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

Anna Russo*

October 15, 2023

PRELIMINARY: CLICK HERE FOR LATEST VERSION

Abstract

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.

*I am grateful to Nikhil Agarwal, Amy Finkelstein, Ben Olken, and Parag Pathak for invaluable advice, guidance, and support. I also thank Samantha Burn, Juan Camilo Castillo, Joseph Doyle, Michael Dinerstein, Josh Gottlieb, Abigail Ostriker, Charlie Rafkin, Maggie Shi, Mike Whinston, Myles Wagner, and Matthew Zahn, for helpful advice and comments, Jonathan Zhang, Kyle Barr, Christine Yee, Steven Pizer, and all members of PEPReC for advice, comments, and conversations about data and institutions, and participants in the MIT PF lunch and NBER Aging and Health research seminar for helpful comments and suggestions. This material is based upon work supported by the National Institute on Aging through grant T32-AG000186 and the National Science Foundation Graduate Fellowship Program under grant number 1745302.

1 Introduction

> "In the end, each nation must decide which style of rationing — by the queue or by price and ability to pay — is most compatible with its culture. Mantras about the virtues of markets are no substitute for serious ethical conviction."
>
> Uwe Reinhardt (2007)

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, 2002).[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, 2002; 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 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.

[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.

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, 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) sub-

[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.

stitution 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, 2002), 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., 2021; 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 effects, 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., 2022), 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 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:

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

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:

$$SW(p,w) = \int g_i \, v_i(x_i(p,w), p, w) f(i)di + g^G p\kappa \tag{2}$$

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

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

follows:

$$SW(0,w^*) - SW(p^*,0) = \int g_i \Bigg(\underbrace{v_i\left(x_i\left(0,w^*\right),0,w^*\right) - v_i\left(x_i\left(p^*,0\right),0,w^*\right)}_{\text{change in allocation}} + \underbrace{v_i\left(x_i\left(p^*,0\right),0,w^*\right) - v_i\left(x_i\left(p^*,0\right),p^*,0\right)}_{\text{change in payoff } \mid \text{ allocation}} \Bigg) f(i)di - \underbrace{g^G p^* \kappa}_{\text{revenue}} \quad (3)$$

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 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 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.

3 Setting and Data

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. Ap-

[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.

proximately 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).

3.1.1 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. 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.

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 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 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

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

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 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 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 4.

Table 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

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

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 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 2a and 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 2c and 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 2c and 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 2c and 2d documents dispersion in exposure, with some clinics substantially "exposed" to the Choice policy.[9]

[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 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 2a and 2b and Figures 2c and 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).

4 Choice Act Policy Analysis: Shifts in p and w

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:

$$y_{it} = \sum_{\tau \neq 2014} \beta_t \cdot \mathbf{1}\{t = \tau\} \cdot 40Miles_i + 40Miles_i + \epsilon_{it} \tag{4}$$

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 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 A.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:

$$y_{it} = \beta \cdot \mathbf{1}\{t > 2014\} \cdot 40Miles_i + 40Miles_i + \epsilon_{it} \tag{5}$$

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 A.2, 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 2c and 2d, in a market- and clinic-level exposure event study design:

$$w_{jt} = \sum_{\tau \neq 2014} \eta_t \cdot \mathbf{1}\{t = \tau\} \cdot ShareEligPre_j + \phi_j + \chi_t + \epsilon_{jt} \tag{6}$$

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 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

$$w_{jt} = \eta \cdot \mathbf{1}\{t > 2014\} \cdot ShareEligPre_j + \phi_j + \chi_t + \epsilon_{jt} \tag{7}$$

pooling together all pre- and post-policy years.

[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).

4.2 Results: direct effects

Effects on choices, utilization, and spending Figure 3 plots event study coefficients from Equation 4 and Table 2 reports coefficient estimates from the pooled difference-in-difference specification in Equation 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 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 3a is that absent data on non-VA financed care, the patterns in Figure 3a could be consistent with either a behavioral response to prices or simply a change in financing for the same choices. Figure 3b rules out a simple change in financing: I plot total community utilization for the TM sample, where I observe the universe of utilization. Figure 3a and Figure 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 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 3d) corresponding to an increase in total VA spending of approximately $28 per eligible enrollee per year (Figure 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 4.3.

Additional results Table 2 presents coefficient estimates of Equation 5, for both the whole sample and the TM sample.[11] Table 2 shows similar patterns regardless of the method of

[11] The full set of event study figures, including Figure 3 replicated for the TM sample, is presented in Appendix A.1.

measuring utilization: visits, relative value units (RVUs), or spending. Table 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 A.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 A.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 A.2 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.

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 A.3 plots the event study figures of Figure 3, focusing on primary care specifically, and documents similar patterns as in Figure 3.[12]

To investigate heterogeneity in responsiveness to p, I estimate Equation 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 4.

Figure 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

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

in healthcare markets, as it highlights that raising prices will screen out lower socio-economic status consumers (Weitzman, 1977).

4.3 Results: equilibrium effects on w

Next, I turn to analyzing the effects of the policy on equilibrium wait times. Figure 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 5 plots coefficient estimates of Equation 6 to test this idea directly. I analyze the effect of Choice on wait times across all specialties and geographies (HRRs) (Figure 5a), and for primary care specifically at the HRR-level (Figure 5b) and the clinic level (Figure 5c). Coefficients in Figure 5 are scaled to represent a move from the 10th to the 90th percentile of the clinic exposure distribution. Figure 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 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 A.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 4, the substitute allocation mechanism is often determined endogenously rather than explicitly designed (Cutler, 2002). 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:

$$w_{jt} = \theta \cdot \mathbf{z_{jt}} + \phi_j^w + \chi_t^w + \epsilon_{jt} \tag{8}$$

$$\ln(s_{jt}^d) = \beta \cdot w_{jt} + \phi_j^s + \chi_t^s + \nu_{jt} \tag{9}$$

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 8, is a more general version of the difference-in-difference estimator discussed in Section 4.1 with a first stage illustrated in Figure 5. However, instead of parameterizing exposure as a continuous interaction, $\mathbf{z_{jt}} = \mathbf{1}\{t > 2014\} \cdot ShareEligPre_j$, which was amenable to event-study plots in Figure 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 5.

Table 3 presents results. Perhaps surprisingly, Table 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 3 documents similar (adverse screening) patterns along measures of health capital, as proxied by lagged utilization.

Appendix Figure A.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.

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 4.2 and Figure 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.

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, is the key structural object for welfare and policy counterfactuals.

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:

$$u_{ijmt} = \underbrace{\beta_{it}'\mathbf{X}_j}_{\text{prefs over clinics}} + \underbrace{\theta_i d_{ij}}_{\text{travel costs}} + \underbrace{\gamma_i w_{jt}}_{\text{waiting costs}} + \underbrace{\alpha_i p_{jmt}}_{\text{price sensitivity}} + \underbrace{\xi_{jmt}}_{\text{demand shock}} + \underbrace{\epsilon_{ijmt}}_{\text{i.i.d. pref shock}} \tag{10}$$

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, 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.

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

[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.

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.

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 A.5 that the covariance between the clinic-level wait times of chosen clinics and patient characteristics 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

[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.

[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.

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 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.

5.3 Equilibrium

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

$$\kappa_{jt} = \sum_i \frac{\exp\left(\beta_{it}'\mathbf{X}_j + \theta_i d_{ij} + \gamma_i w_{jt} + \alpha_i p_{jmt} + \xi_{jmt}\right)}{1 + \Sigma_{j'\in\mathcal{J}_m} \exp\left(\beta_{it}'\mathbf{X}_j + \theta_i d_{ij} + \gamma_i w_{jt} + \alpha_i p_{jmt} + \xi_{jmt}\right)} \tag{11}$$

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

5.4 Estimation

Endogeneity and identification I face two sources of endogeneity when estimating the preferences in Equation 10. First, $\mathbb{E}\left[w_{jt}\xi_{jmt}\right] \neq 0$. This form of endogeneity follows directly from Equation 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.

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]—

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

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 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

$$\beta_{it}'\mathbf{X}_j = \beta_i^{VA} + \beta_i^C + \psi_o^{VA} + \psi_o^C + \chi_t^{VA} + \chi_t^C + \phi_j \tag{12}$$

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 4. Figure 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 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 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 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 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"

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

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 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 B.1 provides more information. In particular, Appendix Figure B.1 recreates Figure 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:

$$\underbrace{\beta_{it}'\mathbf{X}_j + \phi_j + \gamma_0 w_{jt} + \alpha_0 p_{jmt} + \xi_{jmt}}_{\delta_{jmt}\text{: common to all in } jmt} + \sum_b \underbrace{\beta_b^{VA} + \beta_b^C + \gamma_b w_{jt} + \alpha_b p_{jmt} + \left(\theta_0 + \theta_b\right) d_{ij}}_{\lambda_{ijmt}\text{: individual specific parameters}} \quad (13)$$

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 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 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.

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 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. (2010). 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 lead a researcher to (incorrectly) infer that waiting is not costly to consumers and therefore imposes no efficiency costs.

[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.

In column (2), I instrument for w_{jt} using the same $\mathbf{z}_{jt}$ as in the 2SLS exercise in Section 4, using deciles of Choice exposure. When instrumented, γ becomes negative – consistent with Table 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 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, 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 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.1 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 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 7 relate to the raw patterns of screening in Section 4. Section 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

[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.

B.2. However, these descriptive patterns were limited by the fact that I was not comparing willingness-to-wait and 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 7.

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.

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 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 6.2.

My counterfactual analyses only focus on the sub-set of Traditional Medicare dual-eligibles that I use to estimate preferences in Equation 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 1).

6.1.1 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 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 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 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 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.

6.1.2 Quantifying welfare effects

Table 6 quantifies the efficiency and redistributive effects of the two regimes described in Table 5. Column (1) presents the decomposition of the welfare effects of status quo rationing, relative to a price regime from Equation 3 in Section 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 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 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 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 5) — to seemingly high g_i individuals.

6.1.3 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 6) to an individual who prefers the rationed regime (column (2) of Table 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 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 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?

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 3e documented that total spending increased, while Figure 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 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 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 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 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.

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.

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 Philip A. Haile**, "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.

Brot-Goldberg, Zarek C., 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.

Chandra, Amitabh, 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.

Che, Yeon-Koo, 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.

Finkelstein, Amy 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.

Lieber, Ethan M. J. 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.

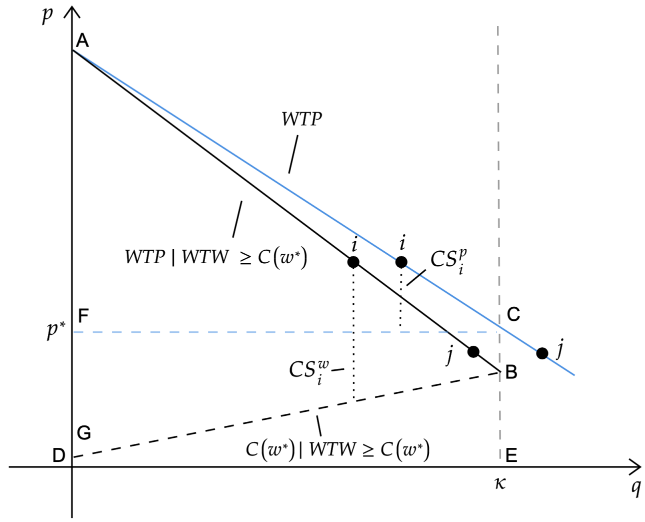
_ , Yevgeniy Feyman, and Steven D. Pizer, "Dually-enrolled patients choose providers with lower wait times: Budgetary implications for the VHA," *Health Services Research*, 2022. Publisher: Wiley Online Library.

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.

Tables and Figures

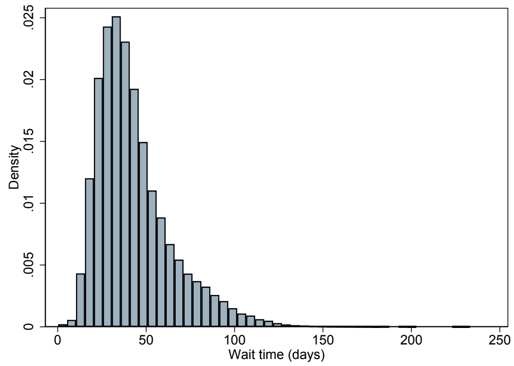
Figure 1: Conceptual framework: graphical illustration



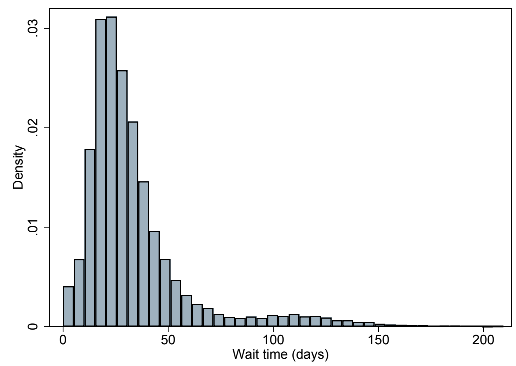
Notes: Figure presents a graphical illustration of the efficiency effects and opportunities for redistribution of surplus under a price-based mechanism and a wait-time-based mechanism. The area between curves AC and AB represents the allocative efficiency loss from rationing by waiting. The curve GB 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: Clinic-level summary statistics

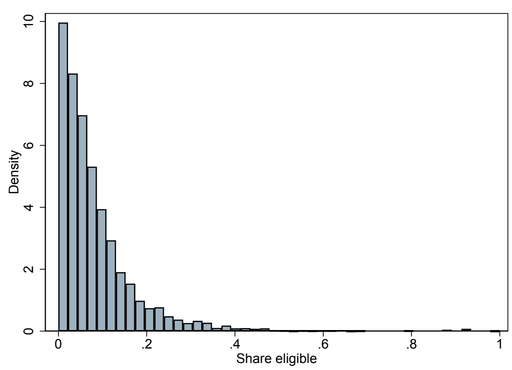
(a) Wait times (days), all specialties



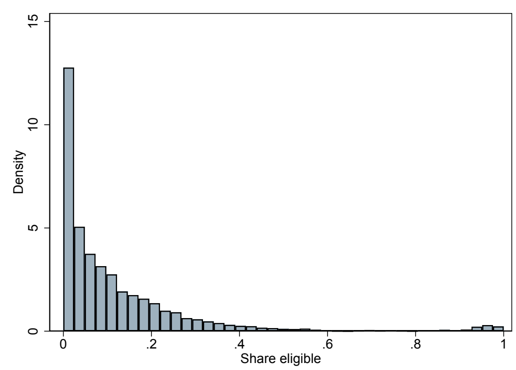
(b) Wait times (days), primary care



(c) Choice exposure



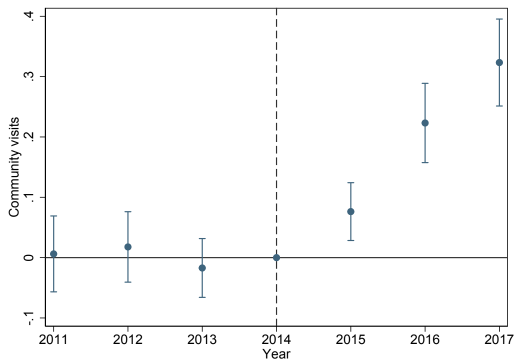
(d) Choice exposure, primary care



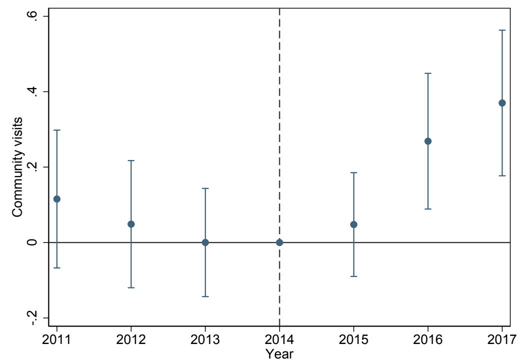
Notes: Figure presents histograms of average wait times and Choice exposure for all outpatient specialties and primary care, both calculated in the pre-Choice (2010-2014) period. Wait times are calculated as the average time between the request date and visit date among all appointments (completed or 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 3: Effect of Choice eligibility on utilization

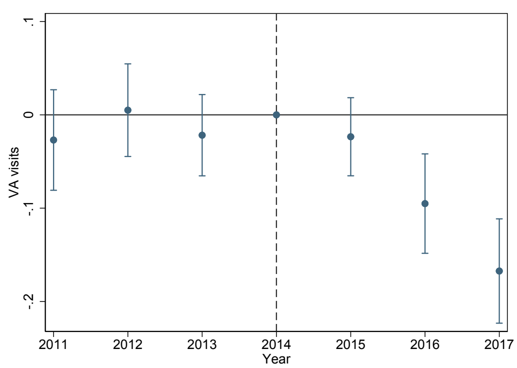
(a) VA-Financed community visits



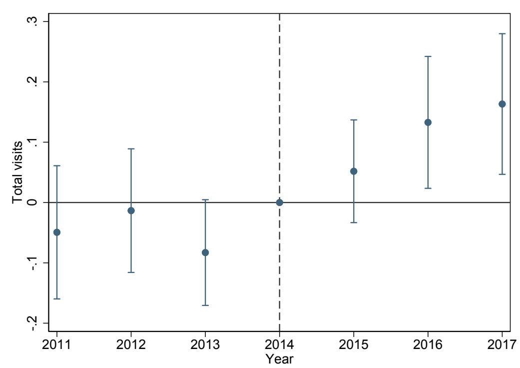
(b) All community visits



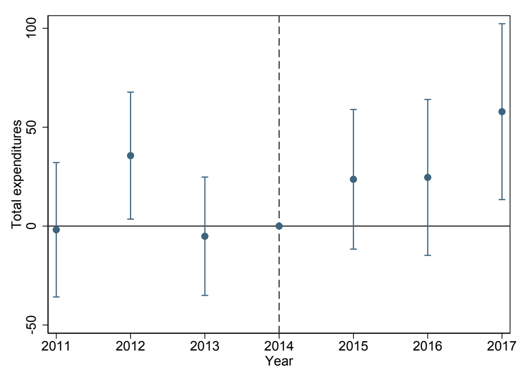
(c) VA visits



(d) Total visits (VA + community)



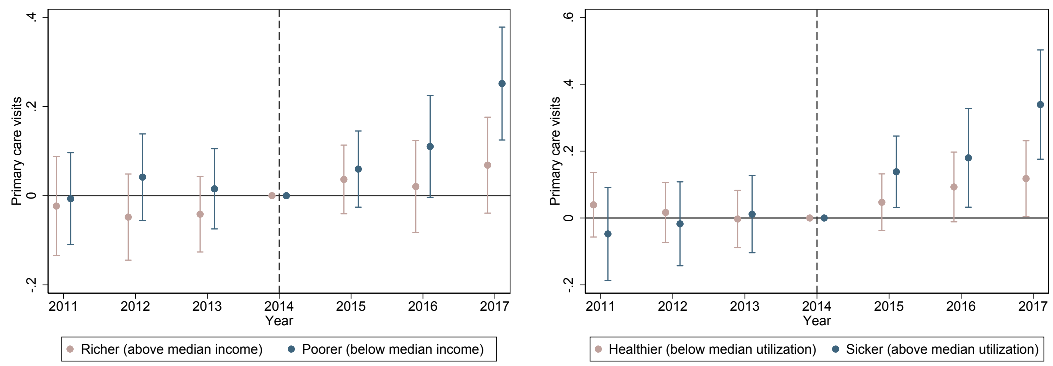
(e) Total VA spending



Notes: Figure presents event study coefficient estimates from Equation 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 4: Heterogeneity on responses to p

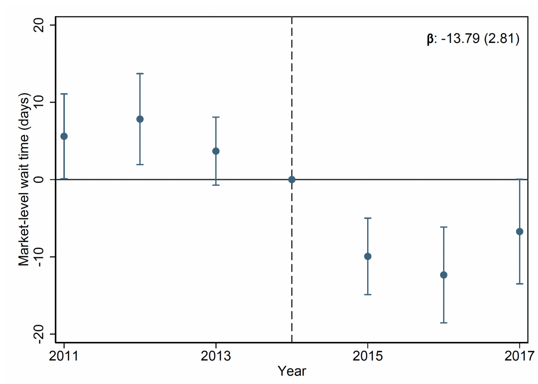
(a) Community PC visits, by income (b) Community PC visits, by health status



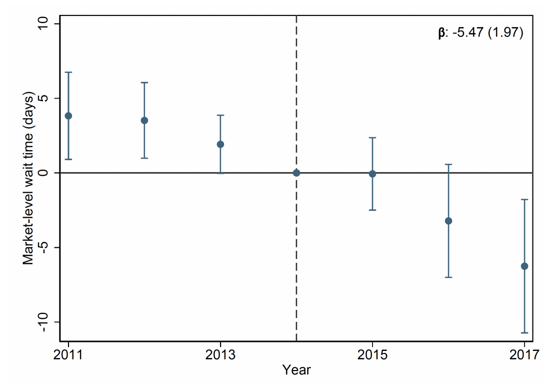
Notes: Figure presents event study coefficient estimates from Equation 4 with the outcome equal to community primary care visits, split by whether a veteran is above versus below median income (a) and health status (b), where health status is measured by the total VA spending in the prior year. The sample 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 5: Effect of choice exposure on wait times

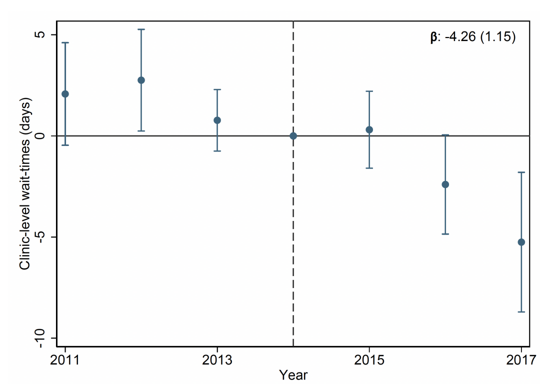
(a) HRR-by-specialty level



(b) HRR-level, primary care

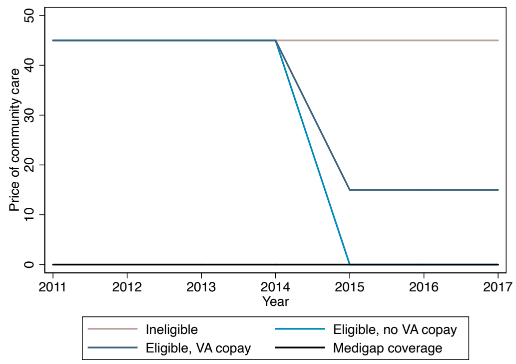


(c) Clinic-level, primary care



Notes: Figure presents event study coefficient estimates from Equation 6 with pooled difference in-difference estimates (from Equation 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.

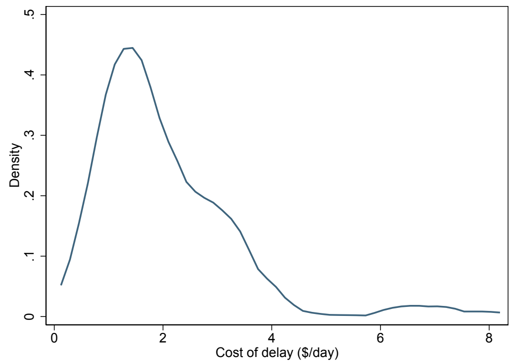
Figure 6: Variation in p



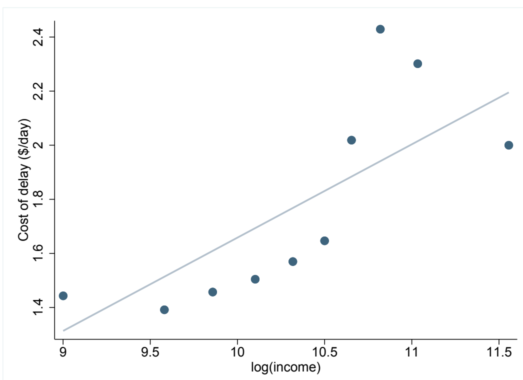
Notes: Figure presents variation in prices for community care among the different out of pocket payment classes for the TM sample. The out of pocket price for TMs without Medigap is calculated as the median out of pocket payment for primary care. Medigap prices are set to zero because the vast majority 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 7: Distribution of the cost of delay, $C_i(w) = \frac{\gamma_i}{\alpha_i}$

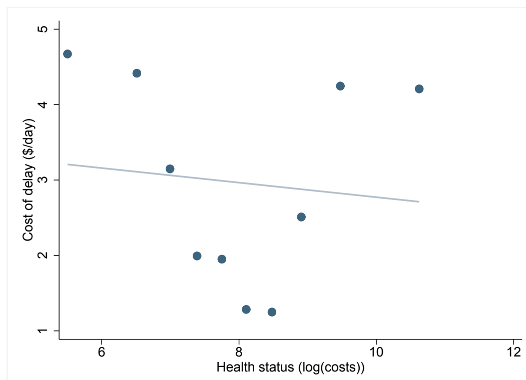
(a) Kernel density plot



(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 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 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: 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 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 3: Heterogeneity in responses to w

	All	By income: Poorer (below medium)	By income: Richer (above medium)	By health status: Sicker (above median utilization)	By health status: 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 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 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 $\left(\chi_t^{VA}, \chi_t^{C}\right)$	✓	✓	✓
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 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 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 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 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 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 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 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 8: The performance of policy instruments

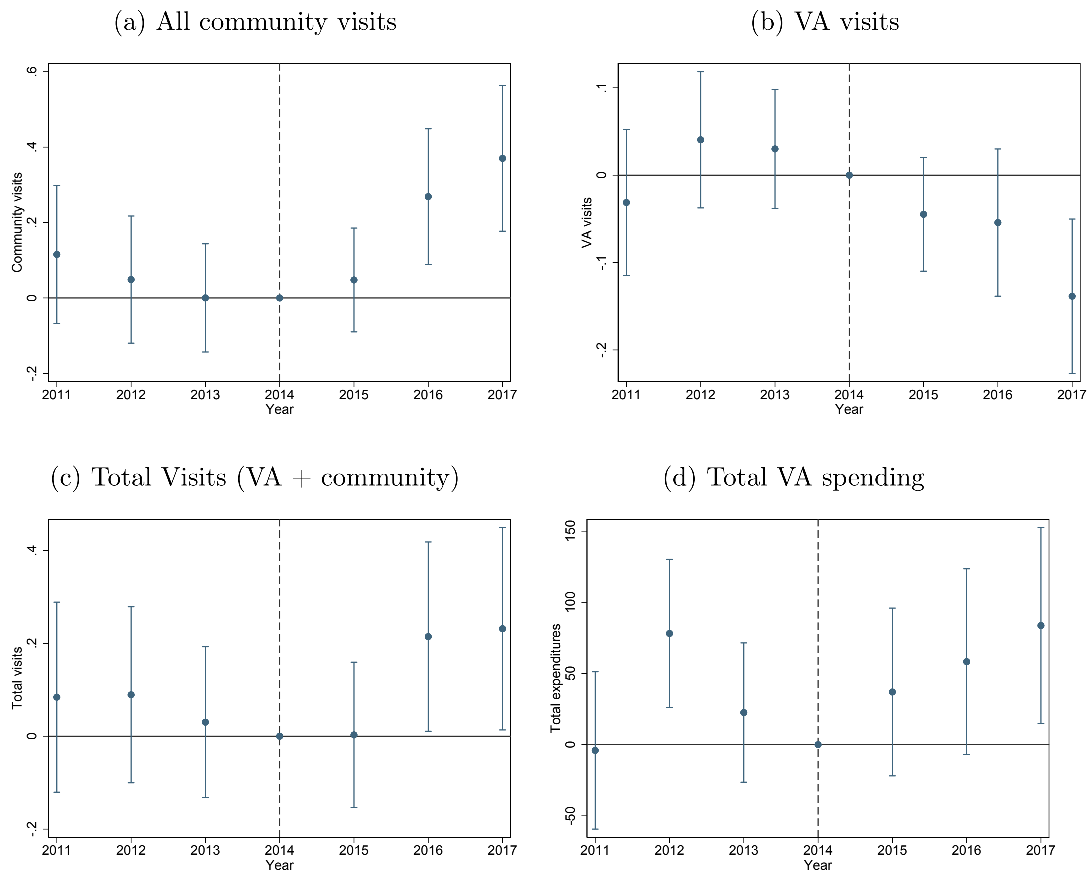
	Price mech.	Choice	Expand to MISSION	Expand to everyone	Copay ↑ from \$15 to \$50
	(1)	(2)	(3)	(4)	(5)
All differences are relative to the status quo					
Δ CS (\$ per veteran, per year)	57.25	1.93	4.28	7.63	11.50
Δ Revenue - Costs	62.90	-3.22	-7.61	-17.59	2.52
Δ Welfare	122.95	-1.84	-4.72	-6.56	14.40

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 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).

A Additional Analyses of the Choice Act

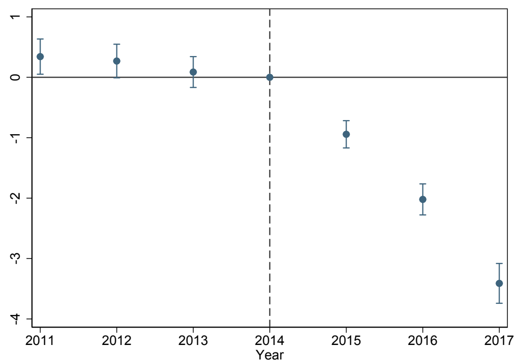
A.1 Robustness and additional results

Figure A.1: Effect of Choice eligibility on utilization, TM sample

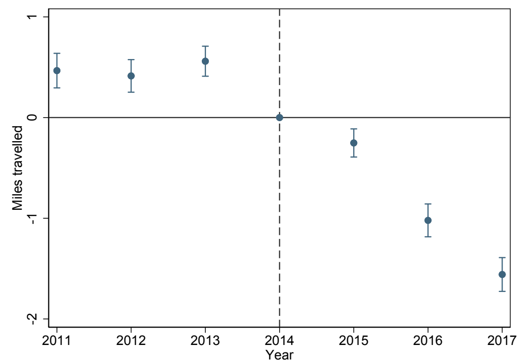


Notes: Notes: Figure presents event study coefficient estimates from Equation 4, estimated on on the TM sample. In sub-figure (a) the outcome is all community outpatient visits per year. Sub-figure (b), (c), and (d) plot VA visits, total visits (VA + community), and total VA spending. Total VA spending 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 A.2: Effects of Choice on characteristics of chosen clinic



(a) Wait times of chosen clinics

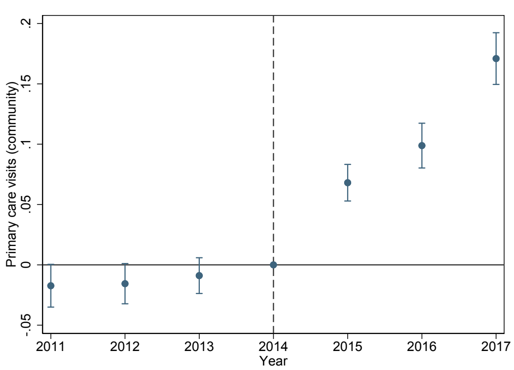


(b) Distance to chosen clinics

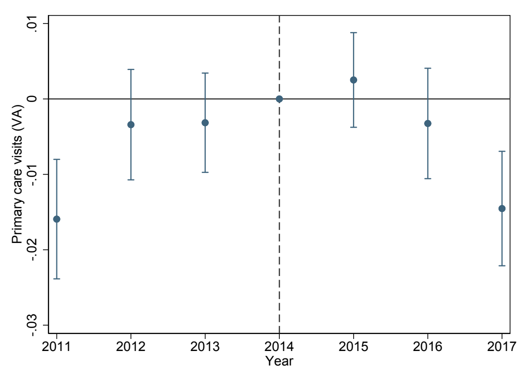
Notes: Figures present event study coefficients from Equation 4 with the outcomes equal to the average wait time (a) and travel distance (b) at the VA and across VA and community care. Wait times at community care are determined by the time between the authorization for community care and the date of 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 A.3: Effect of Choice eligibility on primary care utilization

(a) Community visits: primary care

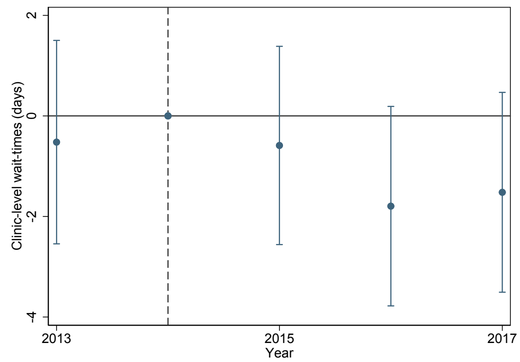


(b) VA visits: primary care



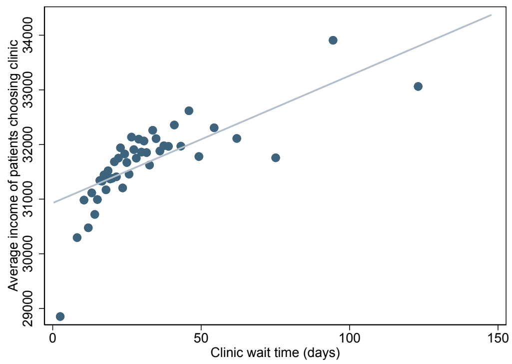
Notes: Figure presents event study coefficient estimates from Equation 4 for primary care specifically. In sub-figure (a) the outcome is all community outpatient visits per year, and in sub-figure (b) I plot VA visits. Effects estimated on the whole sample of enrollees living 10 miles from the 40 mile eligibility threshold.

Figure A.4: Equilibrium effects: new patient wait times

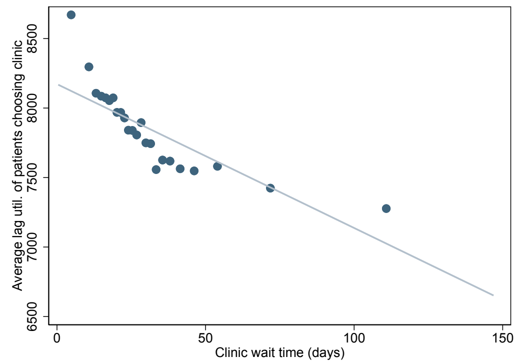


Notes: Figure replicates Figure 5c, calculating wait times only for new patients. The figure excludes 2011 and 2012 because patients are defined as new based on having no visits in two-year look back period.

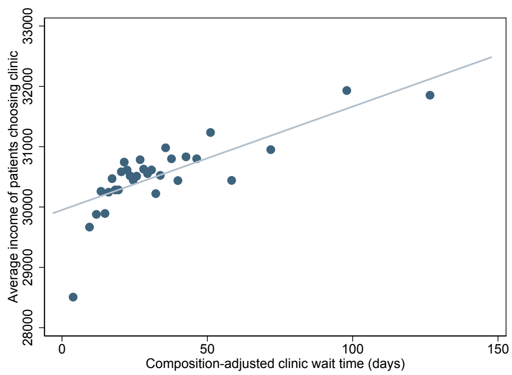
Figure A.5: Screening effects of w: full panel



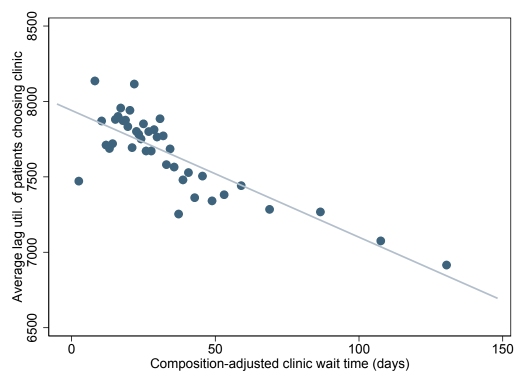
(a) $\mathbb{C}\left(w_{jt}, \overline{y}_{jt}\right)$



(b) $\mathbb{C}\left(w_{jt}, \overline{sick}_{jt}\right) < 0$



(c) $\mathbb{C}\left(\tilde{w}_{jt}, \overline{y}_{jt}\right) > 0$: 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 4.3.

Table A.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 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 7, with the outcome equal to log mortality and log inpatient admissions, among only ineligibles to investigate equilibrium effects on health.

Table A.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 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.

A.2 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:

$$y_{gst} = \beta \mathbf{1}\{WaitElig_{gst}\} + \theta_{gs} + \tau_t + \epsilon_{gst} \tag{14}$$

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 A.3 is therefore not comparable to Table 2 in magnitudes. The primary purpose of this exercise is to examine qualitative patterns.

Table A.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 A.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 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.

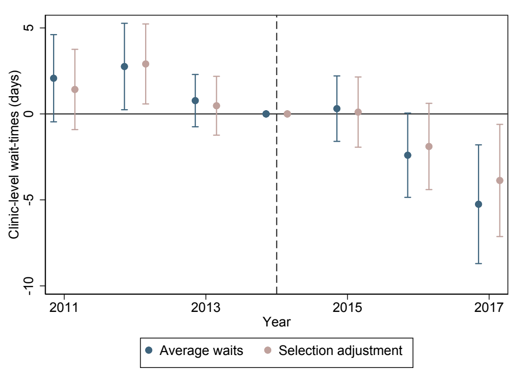
B Model Details

B.1 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.1 re-creates Figure 5c for both the raw mean and the adjusted "unselected" distribution of wait times and shows similar results.

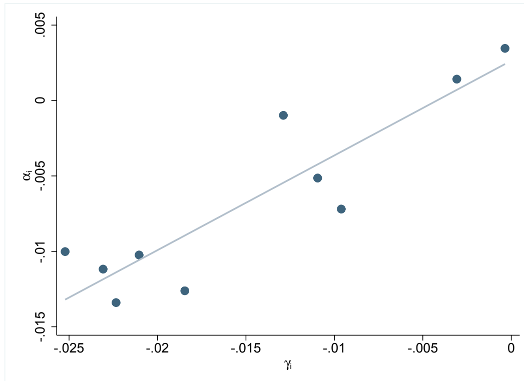
Figure B.1: Effects of Choice exposure on the raw mean and unselected wait times



Notes: Figure plots coefficients from Equation 6 using both raw averages and the constructed unselected distribution of wait times, described above.

B.2 Parameter estimates

Figure B.2: 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.1: 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.