median sales price, in 2010 dollars, of houses sold between 2005 and 2020. For houses sold multiple times, we take the average transaction price across sales. House prices from 2005 to 2020 are derived from administrative sales records from the state of Florida and are restricted to arms length sales. Substantial flood risk is defined as areas with an estimated 100-year flood depth above two feet.

Table 3.2: Spatial Regression Discontinuity Estimates

		Local linear		Fourth-order polynomial	Local linear
	Outside flood zone mean	Triangular kernel, optimal bandwidth	Rectangular kernel, constant bandwidth		Excl. coastal areas
	(1)	(2)	(3)	(4)	(5)
Panel A. Historical land use					
Share (of land) developed in 1980	0.189	-0.005	-0.004	-0.003	-0.005
		(0.002)	(0.002)	(0.002)	(0.002)
		542,302	642,720	3,027,326	543,594
Panel B. Modern land use					
Share (of land) developed in 2016	0.411	-0.038	-0.044	-0.042	-0.035
		(0.005)	(0.005)	(0.005)	(0.005)
		323,362	642,720	3,027,326	277,662
Share covered by a building footprint	0.238	-0.025	-0.026	-0.029	-0.023
		(0.004)	(0.004)	(0.004)	(0.004)
		382,470	642,720	3,027,326	332,234
Share covered by a single family home	0.090	-0.010	-0.013	-0.015	-0.009
		(0.003)	(0.002)	(0.003)	(0.003)
		371,202	642,720	3,027,326	309,866
Share covered by wetlands	0.145	0.072	0.096	0.110	0.075
		(0.006)	(0.007)	(0.007)	(0.007)
		221,336	642,720	3,027,326	187,032
Panel C. Prices					
Log house price	12.278	0.065	0.063	0.064	0.052
		(0.019)	(0.021)	(0.022)	(0.023)
		59,341	35,493	186,171	44,298
Log single-family house price	12.241	0.064	0.060	0.055	0.058
		(0.015)	(0.019)	(0.018)	(0.019)
		52,605	27,984	146,989	28,263

Notes: Table displays estimates of equation 3.1. Outside-of-flood-zone means are calculated within 50 feet of the boundary. Column 2 estimates the MSE-optimal RD bandwidth from Calonico et al. (2014) and fits a local linear regression within that bandwidth using a triangular kernel. Column 3 estimates linear regressions separately on either side of the cutoff, with each point equally weighted within 250 feet of the boundary. Column 4 estimates a fourth order polynomial separately on either side of the boundary, restricted to a window of 2,000 feet on either side of the boundary. Column 5 replicates Column 2, but excluding land less than one mile from the coast. All discontinuities are estimated on the historic boundaries and exclude boundaries that trace a body of water. Observations are grid cells. Robust standard errors (in parentheses) are clustered at the census tract level. Sample sizes are included below each standard error.

Table 3.3: The Effect of Building Standards: Event Study Coefficient Estimates

	Mean in flood zone, Pre-FIRM	Event study
	(1)	(2)
Variable		
Insurance payouts (per $1000 of coverage)	$2.93	$-1.60
		(0.40)
		225,998
Log house price (sold 2005-2020, in 2010 $)	12.21	-0.007
		(0.009)
		173,458

Notes: Table presents variable means and coefficient estimates from equation 3.3 on insurance payouts and log house prices. Insurance payout data come from residential National Flood Insurance Program policies from 2010-2018. Price data come from residential sales prices in 2005-2020. Coefficients with insurance payouts as the outcome are estimated at the policy-year by census-tract by year-built level, weighted by the number of policies. Coefficients with log house price as the outcome are estimated at the house level. Sample includes all single-family residences in Florida. Standard errors (in parentheses) are clustered at the census tract level. Sample size of each regression is listed below standard errors.

Table 3.4: Selected Parameter Estimates

Supply	
Cost of adaptation (ψ)	0.243
	(0.037)
Demand	
Coefficient on SFHA (ϕ)	-0.367
	(0.097)
Coefficient on log price (α^D)	-1.344
	(0.222)
WTP to avoid SFHA (ϕ/α^D)	0.273

Notes: Table presents coefficient estimates in household preferences and housing supply, estimated via two-step GMM (see Appendix 3.10.3 for more details). Standard errors are in parentheses.

Table 3.5: Counterfactuals: Housing Market Outcomes

Outcome	Level	Change Relative to Unregulated Counterfactual		
	Unregulated	Current Policy	Corrective Tax	Improved Targeting
	(1)	(2)	(3)	(4)
New dev on land regulated in current policy				
Approximate N Houses	1,047,291	-181,453	-78,899	-68,890
		-17.3%	-7.5%	-6.6%
New dev on all land				
Approximate N Houses	4,736,719	0	0	0
Number of Adapted Houses	302,349	563,489	198,556	206,070
		186%	66%	68%
Per-house NPV of adaptation-based damages	$14,216	$-7,298	$-7,006	$-7,006
(locations held at unreg. benchmark)		-51.3%	-49.3%	-49.3%
Per-house NPV of all damages	$14,216	$-8,737	$-9,032	$-8,619
(incorporate counterf. locations)		-61.5%	-63.5%	-60.6%
		563,489	198,511	206,091
Price				
Inside observed SFHA	$255,002	$7,112	$-9,564	$-2,061
		2.8%	-3.8%	-0.8%
Outside observed SFHA	$166,522	$12,251	$3,622	$3,199
		7.4%	2.2%	1.9%
Overall	$186,841	$9,097	$-182	$1,242
		4.9%	-0.1%	0.7%

Notes: Table presents estimates of counterfactual outcomes using baseline parameter estimates. Outcomes for the observed policy, corrective tax, and improved targeting counterfactuals are presented in differences from the modeled outcomes under the unregulated benchmark. Percentage differences from the unregulated benchmark are presented below the level differences. N houses is defined by assuming that each developed pixel is equivalent to one house. "Land regulated in current policy" is the area designated as the flood zone in the original flood maps. "Per-house NPV of all damages" is the net present value of damages for newly-built houses divided by the number of all newly-built houses. Areas in our sample counties without digitized historic flood maps are assigned their flood zone status as of 1996, which we estimate overlaps with historic flood zone status in 90% of cases. Prices are weighted by total developed area. The unregulated counterfactual sets $SFHA_z = 0$ everywhere. The corrective tax counterfactual sets $SFHA_z = 0$ everywhere and imposes taxes equal to expected flood damages, conditional on socially-optimal adaptation choices. The improved targeting counterfactual sets $SFHA_z = 1$ only in the locations where it is socially-optimal to adapt. "Adaptation-based damages" compute expected damages under counterfactual adaptation decisions, holding location constant at location choice as modeled under the unregulated benchmark. "All damages" computes damages under counterfactual adaptation and location decisions.

Table 3.6: Counterfactuals: Social Welfare

Outcome (Millions of $)	Level	Change Relative to Unregulated Counterfactual				
		Non-financial demand cost not welfare-relevant			Non-financial demand cost welfare-relevant	
	Unregulated	Current Policy	Corrective Tax	Improved Targeting	Current Policy	Improved Targeting
	(1)	(2)	(3)	(4)	(5)	(6)
Producer Surplus		13,704	4,956	5,513	13,704	5,513
Consumer Surplus		-39,398	-32,163	-14,147	-68,284	-25,221
Expected Flood Damages	68,161	-42,211	-43,577	-41,631	-42,211	-41,631
Government Revenue	3,415	11,519	21,778	2,858	11,519	2,858
Social Welfare		28,036	38,148	35,855	-850	24,781

Notes: Table presents estimates of counterfactual outcomes using baseline parameter estimates. Outcomes are in millions of 2010 dollars. The unregulated counterfactual sets $SFHA_z = 0$ everywhere. Column 1 presents levels under theunregulated benchmark (excluded for producer and consumer surplus). Columns 2-4 present results under the assumption that the magnitude of ϕ in excess of the financial costs of SFHA designation (via increased premiums) is not welfare-relevant. Columns 5-6 present results under the alternative assumption that it is. See main text for a discussion. Columns 2 and 5 present the social welfare effects of the observed policy. Column 3 presents the social welfare effects of the first-best corrective tax counterfactual, which sets $SFHA_z = 0$ everywhere and imposes taxes equal to expected flood damages; homes are adapted when it is socially-optimal to do so. Columns 4 and 6 describe a limited version of the current policy which is better-targeted: only locations where the social benefits of adaptation exceed the costs are designated as SFHAs. Government revenue indicates revenue from insurance premiums and taxes.

3.10 Appendix

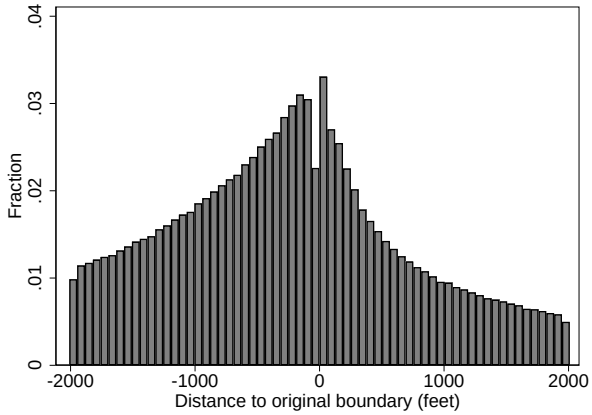
3.10.1 Appendix Tables and Figures

Figure C.1: Building Satisfying Adaptation Requirements in Naples, FL

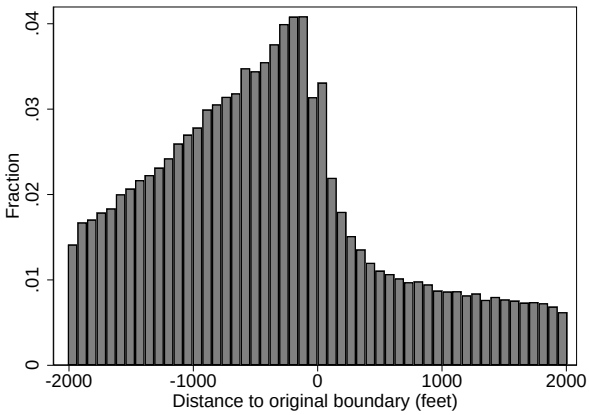
Notes: Figure shows a flood-safe house in Collier County, Florida. At this location, flood zone regulations require the bottom of the lowest (non-basement) floor to be elevated to 10 feet above sea level.

Figure C.2: Histogram of Distance to Flood Zone Boundary

(a) Land (30x30m pixels)



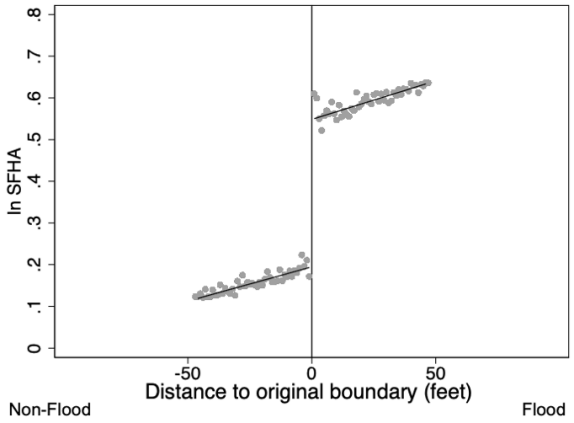
(b) Sales



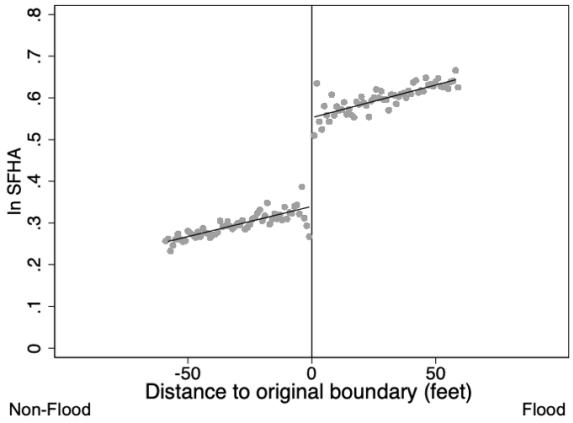
Notes: Figure presents histograms of distance to boundary. Sub-figure (a) shows land observations, at the grid-cell level, and sub-figure (b) shows home sales observations. Distance to boundary is in feet, with positive distance indicating being inside the flood zone. Excludes boundaries that trace a body of water and pixels that overlap with the boundary (leading to the dip at zero).

Figure C.3: Spatial Regression Discontinuity Estimates: Current Flood Zone Status

(a) Share of Land in 1996 Flood Zone

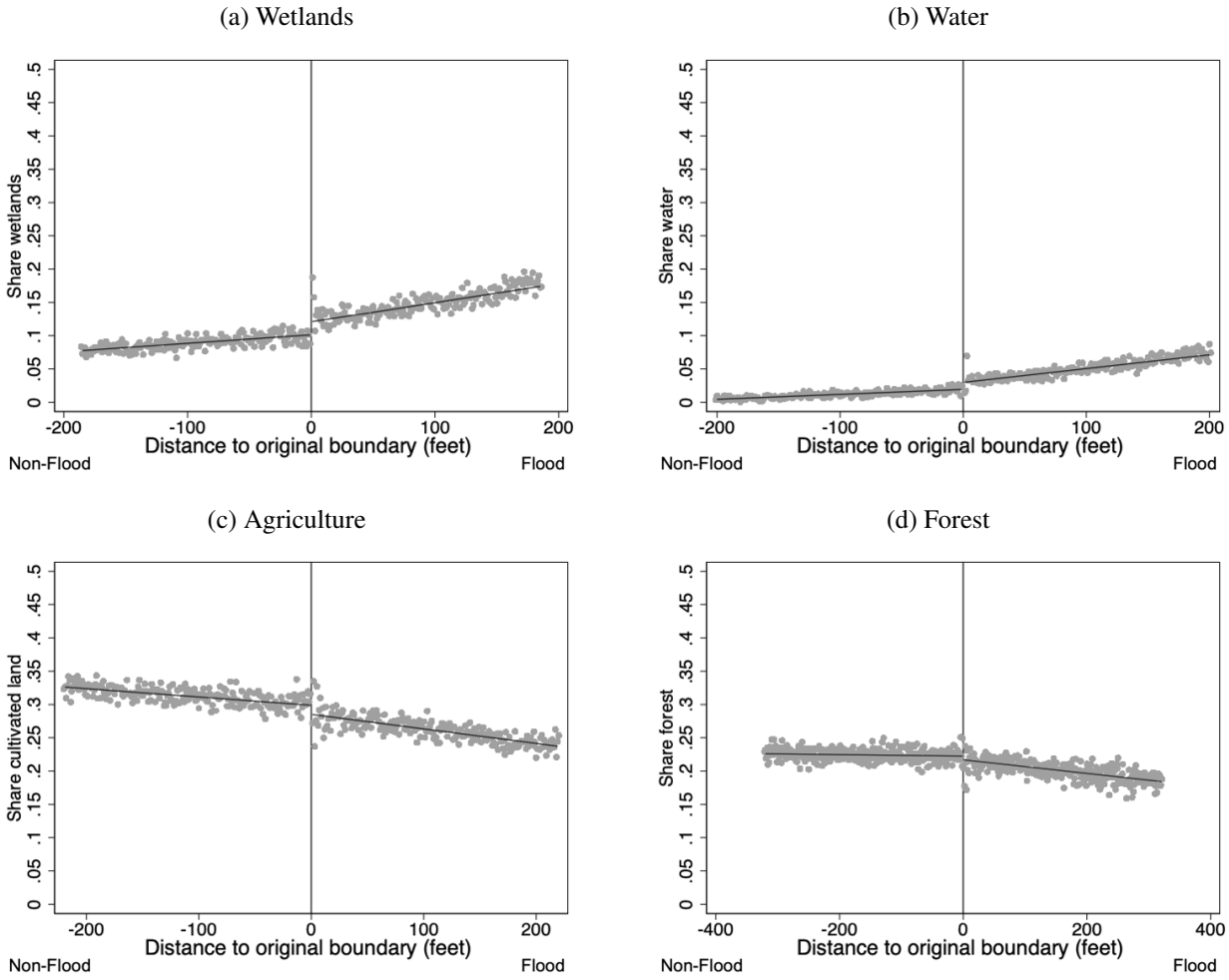


(b) Share of Land in 2017 Flood Zone



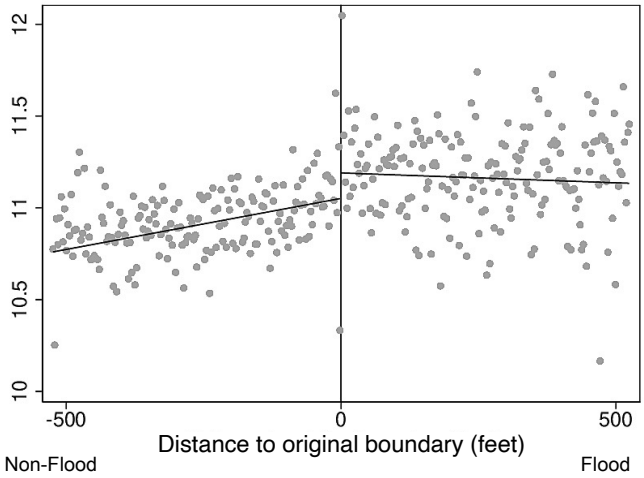
Notes: Figure presents spatial regression discontinuity plots with a local linear fit on either side of the flood zone boundary (equation 3.1). Distance to boundary is measured in feet, with positive distance indicating being inside the flood zone. Sub-figure (a) plots the share of land in the 1996 flood zone, and sub-figure (b) plots the share of land in the 2017 flood zone. Plotted points are binned averages of grid-cell-level observations. Estimates are residualized of census tract fixed effects.

Figure C.4: Spatial Regression Discontinuity Estimates: Other Pre-Period Land Use



Notes: Figure presents spatial regression discontinuity plots with a local linear fit on either side of the flood zone boundary (equation 3.1). Distance to boundary is measured in feet, with positive distance indicating being inside the flood zone. All land use outcomes are measured as of the late 1970s and early 1980s. Plotted points are binned averages of grid-cell-level observations. Estimates are residualized of census tract fixed effects.

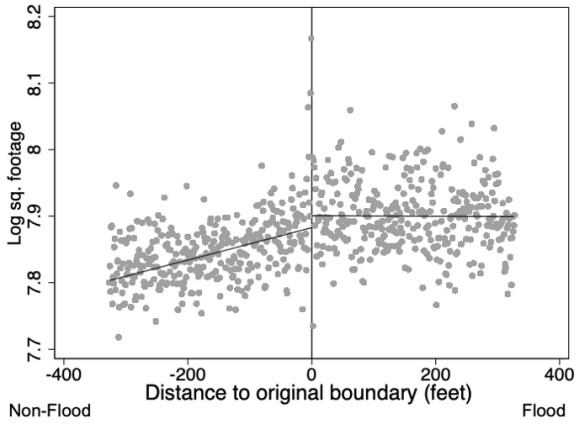
Figure C.5: Spatial Regression Discontinuity Estimates: House Price Net of Attributes



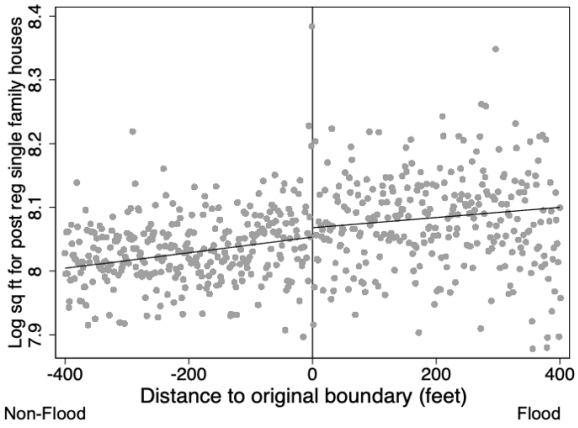
Notes: Figure presents spatial regression discontinuity plot with a local linear fit on either side of the flood zone boundary (equation 3.1). Distance to boundary is measured in feet, with positive distance indicating being inside the flood zone. Outcome is the residual from a regression of log sales prices of residential properties on polynomials in lot size and living area, indicators for construction year, and indicators for county-by-sale-month. Plotted points are binned averages of grid-cell-level observations. Estimates are also residualized of census tract fixed effects.

Figure C.6: Spatial Regression Discontinuity Estimates: Log Sq. Footage

(a) All single-family residential parcels



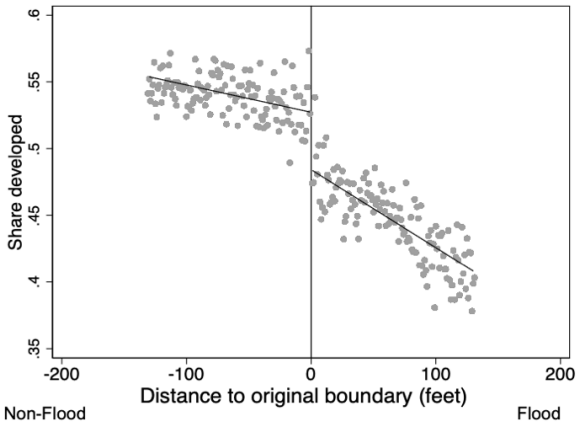
(b) Post-FIRM single-family residential parcels



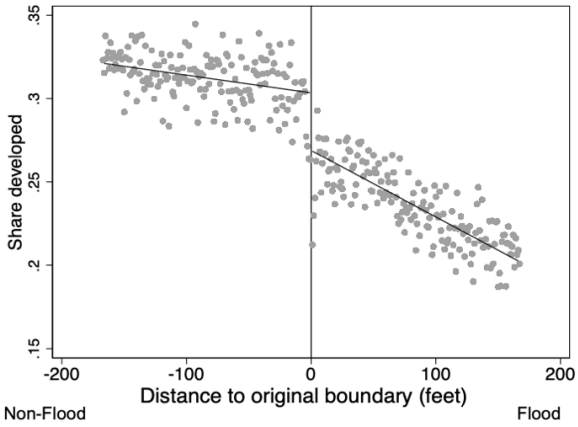
Notes: Figure presents spatial regression discontinuity plots with a local linear fit on either side of the flood zone boundary (equation 3.1). Distance to boundary is measured in feet, with positive distance indicating being inside the flood zone. Outcome is log square footage of homes in two samples: (a) single family houses and (b) and single-family houses built after the introduction of building standards (post-FIRM). Plotted points are binned averages of grid-cell-level observations. Estimates are residualized of census tract fixed effects.

Figure C.7: Spatial Regression Discontinuity Estimates: Effects By Flood Risk Level

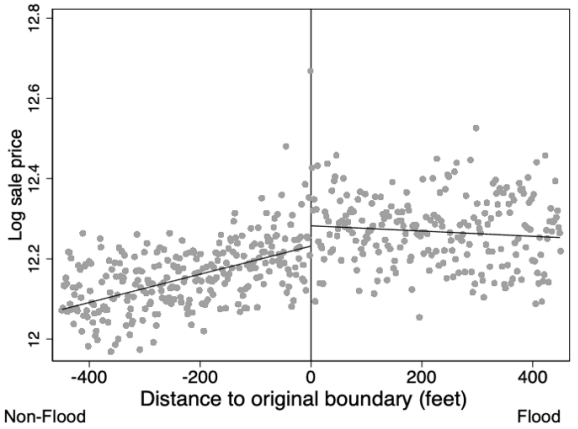
(a) Development in 2016, Low Risk



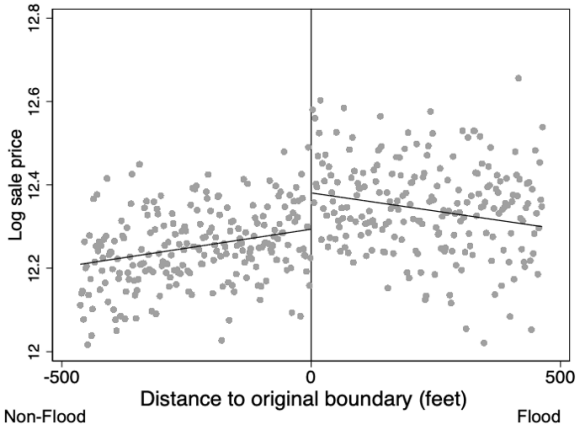
(b) Development in 2016, High Risk



(c) Log Single-Family House Price, Low Risk



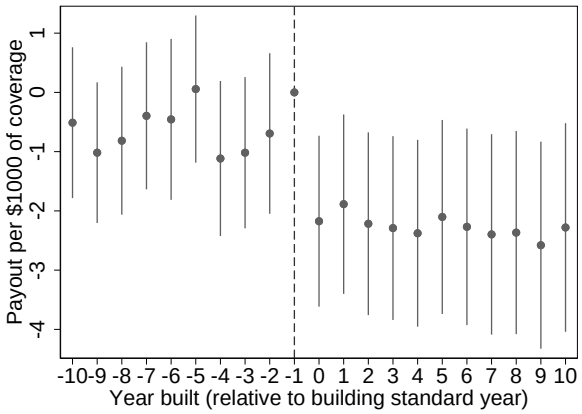
(d) Log Single-Family House Price, High Risk



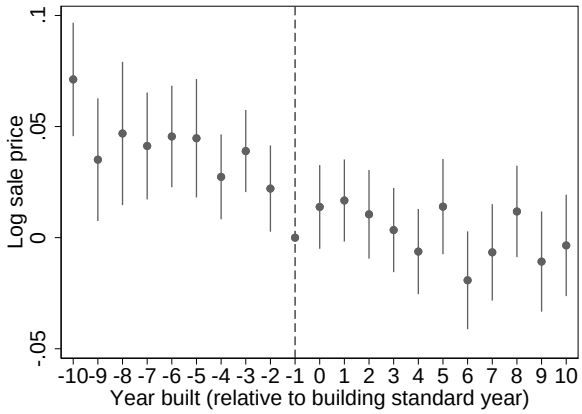
Notes: Figure presents spatial regression discontinuity plots with a local linear fit on either side of the flood zone boundary (equation 3.1). Distance to boundary is measured in feet, with positive distance indicating being inside the flood zone. Sub-figures (a) and (b) plot the share of land developed as of 2016. Sub-figures (c) and (d) plot log sales prices of arms-length sales for single-family homes with structures that sold between 2005 and 2020. Flood risk is calculated at the (1980) census tract level based on the First Street Foundation hydrological model. High-risk denotes an above-median census tract and low-risk denotes a below-median census tract in our sample. Plotted points are binned averages of grid-cell-level observations. Estimates are residualized of census tract fixed effects.

Figure C.8: Introduction of Building Standards: Difference-in-Difference Estimates

(a) Damages



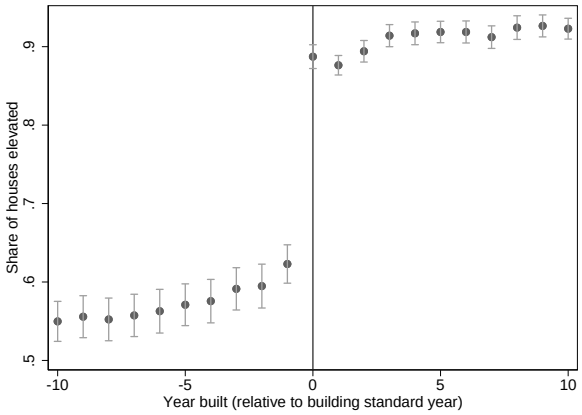
(b) Log Sale Price



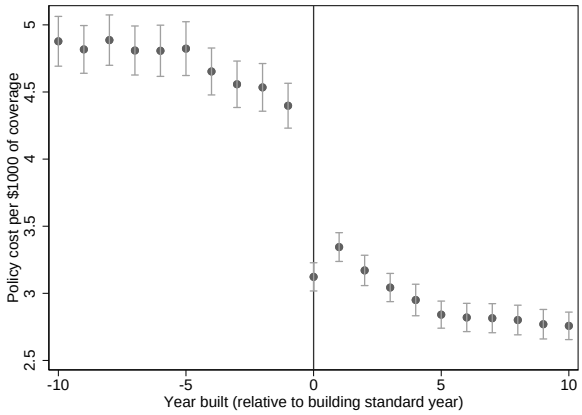
Notes: Figures present coefficients from the difference-in-difference specification comparing houses built before and after the year of building standards introduction, inside and outside of the flood zone (equation 3.12). The Sun-Abraham eventstudyinteract package is used to account for potential heterogeneity in treatment effects across cohorts (Sun, 2021). The effect of regulation is estimated as the average of the year-specific coefficient estimates from years 0-5, less the average of the year-specific estimates from years -6 to -2. Sample is restricted to single-family residences. Sub-figure (a) shows insurance payouts from 2010 to 2018 as a share of total dollars of coverage, aggregated to the policy-year by census-tract by year-built level, and weighted by number of policies. Sub-figure (b) shows the log sale price in 2010 dollars between 2005 and 2020 in each census tract, estimated at the house level. Standard errors are clustered at the census tract level.

Figure C.9: Introduction of Building Standards: Other Outcomes

(a) Share Elevated to Minimum Required Level



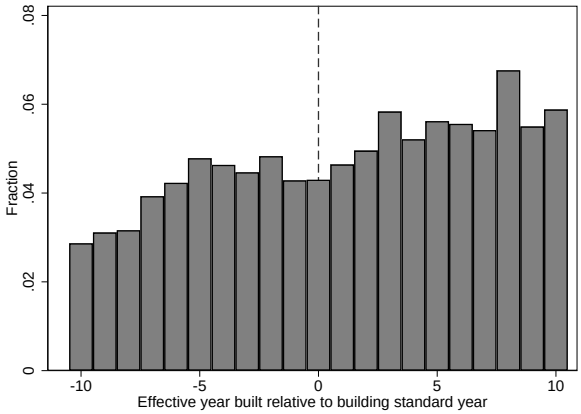
(b) Policy Cost



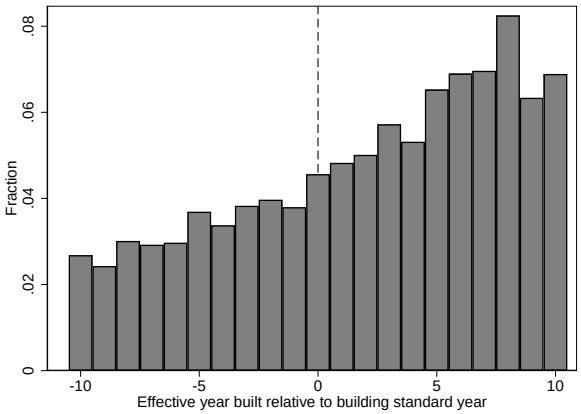
Notes: Figures present coefficients on bins of year built relative to the year of a community's enrollment in the National Flood Insurance Program (at which time building standards began to be imposed on newly-constructed housing) from equation 3.2. Sample is restricted to single-family residences inside the flood zone. Sub-figure (a) shows elevation status reported in National Flood Insurance Policies policies from 2010 to 2018, and sub-figure (b) shows policy cost from 2010 to 2018 as a share of total dollars of coverage (both estimated using the specification in line 1 of equation 3.2). Estimates are estimated at the policy-year by census-tract by year-built level, weighted by the number of policies. Standard errors are clustered at the census tract level.

Figure C.10: Introduction of Building Standards: Histogram of Effective Year Built

(a) Inside Flood Zone

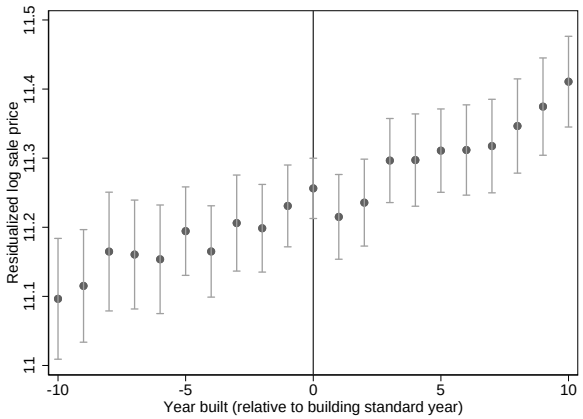


(b) Outside Flood Zone



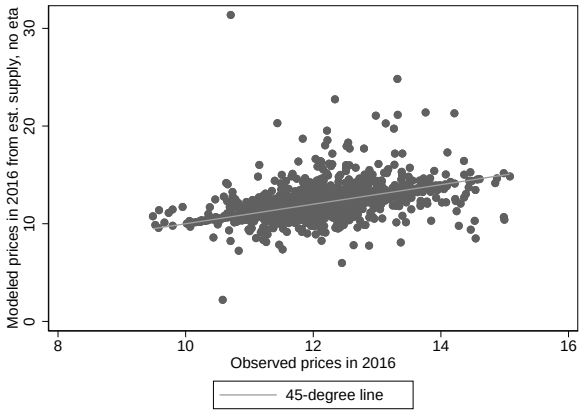
Notes: Figures present histograms of the year houses were built relative to the year building codes were introduced, both inside and outside the flood zone (SFHA).

Figure C.11: Introduction of Building Standards: Residualized Log Sales Price



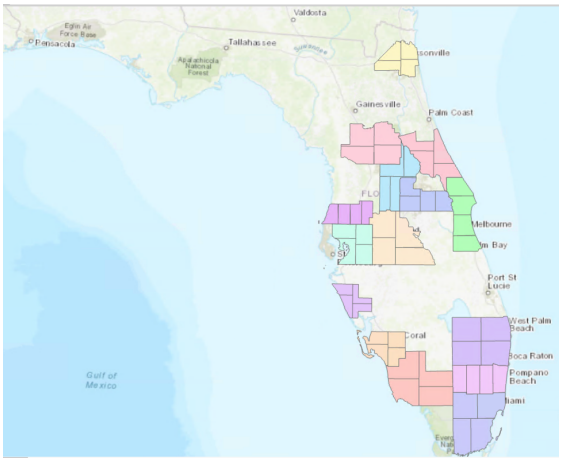
Notes: Figure presents coefficients on bins of year built relative to the year of a community's enrollment in the National Flood Insurance Program (at which time building standards began to be imposed on newly-constructed housing), estimated from equation 3.2 (line 2). Outcome is log sales price, residualized of fourth-degree polynomials in parcel size and total living area, and county-by-sale-month and year built fixed effects. Sample is restricted to properties inside the flood zone.

Figure C.12: Model Fit: Housing Supply Model Reliance on Structural Error Terms



Notes: Figure presents prices that rationalize observed development shares using just the estimated supply curve, without estimated idiosyncratic construction costs. Figure presents the prices in 2016 that would rationalize the observed quantities of development, using the estimated parameters (μ_j, σ_j, ψ) and observed adaptation decisions, if the model omitted η^t_{jz}.

Figure C.13: Areas considered for digitization



Notes: Figure depicts the top 16 counties with most development in Florida, divided into equal-area quadrants.

Table C.1: Discrepancies Between Flood Zone Status and the First Street Model

	Equal weight to each developed pixel		Weighted by number of parcels	
	Inside NFIP flood zone (SFHA)	Outside NFIP flood zone (SFHA)	Inside NFIP flood zone (SFHA)	Outside NFIP flood zone (SFHA)
	(1)	(2)	(3)	(4)
Inside First Street 100 year floodplain	0.188	0.136	0.277	0.104
Outside First Street 100 year floodplain	0.093	0.583	0.136	0.484

Notes: Table calculates the share of all buildings in our eleven counties of interest that fall into each of the four mutually exclusive categories of National Flood Insurance Program flood zone status by modern-day (2017) First Street floodplain status, in our 11 counties of interest. Both "flood zone" categories designate areas that have been determined to be in a 100 year floodplain (i.e. they have a greater than 1 percent chance of flooding per year) by either First Street or the National Flood Insurance Program. Columns (1) and (2) tabulate the share of pixels covered by a building footprint that are in each category. Columns (3) and (4) tabulate the number of parcels (accounting for multiple parcels on the same pixel).

Table C.2: Share Mapped Out of Flood Zone by Land Use

	Share of initial flood zone land that is mapped out
Any Developed	0.230
Developed - Open	0.160
Developed - Low	0.224
Developed - Mid	0.282
Developed - High	0.351
Wetlands	0.004
Water	0.003
Cultivated	0.068
Other Land Use	0.059

Notes: Table presents the share of the land inside the flood zone in 2004 that is remapped out of the flood zone in the next remapping, split by land use category in 2004. Sample is restricted to Marion and Dade Counties, because these were the two counties in our sample that experienced zero remappings between 1996 and 2004 and one remapping between 2004 and 2016. This restriction is motivated by data availability for both map updates and land use outcomes. Land use in 2004 is from the NLCD.

Table C.3: Spatial Regression Discontinuity Estimates: Other Land Use Outcomes (1980)

		Local linear		Fourth-order polynomial	Local linear
	Outside flood zone mean	Triangular kernel, optimal bandwidth	Rectangular kernel, constant bandwidth		Excl. coastal areas
	(1)	(2)	(3)	(4)	(5)
Wetlands	0.098	0.020	0.017	0.006	0.021
		(0.003)	(0.003)	(0.004)	(0.004)
		479,659	642,720	3,027,326	436,080
Water	0.017	0.011	0.010	0.004	0.010
		(0.002)	(0.002)	(0.003)	(0.002)
		518,899	642,720	3,027,326	446,653
Agriculture	0.301	-0.014	-0.014	-0.006	-0.016
		(0.003)	(0.003)	(0.004)	(0.004)
		567,993	642,720	3,027,326	429,803
Forest	0.223	-0.006	-0.006	-0.003	-0.007
		(0.003)	(0.002)	(0.003)	(0.003)
		809,732	642,720	3,027,326	656,567

Notes: Table displays estimates of equation 3.1. Outside-of-flood-zone means are calculated within 50 feet of the boundary. Column 2 estimates the MSE-optimal RD bandwidth from Calonico et al. (2014) and fits a local linear regression within that bandwidth using a triangular kernel. Column 3 estimates linear regressions separately on either side of the cutoff, with each point equally weighted within 250 feet of the boundary. Column 4 estimates a fourth order polynomial separately on either side of the boundary, restricted to a window of 2,000 feet on either side of the boundary. Column 5 replicates Column 2, but excluding land less than one mile from the coast. All discontinuities are estimated on the historic boundaries and exclude boundaries that trace a body of water. Observations are grid cells. Robust standard errors (in parentheses) are clustered at the census tract level. Sample sizes are included below each standard error.

Table C.4: Spatial Regression Discontinuity Estimates: Other Sale Price Estimates

		Local linear		Fourth-order polynomial	Local linear
	Outside flood zone mean	Triangular kernel, optimal bandwidth	Rectangular kernel, constant bandwidth		Excl. coastal areas
	(1)	(2)	(3)	(4)	(5)
Panel A. Log sale price of single-family homes					
Baseline	12.2	0.064	0.060	0.055	0.058
		(0.015)	(0.019)	(0.018)	(0.019)
		52,605	27,984	146,989	28,263
Residualized of characteristics	11.1	0.139	0.108	0.092	0.163
		(0.054)	(0.069)	(0.066)	(0.061)
		17,228	9,095	44,006	10,621
Built pre-regulations (pre-FIRM)	12.0	0.093	0.091	0.064	0.083
		(0.027)	(0.031)	(0.030)	(0.036)
		17,671	10,178	57,037	9,046
Built post-regulations (post-FIRM)	12.5	0.040	0.043	0.035	0.050
		(0.020)	(0.025)	(0.024)	(0.023)
		28,735	13,147	64,723	17,275
Panel B. Log square footage					
Single-family homes	7.9	0.018	0.019	0.024	0.014
		(0.009)	(0.009)	(0.010)	(0.012)
		64,134	49,876	263,945	38,592
Single-family homes built post-regulations	8.1	0.015	0.010	0.023	0.008
(post-FIRM)		(0.012)	(0.014)	(0.013)	(0.014)
		32,176	21,306	104,509	23,285

Notes: Table displays estimates of equation 3.1. Outside-of-flood-zone means are calculated within 50 feet of the boundary. Column 2 estimates the MSE-optimal RD bandwidth from Calonico et al. (2014) and fits a local linear regression within that bandwidth using a triangular kernel. Column 3 estimates linear regressions separately on either side of the cutoff, with each point equally weighted within 250 feet of the boundary. Column 4 estimates a fourth order polynomial separately on either side of the boundary, restricted to a window of 2,000 feet on either side of the boundary. Column 5 replicates Column 2, but excluding land less than one mile from the coast. All discontinuities are estimated on the historic boundaries and exclude boundaries that trace a body of water. Observations are grid cells. Robust standard errors (in parentheses) are clustered at the census tract level. Sample sizes are included below each standard error.

Table C.5: Summary Statistics: Building Standards Event Study Sample

Variable	Inside flood zone		Outside flood zone	
	Pre-Enrollment	Post-Enrollment	Pre-Enrollment	Post-Enrollment
	(1)	(2)	(3)	(4)
Share elevated to minimum level	0.562	0.927		
Building coverage	$188,970	$203,384	$200,232	$203,867
Contents coverage	$42,963	$51,360	$78,908	$80,500
Policy cost	$1,058	$658	$438	$433
Payout	$444	$187	$240	$222
N policy-years	598,961	618,265	236,628	391,829
House price (2010 $)	$263,253	$307,565	$166,752	$188,521
N house sales	69,976	103,620	147,663	304,083

Notes: Table presents variable means in the estimation sample for the analysis of the effect of building codes on elevation, insurance payouts, premiums, and house prices. Elevation, payout, and cost data come from residential NFIP policies from 2010-2018. Price data come from residential sales prices in 2005-2020. We use all single-family residences in Florida. Sample is restricted to houses constructed +/- 10 years around NFIP enrollment.

Table C.6: Effects of Building Standards: Heterogeneity by Flood Risk

	Low-risk			High-risk		
	Mean in flood zone, Pre-FIRM	Event study	Difference-in-Differences	Mean in flood zone, Pre-FIRM	Event study	Difference-in-Differences
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Risk Measured by First Street						
Variable						
Insurance payouts (per $1000 of coverage)	$0.74	$-0.55	$-0.66	$3.36	$-1.54	$-1.33
		(0.22)	(0.18)		(1.11)	(0.45)
N		54,237	117,302		64,569	111,251
Log house price (sold 2005-2020, 2010 $)	12.09	-0.021	-0.045	12.28	-0.010	-0.005
		(0.022)	(0.017)		(0.017)	(0.016)
N		34,832	205,023		53,902	123,547
Panel B. Risk Measured by Insurance Payouts						
Variable						
Insurance payouts (per $1000 of coverage)	$0.10	$-0.05	$-0.03	$7.62	$-4.62	$-3.95
		(0.02)	(0.02)		(1.09)	(0.75)
N		92,237	170,537		89,494	168,763
Log house price (sold 2005-2020, 2010 $)	12.29	-0.019	-0.047	12.21	-0.009	0.009
		(0.013)	(0.013)		(0.017)	(0.001)
N		94,052	224,630		54,869	218,414

Notes: Table presents variable means and coefficient estimates from Eqs. 3.3 and 3.12 on insurance payouts and log house price, split by the risk level of the census tract. Insurance payout data come from residential National Flood Insurance Program policies from 2010-2018. Price data come from residential sales prices in 2005-2020. Coefficients with insurance payouts as the outcome are estimated at the policy-year by census-tract by year built level, weighted by the number of policies. Coefficients with log house price as the outcome are estimated at the house level. The difference-in-difference estimates are estimated as the average of the year-specific coefficient estimates of Eq. 3.12 from years 0-5, less the average of the year-specific estimates from years -6 to -2. Sample includes all single-family residences in Florida. Standard errors (in parentheses) are clustered at the census tract level. In Panel A, census tracts are categorized by whether the census tract's estimated 100-year-flood depth based on the First Street hydrological model falls below (low-risk) or above (high-risk) the median in our sample. In Panel B, census tracts are categorized by whether census tracts have below (low-risk) or above (high-risk) median levels of insurance payouts, conditional on the census tract having any payout event.

Table C.7: Effects of Building Standards: Interacted with Share Elderly

Outcome: Log house price	Threshold for "elderly": 65+	75+	85+
	(1)	(2)	(3)
Post-FIRM	-0.025	-0.027	-0.015
	(0.015)	(0.014)	(0.015)
Post-FIRM x Above Median Share Elderly	0.028	0.030	0.011
	(0.019)	(0.019)	(0.019)
Difference	0.053	0.057	0.026
	(0.032)	(0.031)	(0.032)

Notes: Table presents variable means and coefficients estimated from Equation 3.3 from the event-study analyses of NFIP enrollment on house price. Price data come from residential sales prices in 2005-2020. Indicator for whether a census tract exceeds the overall median share elderly uses data from the 2014-2018 American Community Survey. Sample includes all single-family residences in Florida, restricted to the SFHA. Standard errors are clustered at the census tract level.

Table C.8: Summary Statistics: Model Estimation Sample

	Whole Sample		Balanced Boundary Sample	
	Outside Flood Zone	Inside Flood Zone	Outside Flood Zone	Inside Flood Zone
	(1)	(2)	(3)	(4)
Share developed, 1980	0.65	0.49	0.57	0.50
N developed gridcells, 1980	2,120	886	93	74
Share developed, 2016	0.84	0.68	0.81	0.73
N developed gridcells, 2016	5,831	1,964	239	178
Share elevated pre-regulation	0.00	0.39	0.00	0.45
Share elevated post-regulation	0.00	1.00	0.00	1.00
Home price, 1980	$46,428	$51,389	$49,757	$49,757
Home price, 2017	$206,984	$300,546	$238,114	$283,941
First Street Flood Risk	0.08	0.34	0.09	0.25
First Street AAL (2021)	0.0005	0.0042	0.0007	0.0020
First Street AAL (2051)	0.0013	0.0083	0.0019	0.0046
First-Street future-adjusted AAL	0.0009	0.0063	0.0013	0.0033
N gridcells	20,293	11,251	538	503
N observations	1,043	803	255	255

Notes: Table presents summary statistics of the aggregated sample at the tract-zone-boundary proximity level, used for model estimation and counterfactuals. Columns 1 and 2 describe the whole sample. Columns 3 and 4 describe the subset of the sample used for constructing the boundary regression discontinuity moments. This subset is restricted to paired inside/outside flood zone observations that are within 100 feet of a boundary. Each observation has the same weight regardless of share developed. First Street Flood Risk is the share of observations with an annual risk of flooding more than 2 feet greater than 1%. First Street AAL is the expected Average Annual Losses (as a share of parcel value), calculated using the First Street Foundation model under the "middle" scenario for future flood risk projections.

Table C.9: Estimated Parameters, Alternative Specifications

	Baseline	0 price effect	0 price effect, 1/2 quantity effect	Supply Elast + 1 SD
	(1)	(2)	(3)	(4)
Supply				
Cost of adaptation (ψ)	0.243	0.242	0.141	0.197
	(0.037)	(0.037)	(0.037)	(0.027)
Demand				
Coefficient on SFHA (ϕ)	-0.367	-0.454	-0.430	-0.374
	(0.097)	(0.095)	(0.095)	(0.096)
Coefficient on log price (α^D)	-1.344	-1.378	-1.405	-1.337
	(0.222)	(0.220)	(0.220)	(0.222)
WTP to avoid SFHA (ϕ/α^D)	0.273	0.330	0.306	0.280

Notes: Table presents parameter estimates of our main parameters under alternate assumptions. Column 1 is the baseline (reported in Table 3.4). Column 2 imposes no price effect instead of the price increase estimate used in the baseline. Column 3 additionally assesses the sensitivity to cutting our quantity estimate in half. Column 4 calibrates supply parameters using the baseline estimates of census-tract level supply elasticities plus one standard deviation of the estimates in Baum-Snow and Han (2023). Standard errors (in parentheses) generated from the analytic GMM formula.

3.10.2 Data Appendix

Selecting Counties to Digitize The historic flood maps are available online in a series of scanned images. These maps are organized first by county and then by "community," which can be as small as a village or as large as all unincorporated areas of a county. Each community is mapped in a series of tiles. Tiles vary in size and amount of land covered. Because we faced a fixed per-tile digitization cost and had a limited budget, our goal was to select the fewest number of tiles that give the most useful variation. In particular, we wanted to ensure we digitized tiles that saw substantial development between the 1980s and present day, but focused on sufficiently large areas to avoid any concerns about selecting on an endogenous outcome.

Our process for selecting maps was as follows:

Step 1. Select the top 15 counties with the largest quantity of newly-developed land, according to our digitized land use data.[41]

Step 2. Divide each county into equal-area quadrants. An illustration of the quadrants is shown in Figure C.13.

[41] We restricted to the top 15 counties because each county requires substantial effort to evaluate (manually determining the location of each tile in order to assign it to a quadrant).

Step 3. For each quadrant, compute total area of new development.

Step 4. For each quadrant, count the number of tiles that overlap it.

Step 5. For each quadrant, compute the total area of new development per tiles that would need to be mapped.

Step 6. Sort quadrants by area of new development per tile and drop quadrants with the lowest value until budget constraint is met.

The final sample included 120 tiles from 11 different counties (21 quadrants). An alternative procedure, in which we first dropped all quadrants with more than ten tiles, and then selected the quadrants with largest total area of new development, yielded a very similar set of quadrants.

Computing Distance to Boundaries We compute the distance from each grid cell to the closest point on a flood zone boundary that is not within 100 feet of the border of a body of water. We drop grid cells that fall within 2 miles of county boundaries to avoid including flood zone boundaries that overlap county boundaries.

3.10.3 NFIP Enrollment Event Study: Additional Material

Data Restrictions We restrict to residential policies on single-family homes and drop any policies whose coverage exceeds statutory caps. We measure payouts as the total claims paid out for building and contents insurance. We measure policy cost as the total of the premium and other fees. We measure elevation using an indicator of whether a house's elevation exceeds the base flood elevation (BFE).[42] By definition, this variable is not available outside the floodplain since these areas are assumed to be above the base flood elevation. Inside the floodplain, it is available in about half of pre-period (unregulated) houses and almost 100% of post-period (regulated) houses. We measure a house as elevated if the measured difference between lowest floor elevation and BFE is greater than or equal to 0. We assume that if the elevation is missing, the house is not elevated.

Determining NFIP Enrollment Year We define the year of each community's NFIP enrollment, and therefore the year in which floodplain regulations were imposed, using data on NFIP policies from 2010 to 2018. We use the fact that policies have an indicator for whether the house was built post-FIRM to construct the year of NFIP enrollment at the census tract level. Because enrollment occurred at the community level, and a community is generally larger than a census

[42]The Base Flood Elevation measures the height of the flood that has a greater than 1% chance of happening every year.

tract, characterizing year of enrollment at the census tract level is unlikely to introduce substantial inaccuracies. We define the year of enrollment as the first year within a census tract in which over 50% of homes are coded as post-FIRM. We restrict to census tracts with at least 25 distinct years of construction to avoid classifying the enrollment year based on noise.

Testing for Confounding Disamenities Section 3.4.2 finds that despite reduced flood risk, flood-safe houses are not more expensive than comparable non-adapted houses. One possible explanation for this result is that adaptation, specifically elevating to a minimum level, may introduce stairs, which could be a disamenity for homeowners. To test this, we note that elderly homeowners are most likely to find stairs a disamenity. We therefore split the analysis into tracts with above- and below-median share of elderly adults.[43] We estimate:

$$\begin{aligned} y_{jbt} = \beta_1 Post_{jb} + \beta_2 Post_{jb} \times Elderly_j + \nu_1 r_{jb} + \nu_2 r_{jb} \times Elderly_j \\ + \eta_1 r_{jb} Post_{jb} + \eta_2 r_{jb} Post_{jb} \times Elderly_j + \gamma_j + \varepsilon_{jbt} \end{aligned} \tag{3.11}$$

which replicates equation 3.3, but introduces interaction terms between all relative construction year terms and $Elderly_j$ (an indicator for above-median-elderly-share in each census tract).

Results are displayed in Table C.7. If anything, the effect of building code regulation on house prices is *higher* for more-elderly tracts, rather than lower, but the differences across age groups are not statistically-significant. We view this as suggestive that a dislike for stairs is not confounding our interpretation of our baseline event study estimates.

Difference-in-Difference Specification We expand on our event-study strategy in Section 3.4.2 using the fact that building standards were imposed for houses built inside the flood zone but not for houses built outside of it.[44] Building on equation 3.2, we estimate the following specification:

$$\begin{aligned} m_{jzbt} = \Sigma_r \beta_r \mathbb{1}\{r = b - e_j\} \times SFHA_z + \Sigma_r \gamma_r \mathbb{1}\{r = b - e_j\} + \lambda_b + \gamma_{jz} + \varepsilon_{jzbt} \\ p_i = \Sigma_r \beta_r \mathbb{1}\{r = b_i - e_i\} \times SFHA_{z(i)} + \Sigma_r \gamma_r \mathbb{1}\{r = b_i - e_i\} + \lambda_{b(i)} + \gamma_{j(i)z(i)} + \varepsilon_i \end{aligned} \tag{3.12}$$

where now z indicates flood-zone, γ_{jz} are tract-by-flood-zone fixed effects and $\lambda_{b(i)}$ indicates year-of-construction fixed effects. We cluster standard errors at the census tract level. For outcomes

[43] Using data from the 2014-2018 ACS, we compute the share of population in each census tract that is above 65, 75, or 85. We then define a tract as above-median for each age cutoff if it exceeds the median census tract share above that cutoff across the state.

[44] We use the never-treated units outside of the flood zone to avoid bias in two-way fixed effects estimators, following Sun and Abraham (2021a).

related to insurance claims and policies, we weight each location by the number of policies. Figure C.8 presents results. Results are qualitatively and quantitatively similar to the analysis in Section 3.4.2 of the main text.

Stylized Model of WTP for Adaptation Suppose that houses are either adapted (A) or non-adapted (B), with a fixed supply of each. Denote c as the (total lifetime) savings from living in an adapted house ($c < 0$ indicates savings) and ρ as the share of savings that are internalized by the home-buyer. Let p_A and p_B be the respective house prices (in levels).

Suppose $u_{iA} = \alpha(p_A + \rho c) + \varepsilon_{iA}$ and $u_{iB} = \alpha p_B$ where ε_{iA} is distributed i.i.d Type 1 Extreme Value. This specification embeds the assumption that consumers only care about adaptation through its effects on risk. The share of consumers purchasing an adapted house is

$$s_A = \frac{\exp(\alpha(p_A - p_B + \rho c))}{1 + \exp(\alpha(p_A - p_B + \rho c))}. \tag{3.13}$$

Rearranging terms:

$$\rho = \frac{1}{c}\left(\frac{1}{\alpha}\ln\left(\frac{s_A}{1 - s_A}\right) - (p_A - p_B)\right). \tag{3.14}$$

We assume that $\alpha = -1$ and take $s_A = 0.8$ based on the market share of post-FIRM houses in 2016, and set $c = -6398$ based on our estimates.

An estimate of no price difference between adapted and non-adapted houses yields an internalized share ρ of approximately zero. The upper end of the 95% confidence interval (estimated in Table 3.3) is a price increase of 1.06%, leading to an implied $\rho = 0.36$. Thus, we can reject that consumers internalize more 36% of the expected reduction in flood damages.

Model Estimation Details

Calibration of Supply Elasticities We use as our starting point estimates produced by Baum-Snow and Han (2023) (BSH) for land development elasticities at the census tract level, estimated between 2001 and 2011. The supply elasticities BSH estimate are derived from the following relationship

$$\alpha_j^S = \beta_0 + \beta_1 ShareDev_j + \beta_2 DistCBD_j + \beta_3 Flat\%_j + \varepsilon_j \tag{3.15}$$

which relates the supply elasticity to the share of the census tract that is developed, the distance from the census tract to the center business district (CBD), and the share of the census tract that is

not steep-sloped. The share of land in the census tract that is developed ($ShareDev_j$) is measured in 2001.

The fact that elasticities are estimated from 2001-2010 means that these supply elasticities will likely underestimate supply elasticities in our sample, and differentially so for census tracts that became more developed between 1980 and 2001 as more developed areas have fewer attractive plots on which to build.

We therefore adjust our estimates of α_j^S to accord with 1980s development shares as follows:

1. Estimate equation 3.15 with 2004 development shares, as well as the same $DistCBD_j$ and $Flat\%_j$ used in the original BSH estimates.

2. Predict $\hat{\alpha}_j^S$ using the estimated coefficients from (1) and our measures of 1980s development shares.

3. Replace any negative elasticities with the smallest nonnegative elasticity in our eleven-county sample.

We use the approach above instead of directly using BSH's estimates of β_1, β_2, and β_3 because we measure development slightly differently than BSH.

We then translate our adjusted measures of α_j^S into an implied tract-level μ_j and σ_j for each census tract. Again, to be consistent with the BSH estimates, we estimate μ_j and σ_j treating census tracts as uniform, i.e. not differentiating across flood zones, as this is the level at which the supply elasticities are estimated. We first compute the decrease in share developed in each census tract that would be implied by the BSH elasticities for a price decrease of 10%: $\tilde{q}_j = q_j^{2016} - \alpha_j^S(0.1p_j^{2016})$, and use:

$$\sigma_j = \frac{-0.1p_j^{2016}}{\Phi^{-1}(\tilde{q}_j) - \Phi^{-1}(q_j^{2016})} \tag{3.16}$$

$$\mu_j = p_j^{2016} - \Phi^{-1}(q_j^{2016})\sigma_j \tag{3.17}$$

to obtain census-tract level estimates of μ_j and σ_j (note that because we are matching μ_j and σ_j to the BHS estimates at the census tract level, we ignore η_{jz} and E_{jz}).

Then, we build on this matching exercise by allowing the supply of housing to differ by (1) adaptation status (ψE_{jz}), detailed in the following section, and (2) by a shifter η_{jz}, which we obtain, given estimates of ϕ, as $\eta_{jz} = p_{jz}^{2016} - \psi E_{jz} - \mu_j - \Phi^{-1}(q_{jz}^{2016})\sigma_j$.

Moment Condition Details

Moments Based on the Exogeneity of Building and Land Characteristics The moments based on the exogeneity of building and land characteristics are as follows:[45]

$$\mathbb{E}\left[\xi_{jz}X_{jz}\right]=0 \qquad \mathbb{E}\left[\xi_{jz}\tilde{X}_{jz}\right]=0 \qquad \mathbb{E}\left[\xi_{jz}\tilde{p}_{jz}\right]=0 \tag{3.18}$$

for a vector $\tilde{X}_{jz}$ that averages the observable characteristics X_{jz} of locations in the same housing market that are located more than 3 miles away from geography jz, weighted by land area, and a price vector $\tilde{p}_{jz}$ that rationalizes market shares under no unobserved amenities (i.e., setting $\xi_{jz}=0$).

Boundary Moments: Demand We construct moments to match our spatial regression discontinuity (RD) analysis around the regulatory boundaries. To operationalize this, we first subdivide the tract-zone pairs into tract-zone-band observations, where band $b \in \{close, far\}$ indicates whether an observation is within K feet from a floodplain boundary. Ideally, we would like to construct moments $E[\xi_{jzb}SFHA_z\mathbb{1}\{b=close\}]=0$, taking $K\to 0$. This captures the RD assumption that as the boundary is approached, amenities become uncorrelated with flood zone status. However, taking $K\to 0$ has practical challenges, namely a lack of sufficient observations.

We address this challenge by constructing moments that match our RD estimates in Section 3.4.1 directly. First, we define $K=100$ feet. Then, in order to capture the fact that within a 100-foot boundary, mean amenities may still differ inside and outside the flood zone, we add a boundary-SFHA fixed effect, $\Delta^D SFHA_{zb}$. We redefine unobserved amenities as $\tilde{\xi}_{jzb}=\delta_{jzb}-\left(\alpha^D p_{jzb}+\phi SFHA_z+X_{jzb}\beta+\right.$ where $\phi SFHA_z$ is the causal effect of regulation on choices and $\Delta^D SFHA_{zb}$ is the average difference in amenities between SFHA and non-SFHA locations within 100 feet of the flood zone boundary. Then, we define the following moments to exactly match the RD estimates:

$$\mathbb{E}\left[\left(\frac{1}{\alpha^D}\left(\delta_{j1b}-\delta_{j0b}-\phi-\Delta^D\right)-\beta^{p,2016}\right)\mathbb{1}\{b=close\}\right]=0 \tag{3.19}$$

$$\mathbb{E}\left[\tilde{\xi}_{jzb}SFHA_z\mathbb{1}\{b=close\}\right]=0 \tag{3.20}$$

$$\mathbb{E}\left[\tilde{\xi}_{jzb}\mathbb{1}\{b=close\}\right]=0 \tag{3.21}$$

[45] We include a constant term in X_{jz}.

where $\beta^{p,2016}$ is the RD effect on price. These moments are only calculated for observations close to the flood zone boundary where there is a balanced pair (i.e. an observation both inside and outside the SFHA in the same tract).[46]

Boundary Moments: Supply We follow a similar approach for supply-side moments, constructing moments that match our RD estimates in Section 3.4.1 directly. In order to capture the fact that within a 100-foot boundary, mean construction costs may still differ in the SFHA, versus not, we add a boundary-SFHA fixed effect, $\Delta^{S,t}SFHA_{zb}$. We redefine unobserved construction costs as $\tilde{\eta}^t_{jzb} = p^t_{j1b} - \psi E^t_{j1} - \mu_j - \Delta^{S,t}SFHA_{zb} - \mu^t_b$, where ψE^t_{j1} is the causal effect of regulation on construction costs, $\Delta^{S,t}SFHA_{zb}$ is the average difference in construction costs between SFHA and non-SFHA locations within 100 feet of the flood zone boundary, and μ^t_b is a mean shifter in the boundary sample. Then, we define the following moments to exactly match the RD estimates:

$$\mathbb{E}\left[\Phi\left(\frac{p^{1980}_{j1} - \psi E^{1980}_{j1} - \mu_j - \tilde{\eta}^{1980}_{j1} - \mu^{1980}_b}{\sigma_j}\right) - \Phi\left(\frac{p^{1980}_{j0} - \psi E^{1980}_{j0} - \mu_j - \tilde{\eta}^{1980}_{j0} - \mu^{1980}_b}{\sigma_j}\right) - \beta^{q,1980}\right] = 0 \tag{3.22}$$

$$\mathbb{E}\left[\Phi\left(\frac{p^{2016}_{j1} - \psi E^{2016}_{j1} - \mu_j - \tilde{\eta}^{2016}_{j1} - \mu^{2016}_b}{\sigma_j}\right) - \Phi\left(\frac{p^{2016}_{j0} - \psi E^{2016}_{j0} - \mu_j - \tilde{\eta}^{2016}_{j0} - \mu^{2016}_b}{\sigma_j}\right) - \beta^{q,2016}\right] = 0 \tag{3.23}$$

$$\mathbb{E}\left[\tilde{\eta}^{1980}_{jzb} SFHA_z \mathbb{1}\{b = close\}\right] = 0 \tag{3.24}$$

$$\mathbb{E}\left[\tilde{\eta}^{2016}_{jzb} SFHA_z \mathbb{1}\{b = close\}\right] = 0 \tag{3.25}$$

$$\mathbb{E}\left[\tilde{\eta}^{1980}_{jzb} \mathbb{1}\{b = close\}\right] = 0 \tag{3.26}$$

$$\mathbb{E}\left[\tilde{\eta}^{2016}_{jzb} \mathbb{1}\{b = close\} = 0\right] \tag{3.27}$$

[46] We also include a boundary sample fixed effect in X_{jzb}.

for estimated RD effects on the share of land that is developed in the pre-period, $\beta^{q,1980}$, and the post-period, $\beta^{q,2016}$. All moments are calculated for observations close to the flood zone boundary where there is a balanced pair (i.e. an observation both inside and outside the SFHA in the same tract).

Estimation and Data Details Using the moments specified above, we obtain parameter estimates with two-step optimal GMM and calculate standard errors analytically.

We measure p_{jz}^{2016} as the log of the median sales price for single-family homes from 2014-2019 based on the location of the building footprint. We measure p_{jz}^{1980} as the log of the median value of owner-occupied non-condominium housing units from the 1980 Census. House prices from 1980 are not available at the flood zone level. Because flood zones did not exist in the pre-period we assume that the price does not differ between flood zones within a Census tract. We measure quantity of developed land in 1980 and 2016 as the number of gridcells that are categorized as developed in the 1980 and 2016 land-use datasets.[47] We base our adaptation indicator on a measure of elevation from NFIP policy data, as discussed in Appendix 3.10.3. We define a tract as adapted if more than 50% of insured houses in that tract are elevated. Where we do not observe elevation (including all non-flood-zone tracts), we assume adaptation only occurred when required by regulation. When historic flood zone status is unavailable because of the limited reach of our digitized maps, we use the 1996 flood zone status to calculate market shares, but we restrict to historic maps for the boundary moments.

Appendix Table C.8 presents summary statistics for the model estimation sample. A larger share of our estimation sample is developed than the sample used in Section 3.4, but house prices and flood risk are similar.

Model Fit The structural error terms $\left(\xi_{jz}, \eta_{jz}^{t}\right)$ allow our model to achieve a perfect fit to the observed data. However, given that we calibrate μ_j and σ_j in our model of housing supply from external estimates, we would like to assess fit in our context. To do so, we investigate the extent to which we rely on these structural error terms for model fit. We conduct this exercise in Figure C.12, where we plot observed price in 2016 against the modeled prices that would rationalize observed market shares in each year, *if we omitted the structural error terms*. This exercise isolates the fit of our model of housing supply, as we calculate model-implied prices using our estimated supply parameters. Figure C.12 demonstrates a strong correlation between model-generated and observed

[47] To account for the fact that some locations jz have no developed land in 1980, we calculate share developed $q_{jz} = (Q_{jz}+1)/L_{jz}$ in those locations. Also, in the spirit of Burchfield et al. (2006) we correct for potentially mismeasured growth by measuring the number of developed cells in 1980 as the minimum of the observed number of developed cells in the location in 1980 and the number measured in 2016.

prices, indicating our supply curve is reasonable. Of course, the inclusion of the structural error terms mechanically improves the model fit.

3.10.4 Expected Damages Calculations

We define flood risk using data from the First Street Flood Lab estimates of Average Annual Loss (AAL). AAL expresses expected annual damages as a share of house price. These data come from parcel-specific estimates (as opposed to the raw hazard layer) that combine the raw hazard layer (which generates the parcel-specific inundation depth) with the output of an engineering damage model. The damage model takes as inputs a number of features of the structure, including its market value, number of stories and units, and foundation type, and calculates damages using the HAZUS-MH methodology. The HAZUS-MH methodology was developed for FEMA to calculate estimated damages from natural disasters and is based on a set of depth-damage curves collected from FEMA's Federal Insurance and Mitigation Administration (FIMA) and the USACE Institute for Water Resources (USACE-IWR).[48] Average annual loss is expected to grow over time; we assume risk increases linearly from the estimated 2021 risk to the estimated 2051 risk and then stays constant at the 2051 risk for all future years. Wherever First Street did not provide an AAL estimate but did provide a Flood Factor (another measure of risk), we assumed the AAL was 0.

Expected damages are computed as the product of number of newly-developed gridcells and the PDV of expected damage under a given counterfactual. The expected damage is computed as $0.7 \times P_{jz}^{Obs} \times AAL_{jz} \times M_{jz}^{CF}(E_{jz}^{2016})$, where P_{jz}^{Obs} is the observed (level) price of a house and AAL_{jz} is the observed average annual loss.[49] The term M_{jz}^{CF} is a multiplier that accounts for differences in adaptation in each counterfactual. We assume that if an observation was adapted in the pre-period it will be adapted in the post-period for all counterfactuals. Otherwise, houses that are not observed to be adapted in the post-period but are adapted in a counterfactual will experience 55% lower damages in that counterfactual. This is based on our estimates of the average effect of building code adoption on damages in our event study analysis in Table 3.3 and our heterogeneity analysis in Appendix Table C.6 that indicates that damage reductions are proportional to baseline risk. A similar calculation applies to houses that are observed to be adapted but are counterfactually non-adapted.

We compute two measures of damages:

$D^{All} = \frac{1.05}{.05} N^{CF} D^{CF}$

$D^{Adapt} = \frac{1.05}{.05} N^{UR} D^{CF}$

[48] See First Street Foundation (2021) for more details.

[49] This 70% factor was recommended by First Street, who provided the underlying AAL data.

where N^{CF} denotes the number of newly-developed gridcells under counterfactual CF , N^{UR} denotes the number of newly-developed gridcells under the unregulated benchmark, and D^{CF} denotes the PDV of expected damage under counterfactual CF. The first measure ("all damages") measures total expected damages by multiplying the counterfactual number of newly-developed houses in each area by the expected damages in that counterfactual. The second measure ("adaptation-based damages") holds the number of houses in each location constant at the unregulated benchmark and only changes the expected damage in each location. We then compute per-house damages by dividing the total expected damages of each type by the number of newly-developed houses (which is constant across counterfactuals).

3.10.5 Welfare Calculation Details

We compute consumer surplus differences in each counterfactual scenario relative to the unregulated benchmark. Following Small and Rosen (1981), we calculate per-person consumer surplus in each market m as:

$$CS_i = \frac{-1}{\alpha^D} \ln \Sigma_{j \in J_m, z \in \{0,1\}} \exp(\alpha^D p_{jz} + \phi SFHA_{jz} + X_{jz}\beta + \xi_{jz}) \tag{3.28}$$

where j denotes census tract and z indicates flood zone status. For each market, we compute the change in level price required to make per-person consumer surplus in the counterfactual equivalent to that of the same market in the unregulated benchmark. That is, we solve for ΔP_m^{CF} such that

$$\begin{aligned} &\ln\left(\Sigma_{j \in J_m, z \in \{0,1\}} \exp(\alpha^D p_{jz}^{NoSFHA} + X_{jz}\beta + \xi_{jz})\right) = \\ &\ln\left(\Sigma_{j,z} \exp(\alpha^D \ln(P_{jz}^{CF} + \tau_{jz}^{CF} + \Delta P_m^{CF}) + \phi SFHA_{jz}^{CF} + X_{jz}\beta + \xi_{jz})\right) \end{aligned} \tag{3.29}$$

where P_{jz}^{CF} is the house price in levels in the counterfactual of interest, τ_{jz}^{CF} is the tax in levels, $SFHA_{jz}^{CF}$ indicates the SFHA designation in the counterfactual of interest, and p_{jz}^{NoSFHA} is the house price in logs in the unregulated benchmark. The total consumer surplus associated with new development is then $\Delta CS^{CF} = \Sigma_m N_m \Delta P_m^{CF}$, where N_m denotes the number of new houses in each county m.

We compute producer surplus differences as compensating variation for all developers who did not develop in the pre-period. That is, we solve for ΔP_{jz}^{CF} such that

$$\int_{q_0}^{q_1} \left(\exp(p_{jz}^{NoSFHA}) - \exp(\sigma * \Phi^{-1}(\tilde{q}) + \psi E_{jz}^{0} + \eta_{jz}^{1}) \right) d\tilde{q} = \tag{3.30}$$
$$\int_{q_0}^{q_1} \left(\exp(p_{jz}^{CF}) - \Delta P_{jz}^{CF} - \exp(\sigma * \Phi^{-1}(\tilde{q}) + \psi E_{jz}^{CF} + \eta_{jz}^{1}) \right) d\tilde{q} + \int_{q_1}^{1} \left(-\Delta P_{jz}^{CF} \right) d$$

where p_{jz}^{CF} is the log house price in the counterfactual of interest, E_{jz}^{CF} indicates whether the house is adapted under the counterfactual, E_{jz}^{0} indicates whether the house is adapted in the absence of regulation, p_{jz}^{NoSFHA} is the house price in logs in the unregulated benchmark, q_0 is the share of plots developed as of the end of the pre-period, and q_1 is the share of plots developed as of the end of the post-period. We then compute total change in producer surplus by multiplying by the number of gridcells in each location L_{jz} and summing across locations.

We compute government revenue from the tax policy by adding up all taxes levied on newly-developed houses. We compute government revenue under the flood zone policy by estimating the total amount of insurance premiums. Using our flood insurance policy data, we assume that policies inside the flood zone cost $1484 per year and policies outside the flood zone cost $572 per year. Applying back-of-envelope calculations to recent estimates of take-up in Florida (Lingle and Kousky, 2018) and inside and outside the floodplain nationally (Bradt et al., 2021), we assume take-up is 45% inside the flood zone, 6% outside the flood zone in high-risk areas (areas with positive probability of flooding more than 2 feet in the 100-year flood), and 0% outside the flood zone in low-risk areas. As with the tax revenue, we calculate premium revenue only for new development.