Weighing the Impacts of GLP-1s: Quasi-Experimental Evidence From Provider Adoption

Sam Bock, Jasmin Moshfegh, Jonathan Zhang*

This Version: May 11, 2026

Abstract

The arrival of GLP-1 medications has been described as one of the most important health care innovations in recent years. We provide large-scale quasi-experimental evidence on their real-world impacts by exploiting variation in the eventual prescribing propensities of patients' pre-existing primary care providers. Using a panel intent-to-treat design, we compare outcomes for 1.4 million diabetic or obese veterans based on their 2018 provider's eventual propensity to adopt GLP-1s, leveraging comprehensive electronic health records and biomarker data from the Veterans Health Administration, a setting with minimal insurance attrition and low-cost access to these drugs. Patients whose providers become higher propensity adopters experience substantial improvements in glycemic control and clinically meaningful weight loss; our treatment-on-the-treated estimates closely match estimates from clinical trials. Despite these metabolic benefits, we find no statistically significant average effects on emergency department utilization, or non–GLP-1 medical spending through 2024. However, quantile regression results show heterogeneous effects, with spending declining at lower quantiles and rising at higher quantiles for patients of higher-propensity adopters.

Keywords: GLP-1, diabetes, obesity, cost

*Bock: Duke University, VA Office of Mental Health, sam.bock@duke.edu, Moshfegh: Imperial College London, Veterans Health Administration Palo Alto, j.moshfegh@imperial.ac.uk, Zhang: Duke University, VA Office of Mental Health, and NBER. jonathan.zhang@duke.edu. We thank Jason Abaluck, Pat Bayer, Emily Cuddy, Camille Fredrickson, Jennifer Lee, Matt Maciejewski, Edward Norton, Dan Sacks, Molly Schnell, Duncan Thomas, Jodie Trafton, Winnie van Dijk, Todd Wagner, Coady Wing, and seminar participants at Duke, NBER Economics of Chronic Diseases, Rockwool Foundation, Triangle Applied Microeconomics Workshop, University of Melbourne, and University of Sydney for valuable feedback. We also gratefully acknowledge support from the Department of Veterans Affairs Office of Academic Affiliations Advanced Fellowship, The Big Data Scientist Training Enhancement Program (Moshfegh). The contents of this paper do not represent the views of the U.S. Department of Veterans Affairs or the United States Government.

1 Introduction

GLP-1 receptor agonists are potentially one of the most important pharmaceutical innovations of the past decade. Originally developed as diabetes therapies, these drugs produce large and rapid improvements in glycemic control and body weight, and have generated substantial enthusiasm among clinicians, policymakers, and the public (The Economist, 2024; Moiz et al., 2025). Their diffusion has sparked speculation and debate about the potential downstream effects on health, healthcare costs, and even socioeconomic outcomes (American Enterprise Institute, 2024). As payers evaluate whether and how to reimburse GLP-1 drugs, credible evidence on their real-world impacts has become increasingly needed (Congressional Budget Office, 2024; Hwang et al., 2025).

Despite intense interest, existing clinical evidence is limited in two fundamental ways. First, randomized controlled trials are designed to establish biological efficacy under tightly controlled conditions and selected patient populations (e.g., SUSTAIN trials; Marso et al., 2016); they do not speak to the broader consequences of large-scale adoption among larger and more heterogeneous populations. Second, the trials evaluate short-term clinical endpoints, leaving open the question of whether the costs of GLP-1 therapies are offset by downstream reductions in health care utilization or spending (World Health Organization, 2025). As a result, some of the most policy-relevant questions surrounding GLP-1s remain unanswered: How large are their real-world metabolic benefits? Do they have spillover effects on physical and mental health? And do they translate into meaningful changes in health utilization and medical spending?

We provide one of the first large-scale quasi-experimental estimates of the real-world impacts of GLP-1 therapy on health and health care utilization. Our empirical strategy closely mirrors the way many patients initiate GLP-1s: through their primary care provider's (PCP) adoption decisions (Dixon et al., 2025). We exploit the fact that, when semaglutide (Ozempic) was included in the Veterans Health Administration (VA) formulary in late 2020, veterans were already in long-standing relationships with PCPs. Some PCPs rapidly adopted semaglutide and other GLP-1 therapies, while others did not. Pre-existing patient-PCP relationships generate plausibly exogenous variation in exposure to GLP-1 therapy when semaglutide became widely available.

Our data leverages rich administrative electronic health records from the Veterans Health Administration, an ideal setting for studying GLP-1 adoption at scale. Unlike commercially insured populations, VA patients do not experience insurance-related attrition, allowing us to follow a stable

cohort over time. The VA also contains one of the largest diabetic and obese patient populations in the country (Liu et al., 2017), with over 250,000 veterans receiving a GLP-1 by the end of 2024. Critically, electronic health records provide regular biomarker monitoring—including A1C and weight—offering unusually rich longitudinal measures of health.

We follow a baseline cohort of almost 1.4 million type 2 diabetic or obese veterans as of 2018, actively engaged in primary care, at the time of Ozempic's FDA approval. For each patient, we construct a (leave-out) time-invariant measure of their 2018 PCP's eventual propensity to adopt GLP-1s in 2021–2022 (the period after Ozempic was added to the VA formulary in November 2020). Using an intent-to-treat framework where patients are locked-in to their 2018 PCP, we then compare patients who had a PCP who becomes a higher vs lower GLP-1 adopter, before and after 2021, in an event study framework. This panel framework allows us to control for temporal trends, regional-by-time differences, and importantly, baseline patient and provider characteristics (e.g., differences in patient risk and provider practice behavior).[1]

We document three main findings. First, exposure in 2018 (around FDA's approval of Ozempic) to PCPs who become high GLP-1 adopters leads to large and persistent increases in GLP-1 use, confirming a strong and persistent first stage. A patient whose PCP in 2018 becomes a high-prescriber—for diabetic or obese patients—relative to a low-prescriber, has an approximately 20 percentage point higher likelihood of filling a GLP-1 in 2023 and 2024, and accumulates 400 additional days supply of GLP-1s by the end of 2024.

Next, we show that high GLP-1 adopting PCPs meaningfully improve glycemic control (Hb A1C) and body weight, especially after 2021, closely replicating patterns observed in clinical trials. To translate these reduced form effects (of higher GLP-1 prescribing PCPs) to the causal impact of *GLP-1 therapy*, we require that providers' GLP-1 adoption is not accompanied by concurrent changes in other aspects of diabetes, cardiovascular or preventive care management (exclusion restriction), and that measurement of biomarkers is not selectively observed over time. We present evidence consistent with both assumptions, and that the results are robust to controls and specifications that account for potential violations. With these conditions in place, we estimate average treatment-on-the-treated (ToT) effects per 30-week dose GLP-1 regimen—a common dose duration in clinical trials.

[1] We also implement a stacked difference-in-differences design as a robustness check (Appendix F), though our primary strategy leverages provider-level variation to address endogenous treatment timing.

ToT estimates on blood glucose and body weight are largest in 2021, immediately after semaglutide's inclusion in the VA formulary. A 30-week dose of GLP-1 reduces weight by 3.28% and lowers A1C by 0.32 percentage points. Our ToT estimates for 2021 closely match those from the SUSTAIN clinical trials (a landmark series of semaglutide trial that informed FDA approval): randomized trial estimates fall inside our confidence intervals. Over time, however, the treatment effect of a marginal 30-week regimen dissipates; by 2024, the treatment effects on weight are almost halved and A1C improvements disappear. This attenuation is explained in part by shifting complier composition between 2021 to 2024: as GLP-1 diffusion expands, the marginal patients induced into treatment become progressively healthier. This finding has important policy implications for expanding GLP-1 access to broader, lower risk populations who may have lower marginal benefits relative to higher risk patients.

Finally, despite these metabolic improvements, we find no evidence of broader health improvements or utilization changes, on average. Although we replicate a reduction in heart attacks and strokes among patients with existing cardiovascular disease and chronic kidney failure (SUSTAIN 6 trials; Marso et al., 2016), we find no statistically significant decline in broader ED visits, or in any major diagnostic category. ED visits for alcohol or substance abuse are also unchanged. Moreover, mental health measured via clinical questionnaires (PHQ-9 and AUDIT-C) are unchanged.[2] Notably, we do not see changes in total (non-medication, non-GLP-1) medical spending. ToT estimates allow us to rule out—with 95% confidence—reductions in total non-GLP-1 medical spending of more than 7.3% between 2021 and 2024. While these estimates suggest a null *average* effect of GLP-1s on spending, we uncover substantial heterogeneity across the spending distribution. Quantile regression estimates show that patients of high GLP-1 adopters exhibit spending increases at the upper end of the spending distribution, while spending declines at the lower end of the distribution.

These results have immediate implications for coverage policy. They imply that short-run reimbursement decisions should not rely on the expectation that GLP-1s will quickly pay for themselves through reduced medical utilization. Instead, the case for coverage rests on valuing metabolic and longer-run health benefits, with a potentially important role for targeting treatment toward patients with the highest expected gains.

Recent studies have begun to examine the real-world impacts of GLP-1 therapies using vari-

[2]We estimate a reduction in moderate drinking which supports early promise from phase 2 GLP-1 and alcohol use disorder trials (Hendershot et al., 2025).

ation in treatment timing. Lu et al. (2025) compare early versus later semaglutide adopters across two health care systems and document a statistically significant increase in medical expenditures over the first two years following initiation. Concurrent work by Wing et al. (2025) analyzes health insurance claims from an eight-year continuously enrolled commercial population of adults with type 2 diabetes, implementing a stacked difference-in-differences design that contrasts individuals initiating GLP-1 therapy in 2018 versus 2023. In these settings, early versus late initiation may implicitly reflect underlying heterogeneity in provider adoption of GLP-1 therapies. We complement this literature by directly leveraging provider-level adoption variation in the Veterans Health Administration, allowing us to study GLP-1 diffusion in a large, stable panel without insurance attrition and with unusually rich biomarker data, including A1C and weight. A larger medical literature studies GLP-1 initiation vs non-initiation (e.g., Michalak et al., 2025; Wennberg et al., 2025) and vs initiation of other diabetes medications (e.g., Xie et al., 2025; Cai et al., 2026), relying primarily on matching methods. These approaches match on baseline characteristics rather than pre-treatment trends, which may differ across treatment and comparison groups.[3]

We also contribute to an important and growing literature on the real-world health and economic benefits of medical innovations, outside of clinical trials. A number of empirical studies have estimated the causal impacts of medical innovations on economic benefits (Garthwaite, 2012; Shapiro, 2022; Daysal et al., 2024; Moshfegh, 2025). Recent studies have demonstrated the nuances of extrapolating and interpreting results from clinical trials.[4] Our findings highlight how treatment effects evolve as technologies diffuse: changing complier composition implies that marginal patients increasingly differ from trial populations, with implications for external validity and for policies aimed at more effective treatment targeting.

Finally, our study contributes to the role of technological change in healthcare spending (see

[3]Our study also relates to a large literature on the economics of obesity. Most closely related is work examining the causal effects of interventions on health and behavior among individuals with obesity or diabetes, including health insurance coverage (Bhattacharya et al., 2009), workplace wellness programs, food and beverage taxes or subsidies (see Cawley, 2015, for a review), and information (see e.g., Oster, 2018; Alalouf et al., 2024). Across these interventions, effects on weight are generally modest, with bariatric surgery as a notable exception for a subset of patients (Cawley, 2016). While weight loss medications have been shown to be effective, they often carry significant side effects (Yanovski and Yanovski, 2014). We contribute by providing new evidence on a newer class of drugs—GLP-1 therapies—originally developed for diabetes but increasingly used to treat obesity.

[4]For instance, Alsan et al. (2025) show that a food-as-medicine program to treat diabetes was effective for some individuals despite a null result in the original clinical trial, due to mean reversion across both control and treatment arms. Angrist et al. (2025) find larger risks and benefits of hormone therapy than the original trial when accounting for imperfect compliance. Abaluck et al. (2025) show that clinical trials underestimate the risk of serious adverse events, as a result of under-enrolling patients with these risks.

e.g., Cutler and Huckman, 2003; Cutler, 2004; Chandra and Skinner, 2012). A large body of literature argues that technological change is an important factor in the rise of healthcare expenditures in the U.S. We complement this literature by studying the impacts of a rapidly diffusing innovation that has the potential to affect aggregate healthcare spending, with over 1 in 10 adults worldwide affected by obesity or diabetes. Congressional Budget Office (2024) estimates that coverage of anti-obesity medications by Medicare would increase federal spending by approximately $35 billion from 2026 to 2034, despite a non-negligible cost reduction stemming from health improvements.[5] These estimates are based on projections of uptake, adherence, and savings from improved health, whereas we provide quasi-experimental evidence of real-world spending effects.

The remainder of the paper is organized as follows. We begin by providing background on diabetes and its treatment, the clinical evidence on GLP-1 therapies, and how GLP-1s are prescribed within the VA. Section 3 describes the data, variable construction, and baseline sample. Section 4 outlines our empirical strategy and identifying assumptions. Section 5 presents our main results, beginning with the reduced-form estimates, followed by robustness and sensitivity analyses—including placebo tests among non-diabetic, non-obese, younger patients—that support our identification. We then use the quasi-experimental variation in PCP GLP-1 adoption to estimate treatment-on-the-treated effects of GLP-1 therapy. The final section discusses directions for future work and concludes.

2 Background

In this section we briefly detail diabetes, its treatment, the efficacy of semaglutide and GLP-1s, and how it is prescribed in the VA.

2.1 Diabetes and its Treatment

Diabetes is a chronic metabolic disease characterized by elevated blood glucose resulting from insufficient insulin production (Type 1) or insulin resistance (Type 2; T2D) (WHO, 2024). In 2021, more than 1 in 10 Americans had diabetes—90–95% of whom had T2D—making it the eighth leading

[5]To the best of our knowledge, all cost projection studies arrive at cost savings associated with GLP-1 therapy, typically by relying on strong assumptions about long-run disease progression and causal relationships between weight loss and medical expenditures (e.g., Hong et al., 2019; Gómez Lumbreras et al., 2023; Sexton Ward et al., 2023; Ippolito and Levy, 2024; Hwang et al., 2025).

cause of death in the United States and contributing an estimated $412.9 billion in total costs in 2022 (Parker et al., 2023). Obesity, defined as body mass index (BMI) ≥ 30, is a major modifiable risk factor for T2D; among U.S. adults with diabetes, approximately 90% were overweight or obese (CDC, 2024).

Management of T2D focuses on glycemic control, typically targeting hemoglobin A1C blood glucose level below 7%. Treatment ranges from lifestyle modification to pharmacotherapy. Common drug classes include biguanides (metformin), DPP-4 inhibitors, SGLT2 inhibitors, sulfonylureas, insulin, and most recently, GLP-1 receptor agonists. Semaglutides belong to the GLP-1 class, which mimics endogenous GLP-1 to stimulate insulin secretion and support appetite regulation. Seven GLP-1 agonists are currently FDA-approved for T2D treatment, including semaglutide (Ozempic and Rybelsus; Wegovy is approved for weight loss).[6] The first older generation GLP-1, Exenatide, was approved in the U.S. in 2005.

2.2 What We Know About Semaglutide and GLP-1

The arrival of semaglutides represented a major shift in the perceived efficacy of GLP-1s. The December 2017 FDA approval was based on results from the Semaglutide Unabated Sustainability in Treatment of Type 2 Diabetes (SUSTAIN) clinical research program—a series of clinical trials evaluating the effects of semaglutide in T2D patients. Spanning over 10 clinical trials, across multiple countries, between 2014-2019, the trials established the safety and efficacy of semaglutide in terms of reducing A1C levels and body weight (Sorli et al., 2017; Ahrén et al., 2017; Ahmann et al., 2018; Aroda et al., 2017; Rodbard et al., 2018; Marso et al., 2016; Pratley et al., 2018; Lingvay et al., 2019; Zinman et al., 2019; Capehorn et al., 2020; Frías et al., 2021). Results from the trials demonstrated reductions in A1C levels of ca 0.2-1.5 points and weight losses of about 1-6 kg, depending on the semaglutide dose (e.g., 0.5mg or 1.0mg) and the nature of treatment in the control arm (e.g. metformin or SGLT2s, etc.). In addition, SUSTAIN 6 established cardiovascular benefits of semaglutide treatment for patients with established cardiovascular disease (see Table A.2 for a summary of the findings of the SUSTAIN program and Figure A.12 for treatment effects for each SUSTAIN trial).

[6]These include: Dulaglutide (Trulicity), Exenatide (Byetta) and Exenatide extended-release (Bydureon), Liraglutide (Victoza), Lixisenatide (Adlyxin), Semaglutide injection (Ozempic), Semaglutide tablets (Rybelsus), and Tirzepatide (Mounjaro).

With mounting evidence on the weight loss properties of semaglutides, the FDA broadened its indication for semaglutides to obesity in June 2021, under the brand name Wegovy. This decision was based on evidence from the STEP (Semaglutide Treatment Effect in People with Obesity) clinical trial program. Across four studies enrolling around 4,500 participants, the trials demonstrated that individuals who received Wegovy alongside a reduced-calorie diet and increased physical activity, experienced greater weight loss.[7]

In recent years, many observational studies have used matching methods to compare patients initiating GLP-1s with those starting other diabetes medications (e.g., metformin, insulin). These studies have examined a wide range of outcomes—such as overdose risk (Wang et al., 2024b), cirrhosis (Kanwal et al., 2024), colorectal cancer (Wang et al., 2024a), and 165 health outcomes (Xie et al., 2025). However, these studies have important limitations. They do not account for the endogenous clinical decision to initiate semaglutide rather than another agent, and their matching procedures are based on static, cross-sectional covariates rather than the dynamic trajectories (e.g., trends in health status or treatment response) that may systematically differ before initiation.

2.3 GLP-1s in the VA

Prior to 2020, GLP-1s were not in the VA national formulary and could only be prescribed under special prior authorization. In Figure A.1, the number of patients on older generation GLP-1s (e.g., dulaglutide, liarglutide) was low prior to 2020. In November 2020, Ozempic (semaglutide) was awarded a national contract in the VA as the workhorse GLP-1 for new initiation subject to criteria for use. These criteria for use details can be found in Appendix B. This led to a rapid rise in semaglutide use since late 2020.

Veterans enrolled in VA health benefits generally have access to affordable medications through the VA. When medications, including GLP-1s, are prescribed for a service-connected condition (such as diabetes, among others) it is provided free of charge. Veterans with a service-connected disability rating of 50% or higher (for any disability, not necessarily diabetes) also pay no copays for any

[7]STEP 1, which included individuals that were overweight or obese with at least one comorbidity (excluding T2D), showed that 83.5% of patients who were treated with Wegovy achieved at least a 5% weight reduction compared to 31.1% among placebo recipients (Wilding et al., 2021). The corresponding numbers for STEP 2, which included individuals who were overweight or obese *with* a T2D diagnosis, were 67.4% compared to 30.2%, respectively (Davies et al., 2021).

approved medications.[8] In all other cases, formulary medication (e.g., Ozempic) and approved non-formulary medication (e.g., Wegovy), copayments in 2025 are capped at $11 for a 30-day supply, with an annual out-of-pocket maximum of $700.[9]

3 Data

We utilize data spanning 2000 to 2024 from the Veterans Health Administration (VHA), which delivers healthcare services to over nine million qualified veterans (U.S. Department of Veterans Affairs, 2025). As the nation's largest integrated healthcare system, the VHA operates through a network that includes 170 medical centers and more than 1,300 community-based outpatient facilities across the United States.

The dataset encompasses all healthcare encounters occurring within the VA—ranging from outpatient visits and hospital admissions to prescription records and other services. Because the VA was an early adopter in implementing electronic health records, it provides highly detailed, structured information. This includes diagnosis and procedure codes, identifiers for providers and facilities, patient demographic details, and time-stamped clinical events. Importantly, the data also capture clinical measures that are seldom present in claims datasets, such as laboratory values (e.g., Hb A1C for glucose control), biomarkers and vital signs (weight, blood pressure, etc.), and standardized clinical assessments including PHQ-9 depression and alcohol use disorder (AUD) screens.

For medical care outside the VA, we observe and incorporate VA-reimbursed community care claims. VA reimburses community care delivered in instances of emergency or when VA access is limited by wait times, travel distance, or service availability (U.S. Department of Veterans Affairs, 2023). Unlike VA electronic health records, these data are reimbursement claims with diagnoses, place of service codes, and charges, etc., but do not contain biomarker data.

3.1 Variable Definitions

We briefly describe our outcome variables below and present more details in Appendix C.

[8] Veterans can be rated as having a service-connected disability for a variety of reasons. In 2018, there were 445,566 veterans receiving disability compensation for diabetes, and over 2.6 million veterans with a rating of over 50% (VBA, 2019). These patients all pay no copayments for approved medications.

[9] Ozempic (Semaglutide) is a tier 2 medication subject to a copay of $8 in 2025, Wegovy is a tier 3 medication and subject to a copay of $11 in 2025.

GLP-1 Prescriptions and Other Medications We construct patient-year-quarter indicators for any GLP-1 use from outpatient prescriptions, clinically administered initiation (e.g., hospital setting or supervised dose for educational purposes), and special authorization order requests. In addition to any use, we also measure intensive margin days supply, as well as total cumulative measures of ever use and acumulated days supply. We also construct a measure of any diabetes medication use and categorize individual diabetes drug classes (e.g., metformin, insulin). To characterize the prescribing patterns of high GLP-1 adopting PCPs, we also construct indicators for statins and antihypertensives (e.g., ACE inhibitors, beta blockers, etc.), which are commonly prescribed cardiovascular drugs.

We do not observe GLP-1 fills obtained outside the VA and not reimbursed by the VA; to the extent that patients of low-prescribing PCPs differentially obtain GLP-1s elsewhere, this would attenuate our results. However, given substantially lower out-of-pocket costs within the VA, we expect such behavior to be limited.

Hemoglobin A1C Blood Glucose Levels We construct median A1C for each patient-year-quarter observation, as well as an indicator for median A1C below 7% (the typical target goal for T2D patients; American Diabetes Association, 2025), conditional on measurement. We investigate and address selection into measurement when analyzing A1C outcomes.

Weight We take the logarithm of median weight for each patient-year-quarter, measured in pounds, conditional on measurement. We also report body mass index (BMI) below 30—the clinical cutoff for obesity—as a secondary outcome.

Healthcare Utilization and Spending We construct an outcome for any emergency department (ED) visit (VA and non-VA) in a given year-quarter and categorize ED visits into major diagnostic categories based on the primary diagnosis code. We also analyze healthcare spending. For VA spending, an average cost measure based on Medicare relative value units (RVUs) and VA utilization data (Wagner et al., 2003; VA Health Economics Resource Center, 2025) is available until 2024 Quarter 3, and approximates VA operating costs for providing healthcare. For non-VA spending, we use total charges from non-VA reimbursed claims, which represents the amount charged to and paid by the VA. We sum the VA and non-VA costs to arrive at a total spending which can be

disaggregated into outpatient and inpatient spending.

Mental Health and Substance Use Outcomes We construct year-quarter indicators for drug overdose poisoning and alcohol intoxication occurring in VA or non-VA EDs and hospitalizations. These are generally rare events; we complement these adverse events with measures of acute mental health via Patient Health Questionnaire–9 depression screenings (PHQ-9; Kroenke et al., 2001) which assesses depression and distress over the past two weeks, and alcohol consumption from the Alcohol Use Disorders Identification Test (AUDIT-C; Bush et al., 1998) which assesses frequency and intensity of alcohol consumption over the past year. For clinical screening outcomes, we construct average raw scores and indicators for screening below moderate depression and moderate AUD risk, conditional on observation.

3.2 Sample

We study patients whose pre-period health status placed them at risk of receiving GLP-1 therapy once the drug became available. Our baseline analytic sample includes veterans diagnosed with T2D by 2018 or with obesity between 2015 and 2018 who were actively receiving primary care at the VA. We begin with 2,550,380 patients who either had a type 2 diabetes diagnosis recorded at any time between 2000 and 2018 or an obesity diagnosis between 2015 and 2018 and who were alive at the end of 2018.[10]

Next, we restrict the sample to 1,893,241 patients that were receiving VA primary care and had an assigned primary care provider (PCP) as of February 2018.[11] Finally, we restrict the sample to patients whose 2018 PCP remained employed in the VA through at least early 2021, enabling us to construct each provider's GLP-1 prescribing propensity. The resulting analytic sample contains 1,374,666 veterans with type 2 diabetes or obesity—patients who are plausibly eligible for semaglutide initiation once it becomes available. In Table 3, we report separate findings for diabetic and obese subpopulations. We construct a panel-year-quarter dataset for each patient from 2018 to end of 2024, resulting in 34,293,608, patient-year-quarter observations where the patient is alive.

[10] We apply different look-back windows for diabetes and obesity because diabetes is a chronic, typically lifelong condition, whereas obesity status can change over time. Restricting obesity diagnoses to 2015–2018 helps ensure that baseline classifications reflect patients' contemporary health status.

[11] A key advantage of our data is that we directly observe PCP–patient assignments from the VA's primary care management system rather than inferring them from visit patterns (Currie and Zhang, 2025), which can be endogenous to health shocks or treatment intensity.

Table 1 presents summary statistics for our sample. The average patient is 64.8 years old in 2018, 69% have T2D, while 55% are obese, and 24% are both diabetic and obese. In 2018, average A1C is 7.1%, BMI is 33.1, and 35% were on some form of diabetes medication in the prior year, with metformin being the most common. Average quarterly medical spending was $3,181 with a median of $760.

4 Empirical Strategy

Our empirical strategy exploits variation in the rapid adoption of GLP-1 across PCPs after the VA added semaglutide to its national formulary, leveraging patient–provider relationships established prior to semaglutide's market entry. Approximately 78% of VA-dispensed semaglutide is prescribed in primary care, with the remainder in specialty care and programs (e.g., endocrinology and weight management clinics), making primary care the central setting for semaglutide diffusion within the VA. Specifically, we fix patient–PCP assignments at the start of 2018, at the time of semaglutide approval (December 2017) and nearly three years before its inclusion in the VA formulary in November 2020. We then compare outcomes of patients whose 2018 PCPs ultimately became high vs low GLP-1 prescribers in 2021–2022 (i.e., early adopters). Crucially, we do not require patients to stay with their 2018 PCP. This design yields an intent-to-treat framework that avoids patient self-selection into GLP-1s——an especially important feature given the rapid growth of patient demand and active "shopping" for GLP-1s in recent years.[12]

Our approach differs from the standard provider propensity design (e.g., Eichmeyer and Zhang, 2022, 2023; Bhalotra et al., 2025) in two key respects. First, we adopt a within-patient event study, including patient fixed effects—which encompasses provider fixed effects—that absorb average differences across physicians in practice style, patient panel composition, and baseline clinical outcomes. This ensures that identification comes from differential changes in patient outcomes after semaglutides become available (and hence GLP-1s become more common), rather than from pre-existing cross-sectional differences. Second, we fix patient–provider relationships in 2018, which eliminates concerns that patients might select into high-prescribing PCPs precisely to gain access to semaglu-

[12]It is reported that 1 in 8 Americans have tried GLP-1s by 2024 and 20% of patients receive their GLP-1s from a non-physician/clinic source (Montero et al., 2024). Online pharmacies are becoming increasingly popular—often backed by celebrity endorsements—selling compounded semaglutide that require less FDA oversight (Ashraf et al., 2024; Lupkin, 2025).

tide.[13]

Propensity to Adopt GLP-1 We construct a time-invariant measure of each PCP j's underlying propensity to prescribe GLP-1 receptor agonists during the period of rapid GLP-1 expansion. Operationally, we measure this propensity using prescribing of *semaglutide* in 2021–2022, which provides a high-signal indicator of newer generation GLP-1 adoption because semaglutide's formulary entry generated substantial variation in uptake relative to older generation GLP-1s. Throughout the paper, we refer to this as "GLP-1 adoption"; we show that semaglutide prescribing strongly tracks broader GLP-1 utilization when we present the first stage in Figure 2.

Specifically, we estimate a patient-year-level regression of whether (diabetic or obese prior to 2018) patient i received a semaglutide prescription from their (2018) PCP in 2021 and in 2022, on prior (2016–2017) A1C decile bins, an obesity indicator, station-by-year fixed effects,[14] indicators for prior diagnoses of atherosclerotic cardiovascular disease (ASCVD) or chronic kidney disease (CKD), and indicators for prior metformin or SGLT2 inhibitor use. These covariates capture the primary factors influencing GLP-1 prescribing decisions, including the eligibility criteria in the VA semaglutide Criteria for Use (VA PBM, 2025). We then define $Z_{j(i)}$ as the leave-out mean residual from this regression, representing the average semaglutide prescribing tendency of PCP j (who treats patient i), net of differences in diabetes severity and treatment eligibility.[15]

Our sample consists of 4,737 PCPs. The average PCP has 292 diabetic or obese patients in 2018 (median: 297, SD: 161). Figure A.2 displays the distribution of $Z_{j(i)}$ and the corresponding first-stage relationship. There is substantial variation in PCPs' GLP-1 adoption propensity. For example, patients whose 2018 PCP ultimately falls at the 90th percentile of semaglutide adoption in 2021–2022 are 1.0 percentage point more likely to have ever received a GLP-1 prescription by 2019 compared with those whose PCP is at the 10th percentile (4.0% vs 3.0%). By the end of 2022, this 90–10 percentile gap widens as expected to 2.5 percentage points (9.9% vs 7.4%).

[13] An additional advantage of this design is that the study population remains fixed, avoiding changes in sample composition that could arise if newly eligible veterans—whose underlying health trajectories may differ—entered the VHA following policy expansions (e.g., PACT Act of 2022).

[14] The Veterans Health Administration is comprised of 130 stations, which are parent facilities geographically dispersed across the country.

[15] The outcome variable measures whether the patient received any semaglutide from their PCP, thus the propensity measure accounts for referrals from specialists (e.g., endocrinologists) that get refilled by PCPs.

Within-Patient (Provider) Event Study Equipped with GLP-1 adoption propensity $Z_{j(i)}$, we estimate the following event study between 2018 and 2024, four years before and five years after semaglutide's inclusion in the VA formulary:

$$Y_{it} = \sum_{k=2018}^{2024} \beta_k Z_{j(i)} \times \mathbb{1}\{t = k\} + \alpha_i + \lambda_t + \theta_{s(i),t} + \varepsilon_{it}, \tag{1}$$

where Y_{it} is the outcome of patient i in calendar quarter t (e.g., A1C, weight, utilization). λ_t are calendar time fixed effects. Patient fixed effects, α_i, absorb the main PCP propensity effect because each patient has only one PCP, and PCP fixed effects would absorb the main propensity effect $Z_{j(i)}$. We also control flexibly for station-year-quarter fixed effects, $\theta_{s(i),t}$, which account for regional differences in GLP-1 prescribing and patient outcomes over time.

The event study coefficients, β_k, are the parameters of interest. They trace out the relative outcome paths of patients whose providers ultimately become high vs low GLP-1 adopters. These coefficients correspond to a contrast between always- and never-prescribing PCPs for diabetic or obese patients, which lies outside the empirical support of observed prescribing variation and should be viewed as an extrapolation. In Section 5.7 we present treatment-on-the-treated effects of GLP-1 therapy. Heteroskedastic-robust standard errors are clustered at the PCP-level.

We also estimate and report an aggregate difference-in-differences specification, comparing patient outcomes before vs after 2021 (just after semaglutide was added to the VA formulary in November 2020):

$$Y_{it} = \beta^{DD} Z_{j(i)} \times \mathbb{1}\{t >= 2021\} + \alpha_i + \lambda_t + \theta_{s(i),t} + \varepsilon_{it} \tag{2}$$

Identifying Assumptions The coefficients β_k identify the causal effect of exposure to higher GLP-1 adopting PCPs under a parallel trends assumption. Formally, in the absence of semaglutide entry, patients of high and low propensity providers would have followed similar outcome trajectories after 2021. There must be no other shocks around 2021 that differentially affected patients of high vs low propensity PCPs. Importantly, selection on levels is permissible: patients could have differentially selected into providers prior to 2018, and providers could differ systematically in practice style. Patient fixed effects—which encompass provider fixed effects—absorb these differences. Threats to identification arise only from differential shocks or changes coinciding with semaglutide's entry—for example, if high-Z providers simultaneously altered other aspects of diabetes management and

practice styles around 2021.

We examine the plausibility of this assumption through pre-2021 event study dynamics, asking whether outcomes such as weight and A1C evolve similarly across high vs low GLP-1 adopting providers before 2021. Note that GLP-1 use was already rising prior to 2021 (Figure 1), therefore, we assess whether reduced form outcome trajectories move in line with the first stage adoption trajectories of GLP-1 therapies. In Section 5.6, we further assess and control for potential threats to identification.

5 Results

We begin by documenting the causal effect of high GLP-1 adopting PCPs on patients' GLP-1 utilization. Next, we turn to downstream health outcomes directly treated by GLP-1s—namely blood glucose and weight—followed by impacts on health, healthcare utilization and spending. We then examine heterogeneity in these outcomes in Section 5.5. Finally, in Section 5.6, we discuss mechanisms, arguing that our research design isolates the causal effect of GLP-1 use itself, rather than broader changes in clinical practice among high prescribers, and we compute average treatment effect-on-the-treated (ToT) estimates (Section 5.7). Throughout this section, we focus on event study dynamics (Equation 1), while aggregated post- vs pre-2021 (Equation 2) are reported in Table 2.

5.1 GLP-1 and Semaglutide Use

Figure 2 presents event study estimates from Equation (1) for four measures of GLP-1 utilization: any fill in the quarter (panel a), ever having filled by that quarter (panel b), total days supply in the quarter by that quarter (panel c), and cumulative days supply (panel d). Each panel reports estimates separately for any GLP-1 and specifically for semaglutide.

Patients whose 2018 PCP eventually becomes a high-prescriber exhibit a clear and steadily growing increase in GLP-1 use relative to patients of low-prescribers. By late 2020—immediately prior to semaglutide's addition to the VA formulary—these patients are approximately 10 percentage points (pp) more likely to fill a GLP-1 prescription. Following semaglutide's formulary inclusion, the trajectory steepens: GLP-1 utilization rises rapidly through 2023, reaching and plateauing at about a 20 pp difference—recall that this corresponds to the contrast between always- and never-

prescribing PCPs, and thus reflects an extrapolation beyond the observed range of adoption—by the end of 2024 These dynamics closely match the time series in Figure 1 and broader VA patterns in Figure A.1. Semaglutide shows a similar, but more abrupt, post-2021 jump, consistent with its rapid uptake due to its formulary status.

Cumulative ever-use (panel b) displays the expected smoothed version of these patterns, albeit reaching higher levels (30 pp in 2024). Quarterly days supply (panel c) mirrors the extensive margin until mid-2022, when we observe a noticeable decline. This decline coincides with the national GLP-1 shortage declared by the FDA in March 2022 (U.S. Food and Drug Administration, 2025) and rebounds after three quarters, before experiencing a small dip again in late 2023. This implies that the shortage affects the intensive but not the extensive margin: patients continued receiving GLP-1 prescriptions, but often at reduced dosage or with short gaps in supply.

By the end of 2024, cumulative days supply (panel d) shows large differences in total exposure. Patients whose initial PCP is a high-prescriber accumulate an additional 400 days supply of GLP-1s (and ca 300 for semaglutide) relative to those of low-prescribers. These large cumulative differences imply not only higher initiation rates but also meaningful increases along the intensive margin, with patients sustaining GLP-1 use over multiple years.

To further characterize persistence, Figure A.3 plots the empirical survival function of GLP-1 use since first initiation. While approximately 20% of patients discontinue after the first quarter, over 60% remain on therapy three to four years after initiation.[16]

In aggregate, these results demonstrate a strong and persistent first stage, lending credibility to our research design: a patient's pre-semaglutide-entry PCP identity is a powerful, precise, and persistent predictor of subsequent GLP-1 use through the end of 2024.

5.2 Glucose A1C and Body Weight

Having established a strong first stage effect of PCP adoption on GLP-1 use, we next examine impacts on A1C and weight—key clinical outcomes targeted by GLP-1 therapy. Panel (a) of Figure 3 shows that patients of high GLP-1 adopting PCPs begin to experience statistically significant declines in median A1C after formulary inclusion in late 2020. The post- vs pre-2021 coefficient is

[16]GLP-1 adherence in the VA appears substantially higher than in commercially-insured settings, likely reflecting lower out-of-pocket costs for GLP-1s within the VA system. Gleason et al. (2024) documents 32% adherence rate after one year while Wing et al. (2025) find 40-60% persistence.

−0.11 (column 3 of Table 2), relative to a baseline A1C of 7.12. Consistent with these improvements, the indicator for glycemic control (A1C $< 7\%$) increases around 2021, averaging approximately a 10 pp increase between 2022–2024. Over a baseline mean of 55 percent, this corresponds to an effect size of about 20%.

Weight outcomes exhibit similarly meaningful improvements. Panel (c) shows that weight declines for patients of high adopters, with differences reaching nearly 4 log points ($\approx 4\%$) by 2024 between high- and low-adopting PCPs. The corresponding post- vs pre-2021 average effect is 2.5 log points. For reference, the mean baseline weight in our sample in 2018 is 226.4 pounds. In addition, the indicator for BMI < 30 (the threshold for not being obese) rises steadily, eventually peaking at an improvement of ca 8 pp, almost a quarter of an increase over the baseline mean of 33% (panel d).

The trajectories of both A1C and weight exhibit modest pre-2021 movements, but these differences are small and statistically indistinguishable across patients of high- and low-propensity providers. Following semaglutide's formulary inclusion, outcomes diverge sharply, consistent with the rapid increase in GLP-1 utilization shown in Figure 2. This alignment is consistent with a pathway in which greater GLP-1 access—driven by PCP adoption tendencies—translates into improvements in metabolic outcomes. A complementary stacked difference-in-differences analysis based on initiation timing (Appendix F) yields similar qualitative conclusions, albeit with modest pre-trends in A1C.

It is not straightforward to benchmark these reduced form impacts of GLP-1 propensity on weight and A1C with clinical trial findings. We return to this in Section 5.7 when estimating treated-on-the-treated effects; our quasi-experimental estimates are quantitatively aligned with trial estimates. It is, however, worth noting that the sustained weight loss, yet plateauing—or even slightly rebounding—of A1C glucose are consistent with clinical trial evidence.[17] Finally, unlike medication and healthcare utilization, biomarkers are only observed when measured. We discuss implications of potentially selective measurement next.

[17]For example, SUSTAIN-1 (Sorli et al., 2017) documents rapid initial declines in A1C followed by a plateau after approximately 16 weeks (within a 30-week trial), while SUSTAIN-6 (Marso et al., 2016) reports partial rebound in A1C in week 16 onwards, following an early decline. Although our event studies span multiple years whereas clinical trials span weeks, the timing of cumulative GLP-1 exposure in the VA closely parallels these trial dynamics: panel (d) of Figure 2 shows that patients reach roughly a 16-week equivalent days supply of semaglutide at the end of 2022, and this is when A1C begins to plateau/rebound in both our event studies and the trials. Abstracting from differences in drug adherence, our real-world dynamic patterns align well with those observed in randomized trials. We formally benchmark our treatment-on-the-treated estimates to trial estimates in Section 5.7.

Observability of A1C and weight measurements Figure A.4 examines potential selection into weight and A1C measurements. Panel (a) shows no differential selection into weight measurement, consistent with the fact that weight is easily and routinely collected in primary care. By contrast, the likelihood of having an A1C measurement recorded declines over time (panel b). Clinically, this is expected as glycemic control improves and patients require less frequent A1C testing. However, differential testing intensity could bias our estimated A1C reductions if the observed measurements disproportionately come from a selected group of patients. Panel (c) addresses this concern by showing that, conditional on having an A1C measurement, the tested patients are increasingly those who were ex ante sicker—specifically, those with higher pre-2018 A1C levels.[18] This pattern is consistent with PCPs continuing to monitor glycemic control primarily for patients more severe baseline diabetes while reducing testing among those whose condition has stabilized. Consequently, the A1C declines we document earlier are likely conservative.

We directly address this potential selection with two robustness exercises in Figure A.5. Motivated by the rising share of ex ante high-A1C patients among those tested, panel (a) reweights patient observations in each year-quarter to match the baseline pre-2018 distribution of A1C severity. As expected, the estimated A1C reductions become larger in later years. In another robustness check, panel (b) collapses quarterly A1C values to yearly measures and estimates an event study on the 662,245 patients who have an A1C measurement in every year they are alive (excluding pandemic year 2020). The resulting yearly event study closely mirrors the baseline quarterly estimates in both pattern and magnitude, reinforcing that our main findings are not driven by differential selection into measurement.

5.3 Mental Health and Substance Use

One growing area of interest in GLP-1s is the emerging promise of its potential spillover impacts on mental health and substance use (Tempia Valenta et al., 2024; American Psychological Association, 2025). We examine these outcomes in Figure 4. Beginning with adverse events, we find no statistically significant changes in drug overdoses or alcohol intoxication emergency encounters or

[18] To assess this selection, we construct an outcome $w_{it} = \text{Pre2018A1C}_i \times \mathbb{1}\{\text{A1C measured}_{it}\}$ for patients with A1C measured that quarter (missing otherwise), and estimate $w_{it} = \sum_k \beta_k Z_{j(i)} \times \mathbf{1}\{t = k\} + \alpha_{j(i)} + \lambda_t + \varepsilon_{it}$, where $\alpha_{j(i)}$ are PCP fixed effects. Since each patient's pre-2018 A1C is time-invariant, PCP fixed effects remove cross-sectional differences in baseline composition across providers, so that the β_k coefficients identify whether, within a provider's panel, the composition of patients receiving A1C tests shifts toward those with higher or lower *baseline* A1C.

hospitalizations after 2021 for patients of high–GLP-1–adopting providers. These events are rare, and the corresponding reduced form estimates are imprecise; for example, the post– vs pre–2021 coefficient for any overdose poisoning in a calendar quarter is 0.011 pp (SE: 0.06), over a baseline mean of 0.11 percent.

Turning to clinical questionnaire measures, our estimates are more precise. PHQ-9 depression scores are stable: our post- vs pre-2021 estimate on screening below moderate depression is −0.0215 (SE: 0.059) over a baseline mean of 0.53. We do not detect a change in average alcohol use disorder score. Given the low mean score of 1.1—corresponding to having no more than two drinks, no more than once a month—is well within the low to no risk range, this is not surprising. However, we do detect a statistically significant increase (at the 5% level) of 2.3 pp in the likelihood of being below moderate risk of alcohol use disorder (AUD), over a baseline mean of 91%. Although we are cautious in interpreting this pattern, particularly given the absence of corresponding declines in alcohol-related ED events, it is directionally consistent with phase 2 clinical trial evidence showing reductions in heavy drinking among AUD patients undergoing GLP-1 therapy (Hendershot et al., 2025).

5.4 Healthcare Utilization and Spending

Given the improvements in glycemic control and weight, it is natural to ask whether patients of high GLP-1 adopting PCPs subsequently experience reductions in healthcare utilization. Figure 5 presents event study estimates for four outcomes in panels (a) through (d): any emergency department (ED) visit, outpatient spending, inpatient spending, and total (non-drug, non-GLP-1) medical spending. Each measure includes both VA care and VA-reimbursed community care.

Across all four outcomes, we observe little evidence of meaningful changes following the VA's addition of semaglutide to the national formulary in late 2020. The likelihood of any ED visit trends upward slightly, but the post- vs pre-2021 difference is small and statistically insignificant. Figure A.6 classifies ED visits by the major diagnostic category of the primary diagnosis. Consistent with the aggregate results, no category exhibits a significant decline after 2021. Several categories—including infectious and parasitic diseases, nervous system disorders, respiratory conditions, and symptoms/signs—show statistically significant increases. Clinical trial evidence from SUSTAIN 6 (Marso et al., 2016) find a reduction in major adverse cardiovascular events (myocardial

infarction and stroke) among patients with established cardiovascular disease or chronic kidney disease. We replicate this cardiovascular improvement in Table A.1, especially for those with existing cardiovascular disease or chronic kidney disease.

Outpatient and inpatient spending similarly show no detectable changes attributable to PCP adoption of GLP-1 therapy. Total medical spending exhibits a post- vs pre-2021 difference of −237.3 (standard error of 389.3; column 6 of Table 2) between high- and low-adopting PCPs, relative to a baseline mean of $3,181.

To complement contemporaneous spending, we examine cumulative spending separately for the pre-formulary (2018–2020) and post-formulary (2021-2024) periods. This reduced form cumulative approach improves interpretability by abstracting away from dynamic impacts of differential initiation timing. Figure A.7 shows that cumulative non–GLP-1 medical spending rises similarly across patients of high- and low-propensity PCPs prior to 2021, with no evidence of differential trends. Following 2021, cumulative spending continues to increase at a small comparable rate, with no divergence across groups despite increased GLP-1 use. By the end of 2024, the cumulative difference is small and imprecisely estimated ($1,405; SE: 10,680) relative to a mean of $67,779. The absence of differential cumulative spending both before and after formulary entry indicates that, despite substantial metabolic improvements, increased GLP-1 adoption does not translate into reductions in annual or cumulative healthcare utilization by the end of 2024 for the average diabetic or obese patient.

Distributional Effects While Figure 5 demonstrates a null spending effect of high GLP-1 adopters on average, this masks potentially significant heterogeneity in spending impacts. To address this, we explore quantile treatment effects of having a high versus low GLP-1 adopter using an estimation procedure from Callaway and Li (2019). The details on how we implement this procedure and its assumptions can be found in Appendix E. These quantile treatment effects should not be interpreted as aggregating to the mean effect, as they capture heterogeneous impacts across different points of the outcome distribution. The results in Figure A.8 demonstrate significant heterogeneity in the treatment effect of being treated by a high GLP-1 adopter. For example, at the 20th percentile of the spending distribution, seeing a high GLP-1 adopter is associated with a spending decline of $263 (SE=6.57), relative to seeing a low GLP-1 adopter. However, at the 80th percentile, seeing a high GLP-1 adopter is associated with a spending increase of $630 (SE=41.5).

5.5 Heterogeneity

Table 3 examines whether the impacts of PCP GLP-1 adoption differ across clinically relevant subgroups. Panel (a) splits patients by their 2018 diabetes and obesity status. All three groups show statistically significant increases in GLP-1 use following semaglutide's formulary inclusion, with the largest increases among patients who are both diabetic and obese. These groups also exhibit meaningful improvements in metabolic outcomes, albeit sometimes imprecise: A1C and weight decline across all subsamples. Effects on ED visits remain small and statistically indistinguishable from zero for each group.

Spending effects are generally imprecisely estimated. The only subgroup with a statistically significant change is patients who are both diabetic and obese, who show a large negative and marginally significant estimate for total medical spending. This estimate is somewhat imprecisely estimated (−$1,837 with a standard error of $737), and should be interpreted cautiously, but does indicate a potential benefit for the sickest and most eligible patients. Panel (b) shows broadly similar patterns by age: both younger and older adults exhibit comparable increases in GLP-1 use and improvements in A1C and weight, with no meaningful differences in utilization or spending. A1C reductions are largest for younger adults.

5.6 Assessing the Identifying Assumption

Our identifying assumption requires that patients of high- and low-propensity PCPs would have followed similar outcome trajectories after 2021 in the absence of GLP-1 adoption. A primary threat to this assumption is that GLP-1–adopting providers may have simultaneously changed other aspects of their clinical practice over time, violating the exclusion restriction. In this section, we assess the plausibility of the identifying assumption by examining trends in other prescribing behaviors (e.g., cardiovascular and diabetes medications) and by conducting a falsification test using a placebo cohort of patients who are not candidates for GLP-1 therapy.

Exclusion Restriction In Figure A.9, we show that patients of high GLP-1 adopting PCPs are not meaningfully more likely to be on antihypertensives or statins—two most common non-diabetic medications in our sample—over time or after November 2020. The pre- vs post-2021 coefficients are 0.005 (SE: 0.013) and −0.013 (SE: 0.015) for any antihypertensive and any statins, respectively,

over baseline means of 0.554 and 0.408. Patients also receive same rates of preventive care over time, as measured by take-up rates of seasonal flu vaccinations.

We observe a modest increase in the likelihood that patients receive any diabetes medication, primarily driven by growth in semaglutides, but also SGLT2 inhibitor prescriptions, and a corresponding decline in other diabetes drugs. The decline in older generation diabetes medications reflects substitution toward semaglutide, as expected following the adoption of a new therapeutic innovation and consistent with the "care-as-usual" counterfactual.

The concurrent rise in SGLT2 prescriptions, however, could pose a concern since SGLT2 inhibitors by themselves improve patient outcomes. Reassuringly, we find that this increase primarily reflects concurrent prescribing rather than confounding SGLT2 use: patients receiving GLP-1s are often co-prescribed SGLT2 inhibitors, consistent with clinical evidence supporting the combined use of these agents (Zinman et al., 2019) and with VA prescribing guidance (VA PBM, 2022). Panel (f) of Figure A.9 shows that a large fraction of SGLT2 prescriptions are written for patients also on GLP-1s (42% prior to November 2020 and 79% thereafter). Thus, the observed increase in SGLT2 use is consistent with recommended clinical practice and real-world use of GLP-1, rather than independent adoption behavior. Moreover, GLP-1 use without concurrent SGLT2 use also dominates SGLT2 monotherapy.

Nevertheless, we address any residual exclusion concerns in two ways. First, we conduct heterogeneity analyses among obese and non-diabetic patients in 2018, who are unlikely to receive SGLT2 inhibitors; the results are qualitatively similar to the full sample (Table 3 panel a). Second, we control for SGLT2 prescribing propensity in Equation (1), analogous to the semaglutide propensity and estimate event studies controlling for SGLT2 prescribing propensity interacted with time in Figure A.10.[19] Across both exercises, our main findings remain robust, supporting the interpretation that our results isolate the causal impact of GLP-1 medications rather than broader treatment bundling among GLP-1–adopting providers.[20]

[19]Specifically, we calculate the leave-out propensity to prescribe SGLT2 to pre-2018 diabetes patients over the 2021–2022 (while the first SGLT2 inhibitor was FDA approved in 2013, take-up in the VA and nationally did not rise until many years later; Shen and Farley, 2023), $Z_{j(i)}^{SGLT2}$. Then we estimate the following event study: $Y_{it} = \sum_{k=2018}^{2024} \beta_k^{GLP1} Z_{j(i)} \times \mathbf{1}\{t=k\} + \sum_{k=2018}^{2024} \beta_k^{SGLT2} Z_{j(i)}^{SGLT2} \times \mathbf{1}\{t=k\} + \alpha_i + \lambda_t + \theta_{s(i),t} + \varepsilon_{it}$ and report β_k^{GLP1}. The correlation between PCPs' GLP-1 propensity and SGLT2 propensity is 0.10, indicating while the two are related, there is ample variation across the different practice styles.

[20]We also note that to the extent that SGLT2 inhibitors independently reduce cardiovascular events and healthcare utilization (Bhattarai et al., 2022; Hong et al., 2019), their increased use would bias us against finding a null effect on utilization and spending attributable to GLP-1 adoption.

Placebo Check To assess whether high GLP-1 propensity PCPs changed other aspects of their care during the study period in ways that might confound our results, we conduct a falsification test on a placebo cohort of patients who are highly unlikely to receive GLP-1 prescriptions. If provider propensity is simply proxying for broader changes in provider behavior such as general improvements in clinical practice, we would expect to observe similar health improvements in this placebo group. Conversely, if the event study patterns in the main sample arise from GLP-1 adoption, estimates for this placebo cohort should be flat.

We construct the placebo cohort from VA patients who were younger than 40, non-diabetic, and non-obese in 2018. This group serves as a valid placebo sample because GLP-1 utilization rate reaches only 0.8% by the end of 2024. As in the main analysis, we restrict to patients with an assigned PCP in February 2018 whose PCP remained at the VA through 2021, resulting in a sample of 371,173 individuals. PCP propensity remains constructed on pre-2018 diabetic patients.

Figure A.11 presents the results of this falsification test. The first stage impact of PCP adoption propensity on GLP-1 is very small, peaking at 0.025 at the end of 2024 (panel a), an order of magnitude smaller than our baseline sample. Accordingly, panels (b) and (c) show no relationship between provider propensity and patients' A1C or weight, and panels (d) and (e) show no differential changes to emergency department visits or total medical spending. These null results indicate that high propensity providers are not generating spurious improvements in patient health through other channels, reinforcing that the effects observed in the diabetic and obese cohort are attributable to GLP-1 adoption rather than correlated changes in provider practice.

5.7 Treatment-on-the-Treated Estimates

Given the evidence supporting the exclusion restriction, we convert our reduced form event-study coefficients into average treatment-on-the-treated (ToT) effects—that is, the causal effect of *GLP-1 therapy* on patient outcomes. We estimate marginal dose response effects by scaling the reduced-form coefficient in quarter t by the corresponding first stage coefficient on total cumulative GLP-1 days supply through quarter t (Figure 2 panel d). This Wald-type ratio yields the effect of an additional unit of accumulated GLP-1 exposure, by quarter t, on outcomes in quarter t. For comparability with the SUSTAIN trials—which often evaluate 30-week treatment durations—we scale all effects in terms of a 30-week equivalent dose. Standard errors are computed using a

clustered bootstrap, resampling PCPs 500 times.

Interpreting the ToT estimates causally requires, in addition to our reduced form identifying assumptions, that PCP propensity strongly and monotonically shifts cumulative GLP-1 exposure. The resulting Wald estimator identifies a local average dose response effect of cumulative exposure for compliers at each horizon. Because cumulative exposure aggregates treatment histories, the estimate at a given quarter reflects a weighted average of marginal effects across patients at different durations since initiation. Finally, interpreting these effects in terms of an average 30-week equivalent dose requires that the impacts of exposure scales linearly and that the efficacy of different GLP-1 regimens (e.g., dosage and formulation over our study period) are comparable.

By scaling outcomes by cumulative GLP-1 exposure, the ToT estimates allows us to interpret outcome changes as reflecting the impact of a 30 week dose, regardless of the quarter used to calculate the estimate. Note that, because of this, our ToT estimates are relatively constant moving from 2020 to 2021. Appendix F presents a complementary stacked difference-in-differences design based on treatment initiation timing. That approach yields qualitatively similar conclusions but exhibits modest anticipatory pre-trends, consistent with endogenous treatment timing.

A1C and Weight ToT Effects, and Benchmarking to Clinical Trials Panels (a) and (b) of Figure 6 plot the ToT estimates over time for A1C and weight. Early years (2018–2019) exhibit substantial noise due to limited and specialized GLP-1 use and weak first stage impacts, so we display them in light transparent gray and zoom in on the estimates range beginning in 2020 and especially after 2021, for readability.

Average ToT effects peak shortly after semaglutide's formulary entry. In 2021, an accumulated 30-week GLP-1 regimen reduces weight by 3.28% (95% CI: $-4.55, -2.02$) and lowers A1C by 0.32 percentage points (95% CI: $-0.53, -0.10$).

We benchmark these quasi-experimental ToT estimates against randomized trial evidence. Because our real-world counterfactual is "care-as-usual"—such as SGLT2 inhibitors, insulin, etc.—the appropriate clinical benchmark is *active-comparator* trials (e.g., GLP-1 vs another diabetes

medication, as opposed to vs placebo), as in many SUSTAIN studies.[21,22] Our quasi-experimental estimates closely match the range of effects observed in the SUSTAIN active-comparator trials; indeed, the average effect across the nine active-comparator trials falls squarely within the 95% confidence intervals of our ToT estimates (Figure A.12).

Treatment effects attenuate substantially in later years. In 2024, a 30-week dose reduces weight by 1.82% (95% CI: $-2.23, -1.35$) and improves A1C by only a statistically insignificant 0.05 points (95% CI: $-0.11, 0.02$). This attenuation could be explained by diminishing marginal effects from additional cumulative exposure and/or selection on treatment: the early initiators and compliers in 2021 have larger treatment effects than later compliers.[23]

Indeed, the complier analysis in Figure A.13 confirms that the characteristics of compliers—patients whose GLP-1 use is induced by their PCP—change markedly between 2021 and 2024.[24] In the first year of expanded GLP-1 availability, compliers are twice as likely to have a pre-2018 A1C above 7%, 1.7 times as likely to have used prior diabetes medications, 1.4 times as likely to have a diabetes diagnosis, and 1.3 times as likely to exhibit more than four Elixhauser comorbidities, than the overall sample. Over the subsequent years, however, the complier population—which includes both marginal and inframarginal patients—becomes noticeably healthier. By 2024, the relative likelihoods of prior diabetes medication use, diabetes diagnosis, and high comorbidity burden all fall by 0.2 or more, indicating that later compliers are substantially lower risk than early compliers. Therefore, the diminishing in treatment effects over time can partly be explained by the complier population becoming healthier at baseline over time.

Spending ToT Effects Figure 6 panels (c) and (d) display any ED visit and total medical spending outcomes. Likelihood of any ED visit increases by 3.3 pp per 30-week GLP-1 regimen in 2021 (95% CI: 0.01, 0.06), and 1.7 pp in 2024 (95% CI: 0.01 0.03). ToT estimates on total quarterly

[21] There are some differences in study population. SUSTAIN participants were exclusively T2D patients, often selected for inadequate glycemic control or cardiovascular conditions, and are generally younger on average than our VA sample, which also includes obese patients without diagnosed diabetes.

[22] Some of the SUSTAIN trials compare semaglutide with older generation GLP-1s. Although these head-to-head comparisons capture the incremental benefits of semaglutide relative to earlier GLP-1 agents, we report them alongside active-comparator trials for completeness. Moreover, they aid in understanding heterogeneity across GLP-1 agents and interpreting the attenuation of our ToT effects over time, as the composition of GLP-1 therapies shifts from primarily semaglutide toward including new agents such as tirzepatide and Wegovy.

[23] A third potential explanation is differential adherence across cohorts. If later compliers discontinued treatment more quickly, lower persistence could attenuate average effects. However, Figure A.3 shows that persistence is, if anything, higher among later initiators, making differential adherence an unlikely driver of the attenuation.

[24] Details are provided in Appendix D.

medical spending are a statistically insignificant $1,058 in 2021 (95% CI: −856, 2972) and −$145 in 2024 (95% CI: −808, 519).

While the estimates on utilization and spending are less precisely estimated than those on weight and glucose, our 95% confidence intervals are able to reject sizable reductions in medical spending over the first 4 years of GLP-1 availability in the VA. ToT effects of a 30-week GLP-1 regimen on total medical spending over 2021–2024 (sum of the coefficients) are $5,042 (95% CI: −5076, 15160). Thus, we are able to reject reductions in medical spending (i.e., cost savings not accounting for GLP-1 drug spending) of more than 7.3% over the first four years ($= -\$5,076$ over a baseline mean of $69,067 total spending over 4 years). In other words, among the baseline sample of diabetic or obese veterans engaged in primary care in 2018—many of which have multiple existing comorbidities and are elderly—those induced into ever using a 30-week regimen of GLP-1s due to their PCPs, are unlikely to experience significant cost savings in the short run.

6 Conclusion

This paper provides large-scale quasi-experimental evidence on the real-world impacts of GLP-1 therapies by leveraging variation in the adoption decisions of veterans' pre-existing primary care providers. Using comprehensive health records and biomarker data from the Veterans Health Administration, we document substantial metabolic benefits of GLP-1 therapy. Patients whose providers become higher propensity GLP-1 adopters experienced clinically meaningful reductions in A1C and body weight. Our treatment-on-the-treated estimates for 2021—the first year in which semaglutide (Ozempic) became broadly available in the VA—closely align with efficacy estimates from the SUSTAIN clinical trials. These improvements, however, diminish over time and are substantially attenuated by 2024 as healthier patients initiate GLP-1s. This finding has important considerations for policymakers evaluating broadening access to GLP-1s.

Despite these metabolic gains, we find no evidence that increased GLP-1 use translated into short-run improvements in broader health outcomes or reductions in medical spending. Emergency department visits, inpatient and outpatient expenditures, and total non-GLP-1 spending remain statistically unchanged through 2024. These results suggest that, for the veteran population we study—many of whom are older and have long-standing chronic conditions—the clinical benefits of GLP-1 therapy do not translate into immediate cost savings for the average VA GLP-1 patient.

Our findings also have direct implications for reimbursement policy. A common argument for broad GLP-1 coverage is that large improvements in weight and glycemic control will quickly translate into lower downstream medical spending (Sexton Ward et al., 2023). Our estimates do not support that view in the short run. Although we document clinically meaningful improvements in A1C and body weight, we find no evidence that these gains reduce non-GLP-1 medical spending through 2024, and we can rule out sizable short-run cost offsets. This suggests that broad reimbursement should not be justified on the expectation that GLP-1 therapies will rapidly pay for themselves. Instead, the case for coverage rests on the value of the health gains themselves and any longer-run reductions in morbidity. Moreover, because treatment effects are largest among earlier, higher-risk compliers and attenuate as healthier patients initiate treatment, our results point to an important role for targeting reimbursement toward patients with the highest expected clinical benefit.

References

Abadie, Alberto, "Semiparametric instrumental variable estimation of treatment response models," *Journal of Econometrics*, 2003, *113* (2), 231–263.

Abaluck, Jason, Leila Agha, and Sachin Shah, "Trials Avoid High Risk Patients and Underestimate Drug Harms," Working Paper 34534, National Bureau of Economic Research December 2025.

Ahmann, Andrew J, Matthew Capehorn, Guillaume Charpentier, Francesco Dotta, Elena Henkel, Ildiko Lingvay, Anders G Holst, Miriam P Annett, and Vanita R Aroda, "Efficacy and safety of once-weekly semaglutide versus exenatide ER in subjects with type 2 diabetes (SUSTAIN 3): a 56-week, open-label, randomized clinical trial," *Diabetes care*, 2018, *41* (2), 258–266.

Ahrén, Bo, Luis Masmiquel, Harish Kumar, Mehmet Sargin, Julie Derving Karsbøl, Sanja Hald Jacobsen, and Francis Chow, "Efficacy and safety of once-weekly semaglutide versus once-daily sitagliptin as an add-on to metformin, thiazolidinediones, or both, in patients with type 2 diabetes (SUSTAIN 2): a 56-week, double-blind, phase 3a, randomised trial," *The lancet Diabetes & endocrinology*, 2017, *5* (5), 341–354.

Alalouf, Mattan, Sarah Miller, and Laura R. Wherry, "What Difference Does a Diagnosis Make? Evidence from Marginal Patients," *American Journal of Health Economics*, 2024, *10* (1).

Alsan, Marcella, John Cawley, Joseph J. Doyle, and Nicholas Skelley, "Mean Reversion in Randomized Controlled Trials: Implications for Program Targeting and Heterogeneous Treatment Effects," *AEA Papers and Proceedings*, May 2025, *115*, 221–225.

American Diabetes Association, "Understanding A1C Test," 2025. Accessed October 20, 2025.

American Enterprise Institute, "Ozempic and Your Community," https://www.aei.org/op-eds/ozempic-and-your-community/ 2024. Accessed: 2025-12-06.

American Psychological Association, "A New Era of Weight Loss: Mental Health Effects of GLP-1 Drugs," *Monitor on Psychology*, 2025, *56* (5), 66. Published July 1, 2025.

Angrist, Joshua, Amanda E. Kowalski, Ljubica Ristovska, and Marcia L. Stefanick, "Instrumental Variables Methods Reveal Larger Effects of Menopausal Hormone Therapy in the Landmark Women's Health Initiative Clinical Trial," *AEA Papers and Proceedings*, May 2025,

115, 215–220.

Aroda, Vanita R, Stephen C Bain, Bertrand Cariou, Milivoj Piletič, Ludger Rose, Mads Axelsen, Everton Rowe, and J Hans DeVries, "Efficacy and safety of once-weekly semaglutide versus once-daily insulin glargine as add-on to metformin (with or without sulfonylureas) in insulin-naive patients with type 2 diabetes (SUSTAIN 4): a randomised, open-label, parallel-group, multicentre, multinational, phase 3a trial," *The lancet Diabetes & endocrinology*, 2017, *5* (5), 355–366.

Ashraf, Amir Reza, Tim K. Mackey, János Schmidt, Győző Kulcsár, Róbert György Vida, Jiawei Li, and András Fittler, "Safety and Risk Assessment of No-Prescription Online Semaglutide Purchases," *JAMA Network Open*, 2024, *7* (8), e2428280. Published online August 2, 2024.

Bhalotra, Sonia, N. Meltem Daysal, and Mircea Trandafir, "Antidepressant Treatment in Childhood," Technical Report 2025.

Bhattacharya, Jay, Kate Bundorf, Noemi Pace, and Neeraj Sood, "Does health insurance make you fat?," Report, National Bureau of Economic Research 2009.

Bhattarai, Mukul, Mohsin Salih, Manjari Regmi, Mohammad Al-Akchar, Radhika Deshpande, Zurain Niaz, Abhishek Kulkarni, Momin Siddique, and Shruti Hegde, "Association of Sodium-Glucose Cotransporter 2 Inhibitors With Cardiovascular Outcomes in Patients With Type 2 Diabetes and Other Risk Factors for Cardiovascular Disease: A Meta-analysis," *JAMA Network Open*, 2022, *5* (1), e2142078. Published Online: January 5, 2022.

Bush, Karen, Daniel R. Kivlahan, Mary B. McDonell, Stephan D. Fihn, and Katharine A. Bradley, "The AUDIT-C: A Brief Screening Test for Problem Drinking: Ambulatory Care Quality Improvement Project (ACQUIP)," *Archives of Internal Medicine*, 1998, *158* (16), 1789–1795.

Cai, Miao, Taeyoung Choi, Yan Xie, and Ziyad Al-Aly, "Glucagon-like peptide-1 receptor agonists and risk of substance use disorders among US veterans with type 2 diabetes: cohort study," *bmj*, 2026, *392*.

Callaway, Brantly and Tong Li, "Quantile treatment effects in difference in differences model-swith panel data," *Quantitative Economics*, 2019, *10*, 1759–7323.

Capehorn, MS, A-M Catarig, JK Furberg, A Janez, HC Price, S Tadayon, B Vergès,

and M Marre, "Efficacy and safety of once-weekly semaglutide 1.0 mg vs once-daily liraglutide 1.2 mg as add-on to 1–3 oral antidiabetic drugs in subjects with type 2 diabetes (SUSTAIN 10)," *Diabetes & metabolism*, 2020, *46* (2), 100–109.

Cawley, John, "An Economy of Scales: A Selective Review of Obesity's Economic Causes, Consequences, and Solutions," *Journal of Health Economics*, 2015, *43*, 244–268.

_ , "Does Anything Work to Reduce Obesity? (Yes, Modestly)," *Journal of Health Politics, Policy and Law*, June 2016, *41* (3), 463–472.

Centers for Disease Control and Prevention, "National Diabetes Statistics Report," 2024. Accessed Oct 24, 2025. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight.

Chandra, Amitabh and Jonathan Skinner, "Technology Growth and Expenditure Growth in Health Care," *Journal of Economic Literature*, September 2012, *50* (3), 645–680.

Congressional Budget Office, "How Would Authorizing Medicare to Cover Anti-Obesity Medications Affect the Federal Budget?," Technical Report, Congressional Budget Office 2024. Accessed: 2025-12-06.

Currie, Janet and Jonathan Zhang, "Doing More with Less: Predicting Primary Care Provider Effectiveness," *The Review of Economics and Statistics*, 2025, *107* (2), 289–305.

Cutler, David M., "Are the Benefits of Medicine Worth What We Pay for It? 15th Annual Herbert Lourie Memorial Lecture on Health Policy," Center for Policy Research Reports 27, Center for Policy Research, Maxwell School, Syracuse University 2004.

_ **and Robert S. Huckman**, "Technological Development and Medical Productivity: The Diffusion of Angioplasty in New York State," *Journal of Health Economics*, March 2003, *22* (2), 187–217.

Davies, Melanie, Louise Færch, Ole K Jeppesen, Arash Pakseresht, Sue D Pedersen, Leigh Perreault, Julio Rosenstock, Iichiro Shimomura, Adie Viljoen, and Thomas A Wadden, "Semaglutide 2· 4 mg once a week in adults with overweight or obesity, and type 2 diabetes (STEP 2): a randomised, double-blind, double-dummy, placebo-controlled, phase 3 trial," *The Lancet*, 2021, *397* (10278), 971–984.

Daysal, N. Meltem, William N. Evans, Mikkel H. Pedersen, and Mircea Trandafir, "Do Medical Treatments Work for Work? Evidence from Breast Cancer Patients," NBER Working

Paper 32167, National Bureau of Economic Research 2024.

Dixon, Dave L., Teresa M. Salgado, Amanda Robinson, Salvatore Carbone, Tyler D. Wagner, Haroon Hyder, Bryan Kirschner, Kerri T. Musselman, Tonya M. Buffington, and Roy T. Sabo, "Prescribing Patterns of SGLT2 Inhibitors and GLP-1 Receptor Agonists in Patients With Type 2 Diabetes at Cardiology, Endocrinology, and Primary Care Visits," *Mayo Clinic Proceedings*, 2025, *100* (4), 647–656.

Eichmeyer, Sarah and Jonathan Zhang, "Pathways into Opioid Dependence: Evidence from Practice Variation in Emergency Departments," *American Economic Journal: Applied Economics*, 2022, *14* (4), 271–300.

_ **and** _ , "Primary Care Providers' Influence on Opioid Use and Its Adverse Consequences," *Journal of Public Economics*, 2023, *217*, 104784.

Frías, Juan P, Pernille Auerbach, Harpreet S Bajaj, Yasushi Fukushima, Ildiko Lingvay, Stanislava Macura, Anette L Søndergaard, Tsvetalina I Tankova, Nikolaos Tentolouris, and John B Buse, "Efficacy and safety of once-weekly semaglutide 2· 0 mg versus 1· 0 mg in patients with type 2 diabetes (SUSTAIN FORTE): a double-blind, randomised, phase 3B trial," *The Lancet Diabetes & Endocrinology*, 2021, *9* (9), 563–574.

Garthwaite, Craig L., "The Economic Benefits of Pharmaceutical Innovations: The Case of COX-2 Inhibitors," *American Economic Journal: Applied Economics*, 2012, *4* (3), 116–137.

Gleason, Patrick P., Benjamin Y. Urick, Landon Z. Marshall, Nicholas Friedlander, Yang Qiu, and R. Scott Leslie, "Real-world Persistence and Adherence to Glucagon-Like Peptide-1 Receptor Agonists Among Obese Commercially Insured Adults Without Diabetes," *Journal of Managed Care & Specialty Pharmacy*, 2024, *30* (8).

Hendershot, Christian S., Michael P. Bremmer, Michael B. Paladino, Georgios Kostantinis, Thomas A. Gilmore, Neil R. Sullivan, Amanda C. Tow, Sarah S. Dermody, Mark A. Prince, Robyn Jordan, Sherry A. McKee, Paul J. Fletcher, Eric D. Claus, and Klara R. Klein, "Once-Weekly Semaglutide in Adults With Alcohol Use Disorder: A Randomized Clinical Trial," *JAMA Psychiatry*, 2025, *82* (4). Original Investigation.

Hong, Dongzhe, Lei Si, Minghuan Jiang, Hui Shao, Wai kit Ming, Yingnan Zhao, Yan Li, and Lizheng Shi, "Cost Effectiveness of Sodium-Glucose Cotransporter-2 (SGLT2) Inhibitors, Glucagon-Like Peptide-1 (GLP-1) Receptor Agonists, and Dipeptidyl Peptidase-4 (DPP-

4) Inhibitors: A Systematic Review," *PharmacoEconomics*, 2019, *37*, 777–818.

Hwang, Jennifer H., Neda Laiteerapong, Elbert S. Huang, Dariush Mozaffarian, A. Mark Fendrick, and David D. Kim, "Fiscal Impact of Expanded Medicare Coverage for GLP-1 Receptor Agonists to Treat Obesity," *JAMA Health Forum*, 2025, *6* (4), e250089. Accessed: 2025-12-06.

Ippolito, Benedic and Joseph F. Levy, "Expanding Medicare Coverage of Anti-Obesity Medicines Could Increase Annual Spending by $3.1 Billion to $6.1 Billion," *Health Affairs*, 2024, *43* (9).

Kanwal, Fasiha, Jennifer R. Kramer, Liang Li, Yu-Xiao Yang, Yumei Cao, Xian Yu, Ronald Samuel, Basim Ali, Roxanne Desiderio, George Cholankeril, Mandeep Bajaj, Hashem B. El-Serag, and Steven M. Asch, "GLP-1 Receptor Agonists and Risk for Cirrhosis and Related Complications in Patients With Metabolic Dysfunction-Associated Steatotic Liver Disease," *JAMA Internal Medicine*, 2024, *184* (11), 1285–1296.

Kroenke, Kurt, Robert L. Spitzer, and Janet B. W. Williams, "The PHQ-9: Validity of a Brief Depression Severity Measure," *Journal of General Internal Medicine*, 2001, *16* (9), 606–613.

Lingvay, Ildiko, Andrei-Mircea Catarig, Juan P Frias, Harish Kumar, Nanna L Lausvig, Carel W le Roux, Desirée Thielke, Adie Viljoen, and Rory J McCrimmon, "Efficacy and safety of once-weekly semaglutide versus daily canagliflozin as add-on to metformin in patients with type 2 diabetes (SUSTAIN 8): a double-blind, phase 3b, randomised controlled trial," *The lancet Diabetes & endocrinology*, 2019, *7* (11), 834–844.

Liu, Ying, Sonica Sayam, Xiaonan Shao, Kesheng Wang, Shimin Zheng, Ying Li, and Liang Wang, "Prevalence of and Trends in Diabetes Among Veterans, United States, 2005–2014," *Preventing Chronic Disease*, 2017, *14*, E135.

Lu, Yuan, Yuntian Liu, Tedi Totojani, Chungsoo Kim, Rohan Khera, Hua Xu, John E. Jr. Brush, Harlan M. Krumholz, and Jason Abaluck, "Changes in Cardiovascular Risk Factors and Health Care Expenditures Among Patients Prescribed Semaglutide," *JAMA Network Open*, 2025, *8* (8), e2526013.

Lumbreras, Ainhoa Gómez, Malinda S. Tan, Lorenzo Villa-Zapata, Sabrina Ilham, Jacob C. Earl, and Daniel C. Malone, "Cost-effectiveness analysis of five anti-obesity medications from a US payer's perspective," *Nutrition, Metabolism and Cardiovascular Diseases*, June

2023, *33* (6), 1268–1276.

Lupkin, Sydney, "Thinking of buying obesity drugs online? Read this first," February 2025.

Marso, Steven P, Stephen C Bain, Agostino Consoli, Freddy G Eliaschewitz, Esteban Jódar, Lawrence A Leiter, Ildiko Lingvay, Julio Rosenstock, Jochen Seufert, and Mark L Warren, "Semaglutide and cardiovascular outcomes in patients with type 2 diabetes," *New England Journal of Medicine*, 2016, *375* (19), 1834–1844.

Michalak, Wojciech, Zhenxiang Zhao, Mads Faurby, Sara Alvarez, and Angela Fitch, "Impact of Semaglutide 2.4 mg on Healthcare Resource Utilization and Medical Costs in Patients with Atherosclerotic Cardiovascular Disease in the United States (SHINE-ASCVD)," *Journal of Medical Economics*, 2025, *28* (1), 1075–1085. Epub 2025 Jul 12.

Moiz, Areesha, Kristian B. Filion, Michael A. Tsoukas, Oriana H. Y. Yu, Tricia M. Peters, and Mark J. Eisenberg, "The expanding role of GLP_1 receptor agonists: a narrative review of current evidence and future directions," *eClinicalMedicine*, 2025, *86*, 103363. Accessed: 2025-12-06.

Montero, Alex, Grace Sparks, Marley Presiado, and Liz Hamel, "KFF Health Tracking Poll May 2024: The Public's Use and Views of GLP-1 Drugs," KFF May 2024.

Moshfegh, Jasmin, "Who Gains When Medicine Becomes More Precise? Evidence from Genomic Testing in Breast Cancer," Report, Working Paper 2025.

Oster, Emily, "Diabetes and Diet: Purchasing Behavior Change in Response to Health Information," *American Economic Journal: Applied Economics*, 2018, *10* (4), 308–348.

Parker, Emily D., Janice Lin, Troy Mahoney, Nwanneamaka Ume, Grace Yang, Robert A. Gabbay, Nuha A. ElSayed, and Raveendhara R. Bannuru, "Economic Costs of Diabetes in the U.S. in 2022," *Diabetes Care*, 2023, *47* (1), 26–43.

Pratley, Richard E, Vanita R Aroda, Ildiko Lingvay, Jörg Lüdemann, Camilla Andreassen, Andrea Navarria, and Adie Viljoen, "Semaglutide versus dulaglutide once weekly in patients with type 2 diabetes (SUSTAIN 7): a randomised, open-label, phase 3b trial," *The lancet Diabetes & endocrinology*, 2018, *6* (4), 275–286.

Rodbard, Helena W, Ildiko Lingvay, John Reed, Raymond De La Rosa, Ludger Rose, Danny Sugimoto, Eiichi Araki, Pei-Ling Chu, Nelun Wijayasinghe, and Paul Norwood, "Semaglutide added to basal insulin in type 2 diabetes (SUSTAIN 5): a randomized,

controlled trial," *The Journal of Clinical Endocrinology & Metabolism*, 2018, *103* (6), 2291–2301.

Shapiro, Bradley T., "Promoting Wellness or Waste? Evidence from Antidepressant Advertising," *American Economic Journal: Microeconomics*, 2022, *14* (2), 439–477.

Shen, Tsung-Hua and Joel F. Farley, "Utilization and impact of SLGT2 inhibitors among diabetes patients in a nationally representative survey: Findings from NHANES 2013–2020," *Journal of Diabetes and its Complications*, 2023, *37* (11), 108625.

Sorli, Christopher, Shin ichi Harashima, George M Tsoukas, Jeffrey Unger, Julie Derving Karsbøl, Thomas Hansen, and Stephen C Bain, "Efficacy and safety of once-weekly semaglutide monotherapy versus placebo in patients with type 2 diabetes (SUSTAIN 1): a double-blind, randomised, placebo-controlled, parallel-group, multinational, multicentre phase 3a trial," *The lancet Diabetes & endocrinology*, 2017, *5* (4), 251–260.

The Economist, "GLP-1s like Ozempic are among the most important drug breakthroughs ever," *The Economist*, October 2024. Accessed: 2025-12-06.

U.S. Department of Veterans Affairs, "Annual Benefits Report, Fiscal Year 2018," https://www.benefits.va.gov/REPORTS/abr/docs/2018-abr.pdf 2019. Accessed: 2025-12-04.

_ , "Community Care: Veteran Care Overview," 2023. Accessed: 2025-12-02.

_ , "Corporate Data Warehouse (CDW)," 2025. Accessed: 2025-05 to 2025-12.

U.S. Department of Veterans Affairs, Pharmacy Benefits Management Services, Medical Advisory Panel, and VISN Pharmacist Executive, "Guidance for Semaglutide 2mg (OZEMPIC) Injection," https://www.va.gov/formularyadvisor/DOC_PDF/CG_Semaglutide_2mg_OZEMPIC_Guidance_Jun_2022.pdf June 2022. Accessed: September 19, 2025.

U.S. Department of Veterans Affairs, Pharmacy Benefits Management Services, Medical Advisory Panel, and VISN Pharmacist Executives, "Semaglutide (OZEMPIC) Injection for Type 2 Diabetes Mellitus: Criteria for Use July 2022," https://web.archive.org/web/20230103135947/https://www.va.gov/formularyadvisor/DOC_PDF/CFU_Semaglutide_OZEMPIC_Jul22a.pdf July 2022. Accessed: May 8, 2028.

_ , "Semaglutide (OZEMPIC) Injection for Type 2 Diabetes Mellitus: Criteria for Use," https://www.va.gov/formularyadvisor/DOC_PDF/CFU_Semaglutide_OZEMPIC_Criteria_March_2025.pdf March 2025. Accessed: September 19, 2025.

_ , "Semaglutide (WEGOVY) Subcutaneous Injection for Weight Management Criteria for Use Au-

gust 2025," https://www.va.gov/formularyadvisor/DOC_PDF/CFU_Semaglutide_WEGOVY_for_weight_management_CFU_Rev_Aug_2025.pdf March 2025. Accessed: May 8, 2026.

U.S. Food and Drug Administration, "Declaratory Order: Resolution of Shortages of Semaglutide Injection Products (Ozempic and Wegovy)," https://www.fda.gov/media/185526/download?attachment February 2025. Accessed: September 19, 2025.

VA Health Economics Resource Center, "HERC Average Cost Data for Outpatient Care," 2025. Accessed October 20, 2025.

Valenta, Silvia Tempia, Alba Nicastri, Federica Perazza, Federica Marcolini, Valentina Beghelli, Anna Rita Atti, and Maria Letizia Petroni, "The Impact of GLP-1 Receptor Agonists (GLP-1 RAs) on Mental Health: A Systematic Review," *Current Treatment Options in Psychiatry*, 2024, *11* (4), 310–357.

Wagner, Todd H., Shuo Chen, and Paul G. Barnett, "Using Average Cost Methods to Estimate Encounter-Level Costs for Medical-Surgical Stays in the VA," *Medical Care Research and Review*, 2003, *60* (3_suppl), 15S–36S.

Wang, Lindsey, William Wang, David C. Kaelber, and et al., "GLP-1 Receptor Agonists and Colorectal Cancer Risk in Drug-Naive Patients With Type 2 Diabetes, With and Without Overweight/Obesity," *JAMA Oncology*, 2024, *10* (2), e235514.

Wang, William, Nora D. Volkow, Quangqiu Wang, Nathan A. Berger, Pamela B. Davis, David C. Kaelber, and Rong Xu, "Semaglutide and Opioid Overdose Risk in Patients With Type 2 Diabetes and Opioid Use Disorder," *JAMA Network Open*, 2024, *7* (9), e2431129.

Ward, Alex Sexton, Adam Stefanski, Bryan Tysinger, Phuong Nguyen, Dana Goldman, and Darius Lakdawalla, "Benefits of Medicare Coverage for Weight Loss Drugs," White Paper, USC Schaeffer Center for Health Policy and Economics 2023.

Wennberg, David, Henriette Coetzer, Amy Marr, Caroline Margiotta, Razia Hashmi, Jason Kay, Vikrant Vats, Stuart Hagen, Laura Anatale-Tardiff, Sandhya Rao, Mark Mugiishi, and Jonathan Skinner, "The Real-World Costs of GLP-1 Receptor Agonist Treatment," *medRxiv*, 2025. Preprint.

Wilding, John PH, Rachel L Batterham, Salvatore Calanna, Melanie Davies, Luc F Van Gaal, Ildiko Lingvay, Barbara M McGowan, Julio Rosenstock, Marie TD Tran, and Thomas A Wadden, "Once-weekly semaglutide in adults with overweight or obesity," *New*

England Journal of Medicine, 2021, *384* (11), 989–1002.

Wing, Coady, Sih-Ting Cai, Daniel Sacks, and Kosali Simon, "Do GLP-1 Medications Pay for Themselves?," 2025. Working Paper.

World Health Organization, "Diabetes," 2024. Accessed Aug 26, 2025. https://www.who.int/health-topics/diabetes#tab=tab_1.

_ , "WHO Guideline on the Use of Glucagon-Like Peptide-1 (GLP-1) Therapies for the Treatment of Obesity in Adults," Licence: CC BY-NC-SA 3.0 IGO 2025. Accessed: YYYY-MM-DD.

Xie, Yan, Taeyoung Choi, and Ziyad Al-Aly, "Mapping the Effectiveness and Risks of GLP-1 Receptor Agonists," *Nature Medicine*, 2025, *31*, 951–962. Published January 20, 2025.

Yanovski, Susan Z. and Jack A. Yanovski, "Long-term Drug Treatment for Obesity: A Systematic and Clinical Review," *JAMA*, January 2014, *311* (1), 74–86.

Zinman, Bernard, Vaishali Bhosekar, Robert Busch, Ingrid Holst, Bernhard Ludvik, Desirée Thielke, James Thrasher, Vincent Woo, and Athena Philis-Tsimikas, "Semaglutide once weekly as add-on to SGLT-2 inhibitor therapy in type 2 diabetes (SUSTAIN 9): a randomised, placebo-controlled trial," *The lancet Diabetes & endocrinology*, 2019, *7* (5), 356–367.

Tables

Table 1: Summary Statistics

	Mean	SD	Q1	Median	Q3
Female	0.079				
Age in 2018	64.8	13.8	57	68	73
Black	0.196				
White	0.708				
Asian American/Pacific Islander	0.023				
Native American or Alaska Native	0.010				
Hispanic Ethnicity	0.073				
Type 2 Diabetes	0.686				
Obesity	0.554				
Diabetic and Obese	0.241				
ASCVD or CKD	0.271				
Elixhauser Comorbidities, Count	3.6	2.2	2	3	5
Outcomes in 2018:					
A1C, %	7.1	1.5	6.0	6.7	7.8
Weight, lbs	226	48	193	222	254
BMI	33.1	6.3	28.7	32.4	36.6
GLP-1	0.011				
Biguanides (Metformin)	0.227				
Other Diabetes Medication	0.243				
Any Diabetes Medication	0.351				
Any Emergency Department Visit, Quarterly	0.090				
Total Medical Spending, Quarterly	3,181	26,763	56	760	2,280
Outcomes in 2022:					
A1C, %	7.0	1.5	5.9	6.7	7.7
Weight, lbs	221	49	188	217	250
BMI	32.3	6.5	27.8	31.7	36
GLP-1	0.054				
Biguanides (Metformin)	0.229				
Other Diabetes Medication	0.267				
Any Diabetes Medication	0.372				
Any Emergency Department Visit, Quarterly	0.120				
Total Medical Spending, Quarterly	4,818	20,586	0	961	3026
Outcomes in 2024:					
A1C, %	6.9	1.4	5.9	6.6	7.6
Weight, lbs	219	48	185	214	247
BMI	32	6.5	27.5	31.3	35.7
GLP-1	0.101				
Biguanides (Metformin)	0.224				
Other Diabetes Medication	0.295				
Any Diabetes Medication	0.401				
Any Emergency Department Visit, Quarterly	0.135				
Total Medical Spending, Quarterly	5,531	17,307	82	1,190	3,741
Number of observations:					
Patients	1,374,666				
PCPs	4,737				

Notes: This table presents summary statistics for our baseline sample of patients diagnosed with diabetes or obesity prior to 2018 and were receiving VA primary care in 2018. We report summary statistics on demographics and existing comorbidities (prior to 2018), as well as key outcomes (A1C, weight, body mass index, medication use, any quarterly emergency department visits, and total quarterly medical spending) in 2018, 2022, and 2024. Statistics for outcomes are conditional on being alive. Spending is measured in 2025 dollars.

Table 2: Aggregate Outcomes, Post vs Pre Semaglutide Included in VA Formulary

	Dependent variable:					
	Any GLP-1	Any Semaglutide	A1C	Log Weight	Any ED	Total Spending
	(1)	(2)	(3)	(4)	(5)	(6)
Propensity × Post2021	0.1271***	0.1838***	−0.1121**	−0.0249***	0.0083	−237.3
	(0.0133)	(0.0144)	(0.0465)	(0.0048)	(0.0079)	(389.3)
N (patient-year-quarter)	34,293,608	34,293,608	10,435,204	15,131,936	34,293,608	33,246,463
Mean Dep. Var. (2018 Q1)	0.011	0.00001	7.119	5.400	0.090	3,181

Notes: This table displays the aggregate post-2021 vs pre-2021 coefficient (β^{DD}) estimated from Equation (2) on select main outcomes. The main effects on PCP propensity and post-2021 are absorbed by patient and station-by-year-quarter fixed effects and therefore cannot be reported. Spending data is only available until 2024 Q3, resulting in one quarter fewer observations. Mean dependent variables are calculated at baseline (2018 Q1). Patient-year-quarter observations where the patient is dead are dropped. Heteroskedastic-robust standard errors are clustered at the PCP-level. ***: 0.01, **: 0.05, *: 0.1.

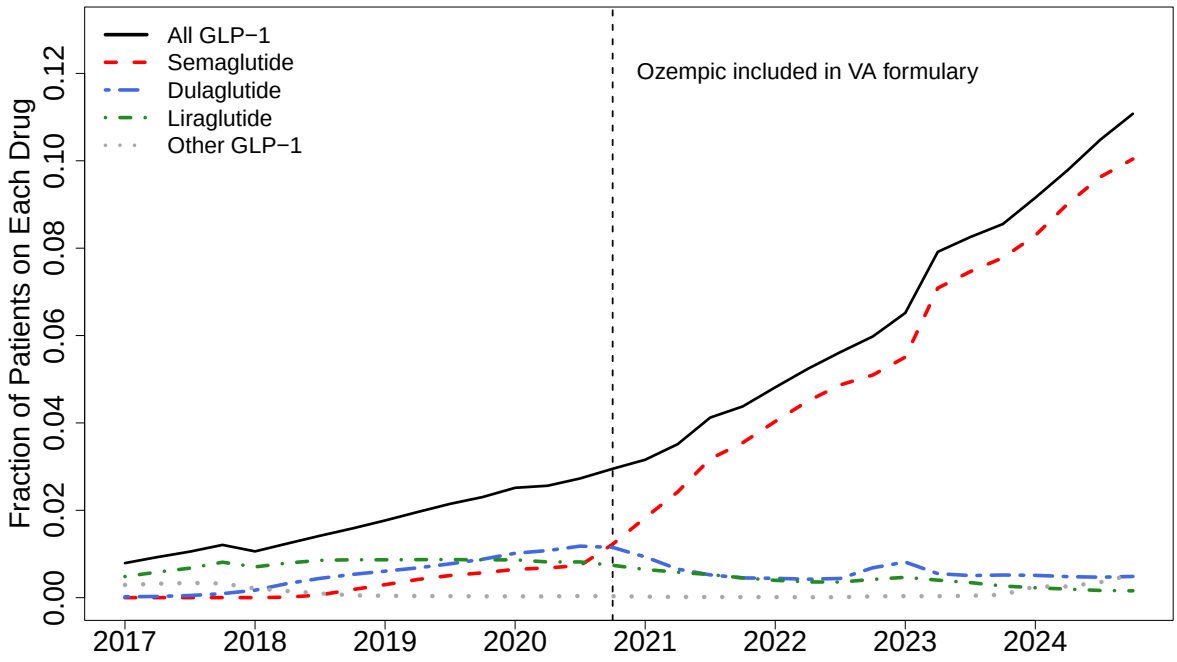
Table 3: Heterogeneity: Aggregate Outcomes, Post- vs Pre-Semaglutide Included in VA Formulary

	Dependent variable:					
	Any GLP-1	A1C	Log Weight	Any ED	Total Spending	N
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel (a): Diabetic vs Obese			
Diabetic only	0.1288***	−0.0458	−0.0158***	0.0127	−174.2	612,428
	(0.0167)	(0.0617)	(0.0046)	(0.0096)	(503.5)	
Mean Dep. Var.	0.008	7.36	5.29	0.083	2,966.9	
Obese only	0.0440***	−0.1551***	−0.0102	−0.0019	12.8	431,039
	(0.0084)	(0.0483)	(0.0062)	(0.0094)	(359.0)	
Mean Dep. Var.	0.0005	5.71	5.47	0.083	2,410.1	
Diabetic & Obese	0.2192***	−0.1197	−0.0249***	−0.0063	−1,836.5**	331,199
	(0.0267)	(0.0811)	(0.0066)	(0.0128)	(737.2)	
Mean Dep. Var.	0.028	7.53	5.50	0.113	3,982.7	
			Panel (b): Age			
Age $\leq$ 55	0.1297***	−0.3010***	−0.0200***	−0.0016	60.36	311,039
	(0.0181)	(0.0968)	(0.0071)	(0.0106)	(377.6)	
Mean Dep. Var.	0.008	6.91	5.47	0.099	2,479	
Age $>$ 55	0.1265***	−0.0537	−0.0232***	0.0092	−479.5	1,063,627
	(0.0140)	(0.0469)	(0.0042)	(0.0085)	(454.4)	
Mean Dep. Var.	0.011	7.16	5.38	0.088	3,200.4	

Notes: This table explores heterogeneity by displaying the aggregate post-2021 vs pre-2021 coefficient (β^{DD}) estimated from Equation (2) on select main outcomes (as in Table 2) estimated on various subsamples. Panel (a) splits the sample by prior 2018 diabetes and obesity status, and panel (b) splits the subsample by younger (under 55) and older (over 55) patients in 2018. Mean dependent variables are calculated at baseline (2018 Q1). Column 6 displays the number of patients in each subsample. Patient-year-quarter observations where the patient is dead are dropped. Heteroskedastic-robust standard errors are clustered at the PCP-level. ***: 0.01, **: 0.05, *: 0.1.

Figures

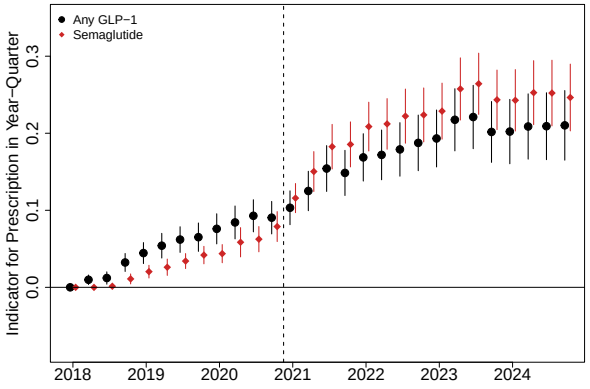
Figure 1: Rates of GLP Use Among Baseline Diabetic or Obese Sample



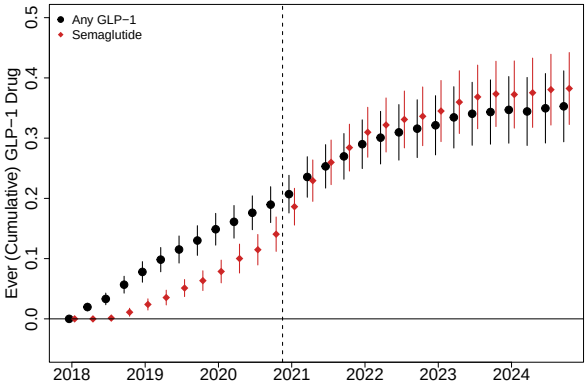
Notes: This figure shows the rise in GLP-1 use among the baseline sample (patients diagnosed with diabetes or obesity prior to 2018 and were receiving VA primary care in 2018). The fraction is calculated for patients who are alive in that year-quarter. Other GLP-1s include albiglutide, exenatide, lixisenatide, and tirzepatide. The dashed vertical line corresponds to November 2020, when semaglutide (Ozempic) was added to the VA national formulary. Figure A.1 shows the time series for the entire VA population, which follows a nearly identical pattern.

Figure 2: GLP-1 Medication Utilization

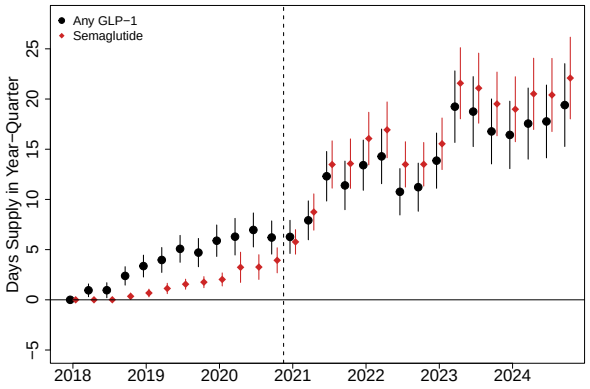
(a) Any Quarterly Utilization



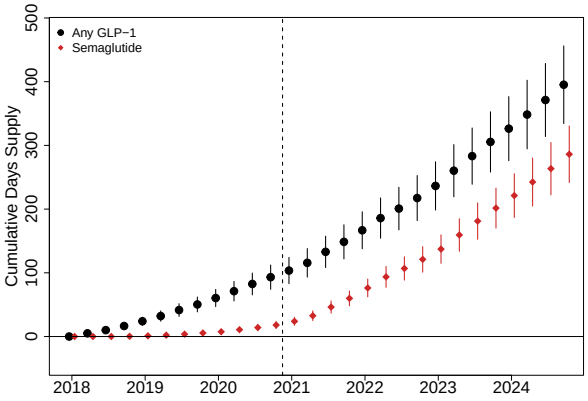
(b) Any Utilization, Cumulative



(c) Quarterly Days Supply

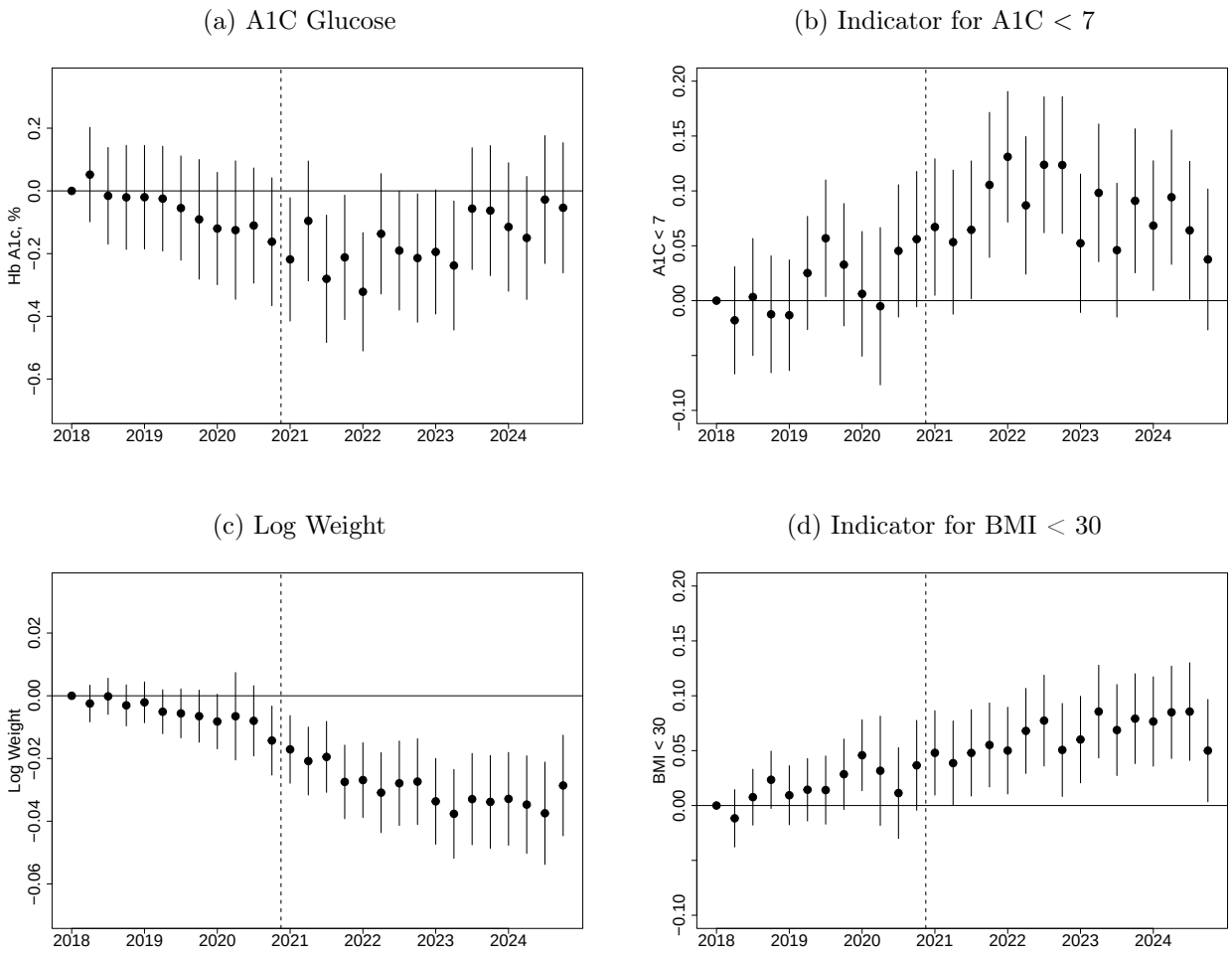


(d) Cumulative Days Supply



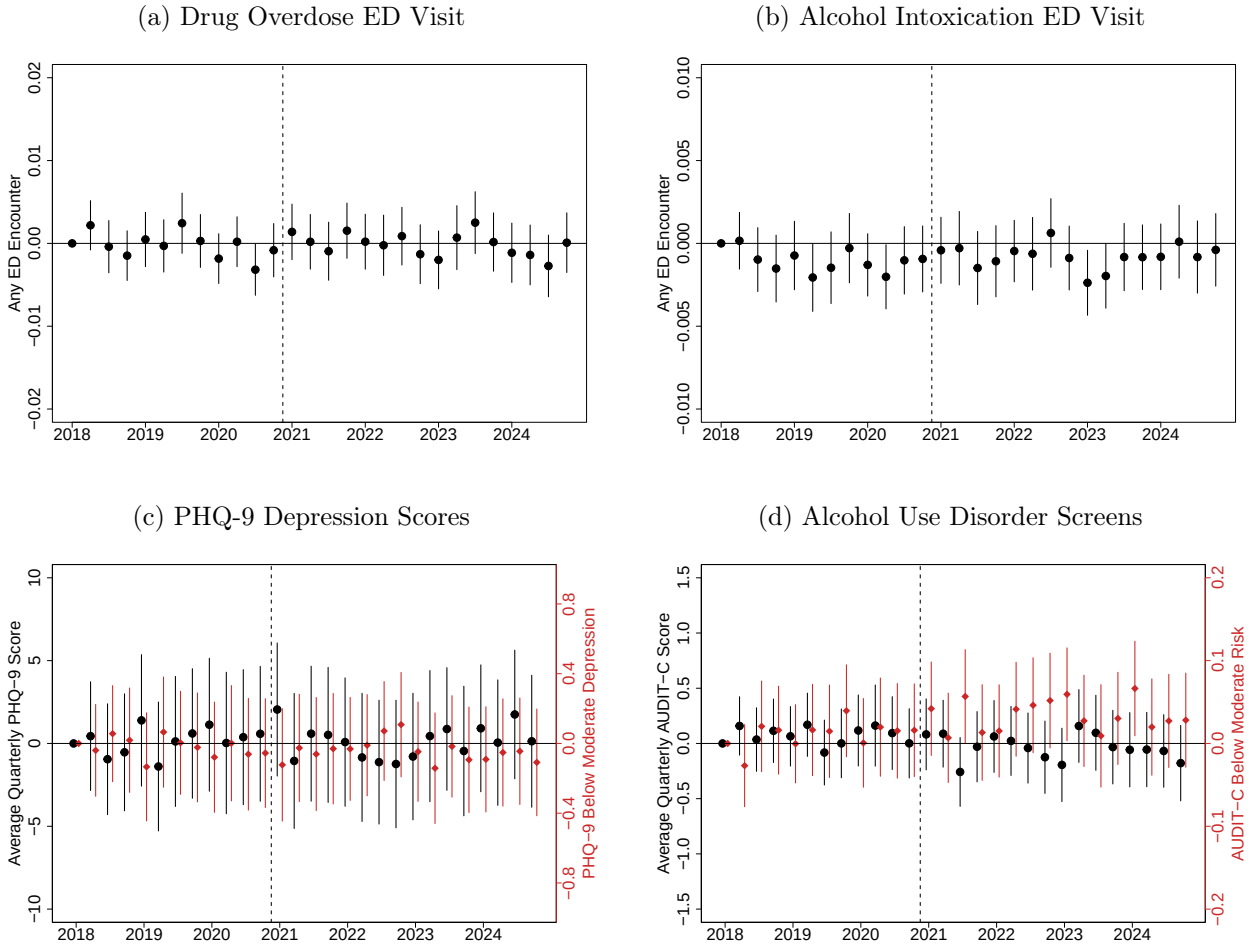
Notes: This figure shows the first stage impact of PCP propensity to adopt GLP-1s on GLP-1 outcomes estimated from Equation (1). Panel a shows any GLP-1 and any semaglutide (a specific GLP-1) in that year-quarter as the outcome variable. Panel b shows whether the patient has received any GLP-1 (and any semaglutide) up to that year-quarter (cumulative). Panel c and d reports the same contemporaneous and cumulative measures for total days supply. Patient-year-quarter observations where the patient is dead are dropped. Heteroskedastic-robust standard errors are clustered at the PCP-level, and 2018 Q1 is the omitted base period. The dashed vertical line corresponds to November 2020, when semaglutide (Ozempic) was added to the VA national formulary.

Figure 3: Blood Glucose and Weight Outcomes



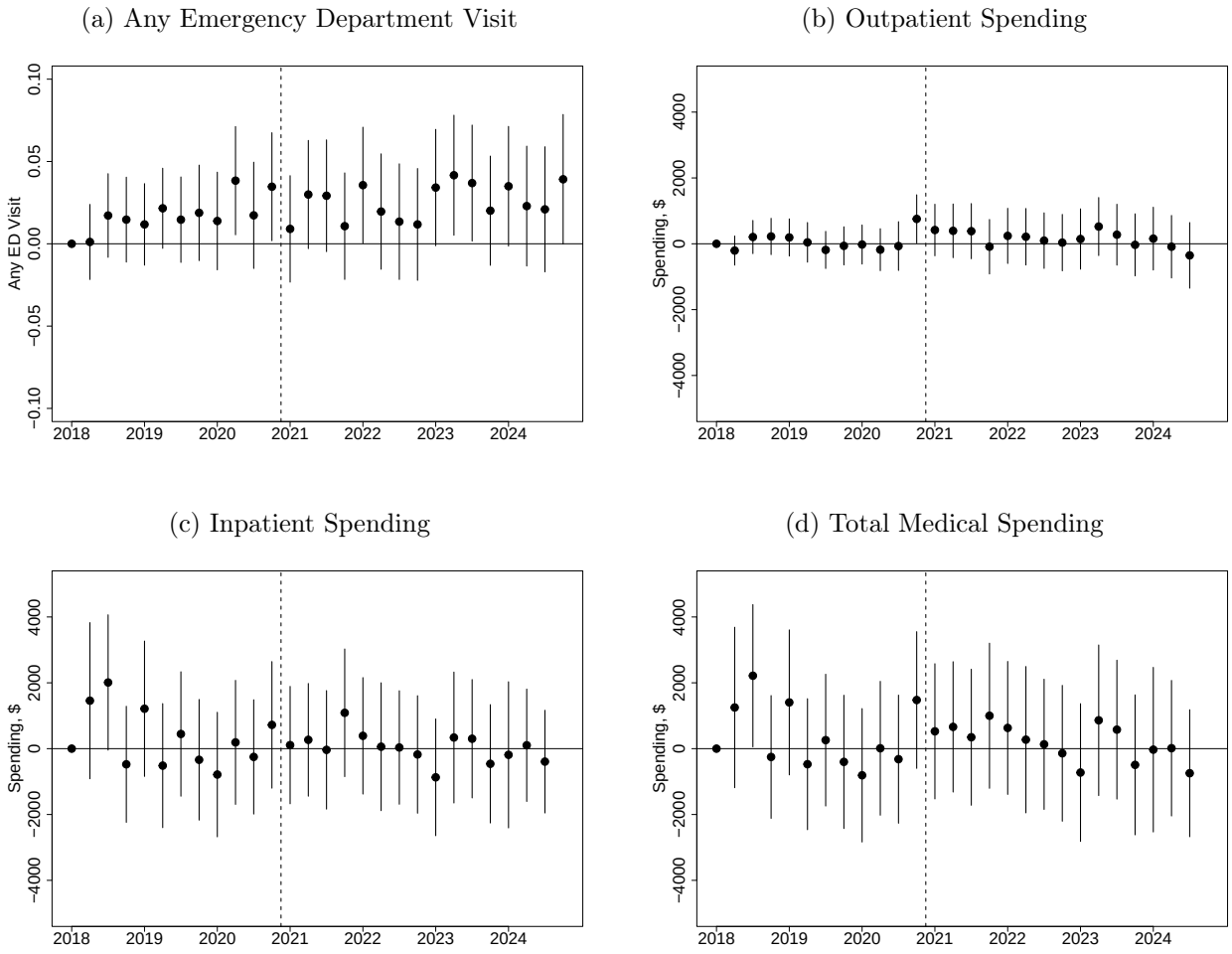
Notes: This figure shows the reduced form impact of PCP propensity to adopt GLP-1s on hemoglobin A1C blood glucose and weight outcomes estimated from Equation (1). Panels a to d display median quarterly A1C, indicator for median A1C that quarter below 7%, log of weight (in pounds), and indicator for body mass index below 30 (threshold for obesity). Means in the base period (2018 Q1) are 7.12 for A1C, 0.55 for A1C < 7, 5.4 for log weight (226.4 lbs), and 0.33 for BMI < 30. The specification is estimated on a sample of patient-year-quarter observations with non-missing weight or A1C observations, and patient is alive; selection into measurement is shown in Figure A.4. Heteroskedastic-robust standard errors are clustered at the PCP-level, and 2018 Q1 is the omitted base period. The dashed vertical line corresponds to November 2020, when semaglutide (Ozempic) was added to the VA national formulary.

Figure 4: Mental Health and Substance Use Outcomes



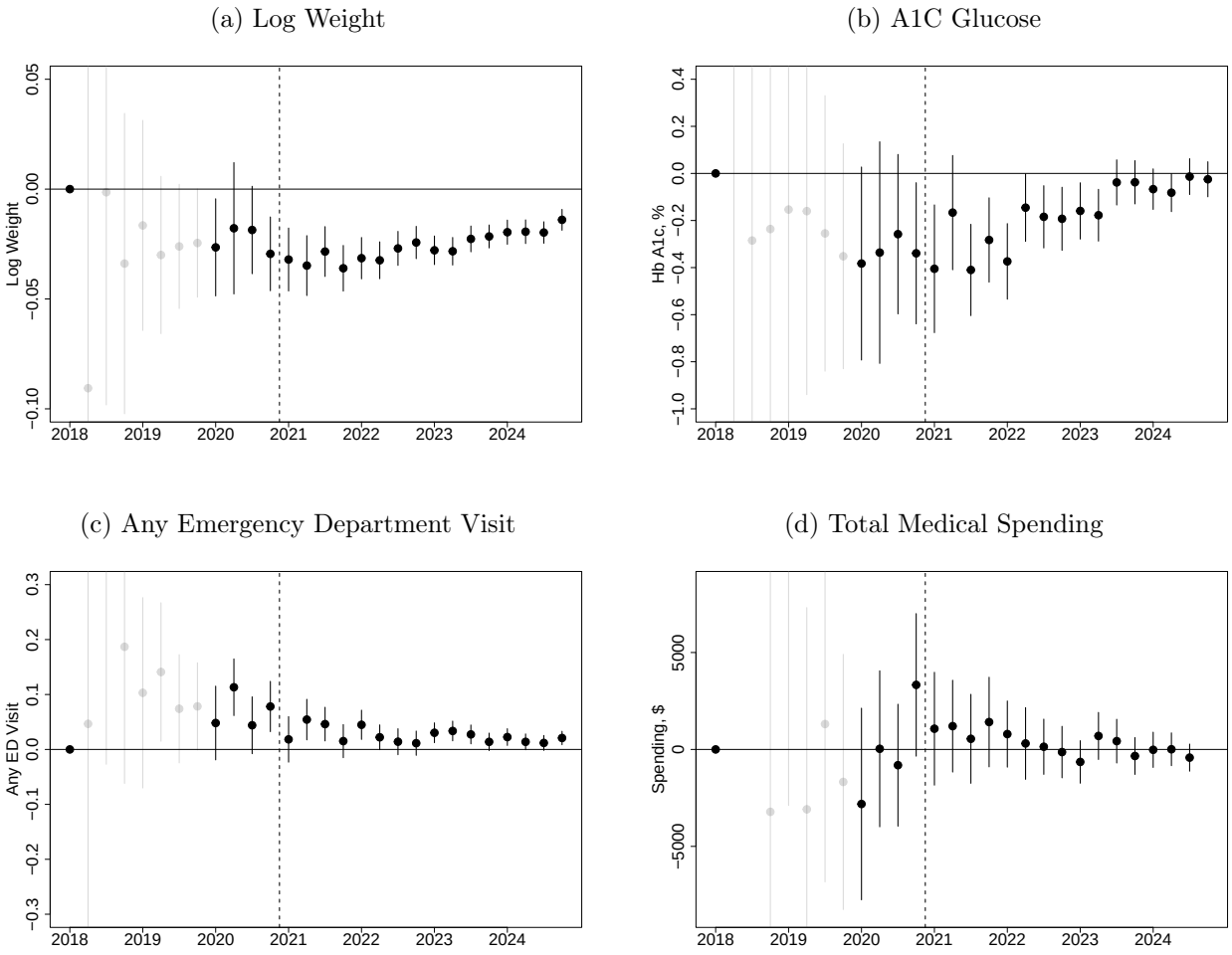
Notes: This figure displays shows the reduced form impact of PCP propensity to adopt GLP-1s on mental health and substance use outcomes estimated from Equation (1). Panels a to d display any drug overdose ED encounter or hospitalization (panel a), any alcohol intoxication ED encounter or hospitalization (panel b), average PHQ-9 depression scores and PHQ-9 score below moderate depression in black and red (panel c), and average alcohol use disorder screen (AUDIT-C) and below moderate risk in black and red (panel d). Means in the base period (2018 Q1) are 0.0011 for overdose, 0.0005 for alcohol intoxication, 9.55 for average PHQ-9, 0.53 for PHQ-9 below moderate, 1.09 for average AUDIT-C, and 0.91 for AUDIT-C below moderate risk. The specification is estimated on a sample of patient-year-quarter observations when the patient is alive, and for clinical questionnaires, when they are conducted. Heteroskedastic-robust standard errors are clustered at the PCP-level, and 2018 Q1 is the omitted base period. The dashed vertical line corresponds to November 2020, when semaglutide (Ozempic) was added to the VA national formulary.

Figure 5: Utilization and (Non-GLP-1) Spending Outcomes



Notes: This figure shows the reduced form impact of PCP propensity to adopt GLP-1s on healthcare utilization and cost outcome estimated from Equation (1). Panels a to d display indicator for any emergency department (ED) visit, total outpatient spending, total inpatient spending, and total outpatient and inpatient spending. All outcomes include VA utilization and non-VA community utilization reimbursed by the VA and do not include prescription drugs (and hence GLP-1 spending is not included). All spending are in 2025 dollars. Means in the base period (2018 Q1) are 0.090 for any ED visit, $2,188 for outpatient spending, $993 for inpatient spending, and $3,181 for total medical spending. Heteroskedastic-robust standard errors are clustered at the PCP-level, and 2018 Q1 is the omitted base period. The dashed vertical line corresponds to November 2020, when semaglutide (Ozempic) was added to the VA national formulary.

Figure 6: Average Treatment-on-the-Treated Effects of 30-Week Dose of Accumulated GLP-1



Notes: The figure plots the treatment-on-the-treated (ToT) estimates of the impact of a marginal accumulated 30-week dose GLP-1 treatment by quarter t, on outcomes in quarter t. These ToT are estimated via a Wald-type ratio of the reduced form impact of higher propensity GLP-1 adopting PCPs divided by the first stage impact measured in terms of accumulated 30-week days supply of GLP-1s. Standard errors are computed using a clustered bootstrap, resampling PCPs 500 times. ToT estimates prior to 2020 are unstable and noisy due to limited GLP-1 use and a weaker first stage. Therefore, for readability, we plot those in transparent gray points and zoom in on the y-axis values that span post-2020 ToT estimates. The dashed vertical line corresponds to November 2020, when semaglutide (Ozempic) was added to the VA national formulary.

Appendix

Table of Contents

A Appendix Tables and Figures **A.2**

B Details on the Institutional Setting **A.18**

C Details on Variable Definitions **A.19**

D Complier Analysis **A.21**

E Quantile Regressions **A.22**

F Difference-in-Differences: Early- vs Later-Treated **A.24**

A Appendix Tables and Figures

Table A.1: Major Adverse Cardiovascular Events: Myocardial Infarction and Stroke

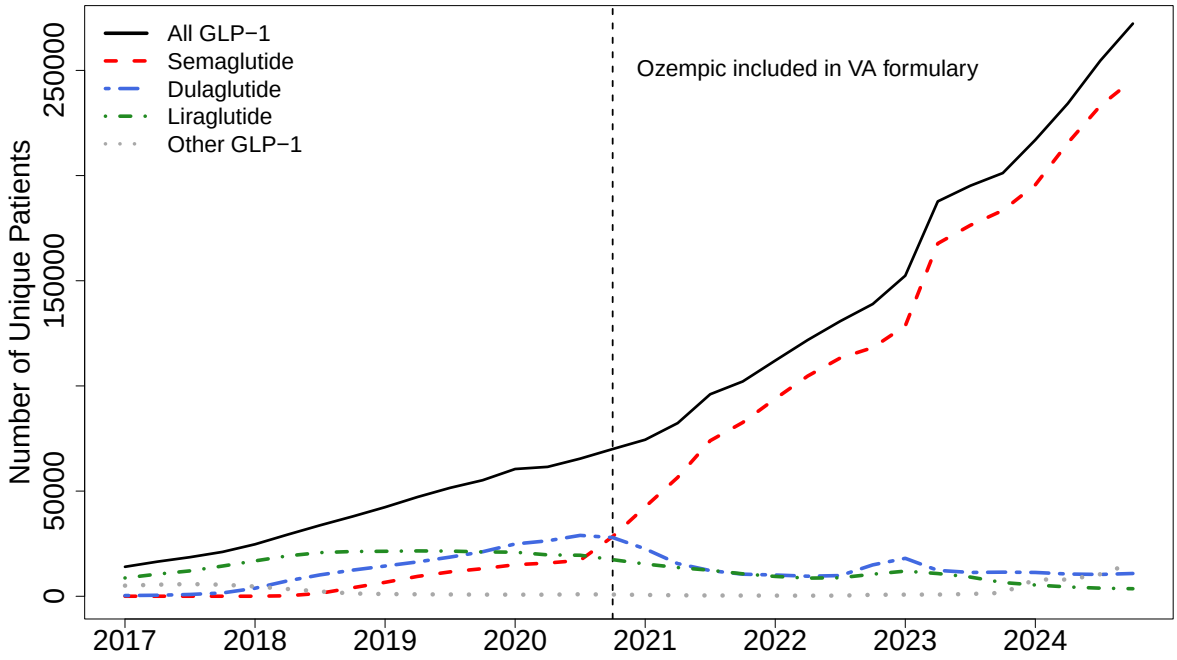
	Sample:	
	Full Sample	Prior ASCVD or CKD
	(1)	(2)
Propensity × Post2021	−0.0036	−0.0157**
	(0.0026)	(0.0065)
N (patient-year-quarter)	34,293,608	8,425,648
Mean Dep. Var.	0.011	0.021

Notes: This table displays the aggregate post-2021 vs pre-2021 coefficient on an indicator variable of whether the patient had an myocardial infarction (heart attack) or stroke, in an ED or hospitalization, in a given quarter. The outcome variable and the subsample by prior Atherosclerotic Cardiovascular Disease or Chronic Kidney Disease closely approximate the SUSTAIN 6 clinical trial (Marso et al., 2016) which focuses on patients with established cardiovascular disease, chronic heart failure, or chronic kidney disease. Mean dependent variables are calculated at baseline (2018 Q1). Patient-year-quarter observations where the patient is dead are dropped. Heteroskedastic-robust standard errors are clustered at the PCP-level. ***: 0.01, **: 0.05, *: 0.1.

Table A.2: Summary of Results from the SUSTAIN Clinical Trial Program

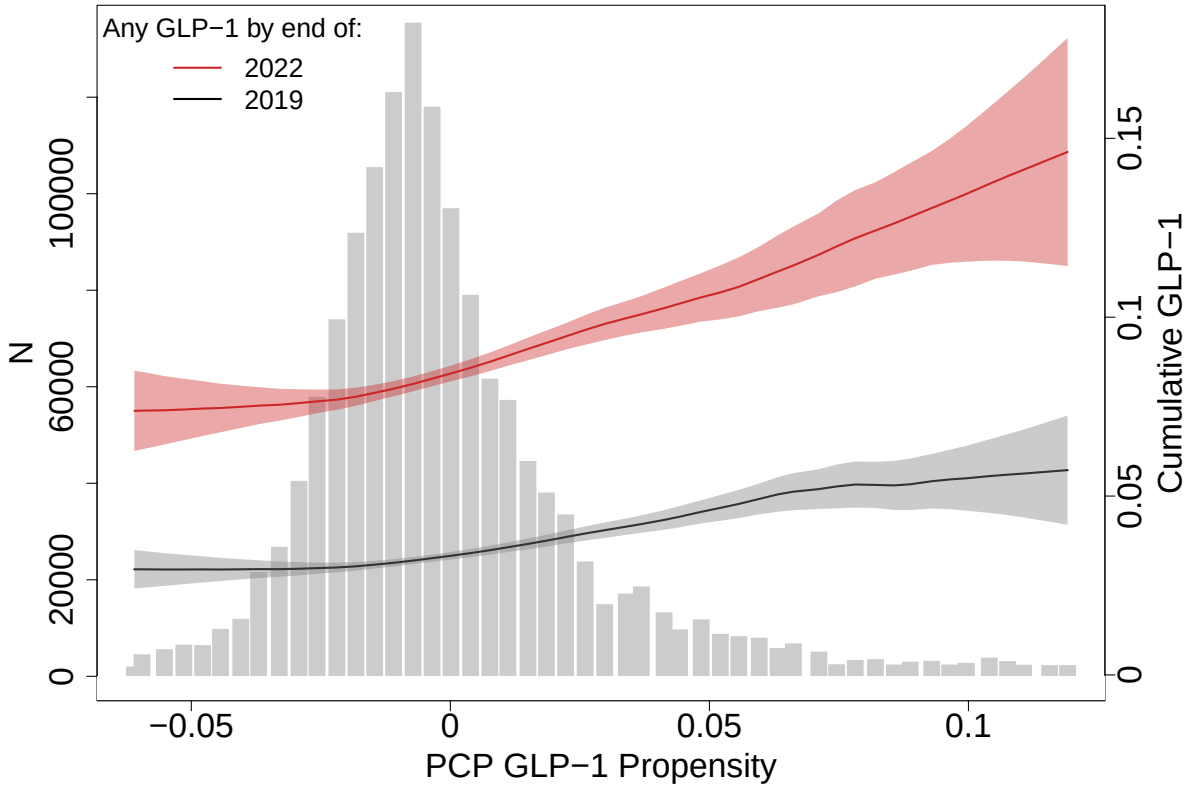
Study	Sample	Design	Results	Conclusion
Sorli, Christopher, et al. "Efficacy and safety of once-weekly semaglutide monotherapy versus placebo in patients with type 2 diabetes (SUSTAIN 1): a double-blind, randomised, placebo-controlled, parallel-group, multinational, multicentre phase 3a trial." The lancet Diabetes & endocrinology 5.4 (2017): 251-260.	* Treatment-naive patients with type 2 diabetes who had insufficient glycaemic control with diet and exercise alone. * International study. * Between February 3, 2014, and August 21, 2014 * 388 participants	* NCT02054897 *Random assignment to once-weekly subcutaneously injected semaglutide (0·5 mg or 1·0 mg), or volume-matched placebo (0·5 mg or 1·0 mg), for 30 weeks via prefilled PDS290 pen-injectors	* Primary outcome (HbA1c) reduction of 1.43percent in 0.5mg arm, and 1.53percent in 1 mg arm * Body weight reduction 2.75 and 3.56 kg	*"Semaglutide significantly improved HbA1c and bodyweight in patients with type 2 diabetes compared with placebo, and showed a similar safety profile to currently available GLP-1 receptor agonists, representing a potential treatment option for such patients"*
Ahrén, Bo, et al. "Efficacy and safety of once-weekly semaglutide versus once-daily sitagliptin as an add-on to metformin, thiazolidinediones, or both, in patients with type 2 diabetes (SUSTAIN 2): a 56-week, double-blind, phase 3a, randomised trial." The lancet Diabetes & endocrinology 5.5 (2017): 341-354.	*Type 2 diabetes with insufficient glycaemic control and treatment with metformin, thiazolidinediones, or both *International study *Between Dec 2, 2013, and Aug 5, 2015 *1231 participants	* NCT01930188 * Random assignment to 1) subcutaneous semaglutide 0.5mg once weekly + oral sitagliptin once daily, 2) subcutaneous semaglutide 1.0mg once weekly + oral sitagliptin once daily, 3) oral sitagliptin 100mg once daily plus subcutaneous semaglutide placembo 0.6mg once weekly, 4) oral sitagliptin 100mg once daily plus subcutaneous semaglutide placembo 1.0 mg once weekly. The latter two were pooled for the analysis	*Primary outcome (HbA1c) reduction -0.77percent with semaglutide 0.5mg, -1.06percent with semaglutide 1.0mg * Body weight -2.35kg with semaglutide 0.5mg, 4.20kg with semaglutide 1.0mg	*"Once-weekly semaglutide was superior to sitagliptin at improving glycaemic control and reducing bodyweight in participants with type 2 diabetes on metformin, thiazolidinediones, or both, and had a similar safety profile to that of other GLP-1 receptor agonists. Semaglutide seems to be an effective add-on treatment option for this patient population."*
Ahmann, Andrew J., et al. "Efficacy and safety of once-weekly semaglutide versus exenatide ER in subjects with type 2 diabetes (SUSTAIN 3): a 56-week, open-label, randomized clinical trial." Diabetes care 41.2 (2018): 258-266.	*Type 2 diabetes taking oral antidiabetic drugs (metformin, and/or sulfonylureas, and/or thiazolidinedione) *International study *Between December 2013 and July 2015 *813 participants	*NCT01885208 *Random assignment to semaglutide 10.mg or exenatide ER 2.0mg	*Primary outcome (HbA1c) reduction -0.62percent *Body weights -3.78kg	*"Semaglutide 1.0 mg was superior to exenatide ER 2.0 mg in improving glycemic control and reducing body weight after 56 weeks of treatment; the drugs had comparable safety profiles. These results indicate that semaglutide treatment is highly effective for subjects with type 2 diabetes who are inadequately controlled on oral antidiabetic drugs."*
Aroda, Vanita R., et al. "Efficacy and safety of once-weekly semaglutide versus once-daily insulin glargine as add-on to metformin (with or without sulfonylureas) in insulin-naive patients with type 2 diabetes (SUSTAIN 4): a randomised, open-label, parallel-group, multicentre, multinational, phase 3a trial." The lancet Diabetes & endocrinology 5.5 (2017): 355-366.	*Type 2 diabetes with insufficient glycaemic control with metformin alone or in combination with a sulfonylurea *International study *Between Aug 4, 2014, and Sept 3, 2015 *1089 partcipants	*NCT02128932 *Random assignment to subcutaneous once weekly 0.5mg semaglutide, subcutaneous once weekly 0.5mg semaglutide, once-daily insulin glargine	*Primary outcome (HbA1c) reduction -0.38percent with 0.5mg semaglutide, -0.81percent with 1.0mg semaglutide *Body weight -4.62 kg with 0.5mg semaglutide, 6.33 kg with 1.0 mg semaglutide	"Compared with insulin glargine, semaglutide resulted in greater reductions in HbA1c and weight, with fewer hypoglycaemic episodes, and was well tolerated, with a safety profile similar to that of other GLP-1 receptor agonists."
Rodbard, Helena W., et al. "Semaglutide added to basal insulin in type 2 diabetes (SUSTAIN 5): a randomized, controlled trial." The Journal of Clinical Endocrinology & Metabolism 103.6 (2018): 2291-2301.	*Uncontrolled type 2 diabetes with basal insuling or basal insulin+metformin *International study *Between 1 December 2014 through 21 Novemeber 2015 *397 participants	*NCT02305381 *Subcutaneous semaglutide 0.5mg or 1.0mg once weekly or volume -matched placebo	*Primary outcome (HbA1c) reduction -1.35percent in semaglutide 0.5mg arm and -1.75percent in semaglytide 1.0mg arm *Body weight -2.31kg in semaglutide 0.5mg arm and -5.06kg in semaglutide 1.0mg arm	*"Semaglutide, added to basal insulin, significantly reduced HbA1c and body weight in patients with uncontrolled T2D vs placebo"*
Marso, Steven P., et al. "Semaglutide and cardiovascular outcomes in patients with type 2 diabetes." New England Journal of Medicine 375.19 (2016): 1834-1844. (SUSTAIN 6)	*Type 2 diabetes on standard care regimen *International study *Between February 2013 through December 2013 *2375 participants	*NCT01720446 *Once weekly semaglutide (0.5 mg or 1.0mg) or placebo for 104 weeks	*Primary outcome (first occurrence of cardiovascular death, nonfatal myocardial infarction, or nonfatal stroke): hazard ratio of 0.74	*"In patients with type 2 diabetes who were at high cardiovascular risk, the rate of cardiovascular death, nonfatal myocardial infarction, or nonfatal stroke was significantly lower among patients receiving semaglutide than among those receiving placebo, an outcome that confirmed the noninferiority of semaglutide."*
Pratley, Richard E., et al. "Semaglutide versus dulaglutide once weekly in patients with type 2 diabetes (SUSTAIN 7): a randomised, open-label, phase 3b trial." The lancet Diabetes & endocrinology 6.4 (2018): 275-286.	*Type 2 diabetes and HbA1c 7.0-10.5percent with metformin monotherapy *International study *Between Jan 6, 2016, and June 22, 2016 *1201 participants	*NCT02648204 *Once weekly semaglutide 0.5 mg, dulaglutide 0.75 mg, semaglutide 1.0mg, or dulaglutide 1.5 mg subcutaneously	*Primary outcome (HbA1c) reduction of -0.40ppt for semaglutide 0.5mg arm compared to dulaglutide 0.75mg arm, reduction of -0.41ppt for semaglutide 1.0mg arm compared to dulaglutide 1.5 mg arm *reduction of -2.26kg for semaglutide 0.5mg arm compared to dulaglutide 0.75mg arm, reduction of -3.55kg for semaglutide 1.0mg arm compared to dulaglutide 1.5 mg arm	*"At low and high doses, semaglutide was superior to dulaglutide in improving glycaemic control and reducing bodyweight, enabling a significantly greater number of patients with type 2 diabetes to achieve clinically meaningful glycaemic targets and weight loss, with a similar safety profile."*
Lingvay, Ildiko, et al. "Efficacy and safety of once-weekly semaglutide versus daily canagliflozin as add-on to metformin in patients with type 2 diabetes (SUSTAIN 8): a double-blind, phase 3b, randomised controlled trial." The lancet Diabetes & endocrinology 7.11 (2019): 834-844.	*Uncontrolled type 2 diabetes on daily metformin therapy *International study *Between March 15 2017, and Nov 16, 2018 *788 patients	*NCT03136484 * Random assignment to semaglutide 1.0mg once weekly or oral canagliflozin (an SGLT2 inhibitor) 300 mg once daily	* Primary outcome (HbA1c) reduction of -0.49 ppt in semaglutide arm * Body weight -1.06 kg	*"Once-weekly semaglutide 1·0 mg was superior to daily canagliflozin 300 mg in reducing HbA1c and bodyweight in patients with type 2 diabetes uncontrolled on metformin therapy. These outcomes might guide treatment intensification choices."*
Zinman, Bernard, et al. "Semaglutide once weekly as add-on to SGLT-2 inhibitor therapy in type 2 diabetes (SUSTAIN 9): a randomised, placebo-controlled trial." The lancet Diabetes & endocrinology 7.5 (2019): 356-367.	* Patients with type 2 diabetes and HbA1c 7.0-10.0percent with at least 90 days of treatment with an SGLT-2 inhibitor *International study. *Between March 15, and dec 4, 2017 * 302 participants	* NCT03086330 *Random assignment to once weekly semaglutide 1.0mg, afer a dose-escalation schedule of 4 weeks of 0.25 mg, and 4 weeks of 0.5mg semalgutide	* Primary outcome (HbA1c) reduction of 1.42percent in treatment vs control * Body weight -3.81 kg	*"Adding semaglutide to SGLT-2 inhibitor therapy significantly improves glycaemic control and reduces bodyweight in patients with inadequately controlled type 2 diabetes, and is generally well tolerated"*
Capehorn, M. S., et al. "Efficacy and safety of once-weekly semaglutide 1.0 mg vs once-daily liraglutide 1.2 mg as add-on to 1–3 oral antidiabetic drugs in subjects with type 2 diabetes (SUSTAIN 10)." Diabetes & metabolism 46.2 (2020): 100-109.	* Patients with type 2 diabetes and HbA1c 7.0-11.0percent on 1-3 oral antidiabetic drugs *International study *Between June and November 2017 * 577 participants	*NCT03191396 * Random assignment to once-weekly subcutaneous semaglutide 1.0mg or once-daily subcutaneous liraglutide 1.2mg	*Primary outcome (HbA1c) reduction of -0.69 percent *Body weight -3.83 kg	*"Semaglutide was superior to liraglutide in reducing HbA1c and body weight. Safety profiles were generally similar, except for higher rates of gastrointestinal AEs with semaglutide vs liraglutide."*
Frías, Juan P., et al. "Efficacy and safety of once-weekly semaglutide 2· 0 mg versus 1· 0 mg in patients with type 2 diabetes (SUSTAIN FORTE): a double-blind, randomised, phase 3B trial." The Lancet Diabetes & Endocrinology 9.9 (2021): 563-574.	* Inadequately controlled type 2 diabetes with metformin * International study. * Between June 19, 2019, and Nov 28, 2019 * 961 participants	* NCT03989232; EudraCT, 2018-004529-96; and WHO, U1111-1224-5162. * Random assignment to once-weekly semaglutide 2.0mg or 1.0mg	Primary outcome (HbA1c) reduction of 0.23 ppt (trial product estimand) or 0.18 ppt (treatment policy estimand)	*"Semaglutide 2·0 mg was superior to 1·0 mg in reducing HbA1c, with additional bodyweight loss and a similar safety profile."*

Figure A.1: Rates of GLP-1 Use Among Baseline Diabetic or Obese Sample



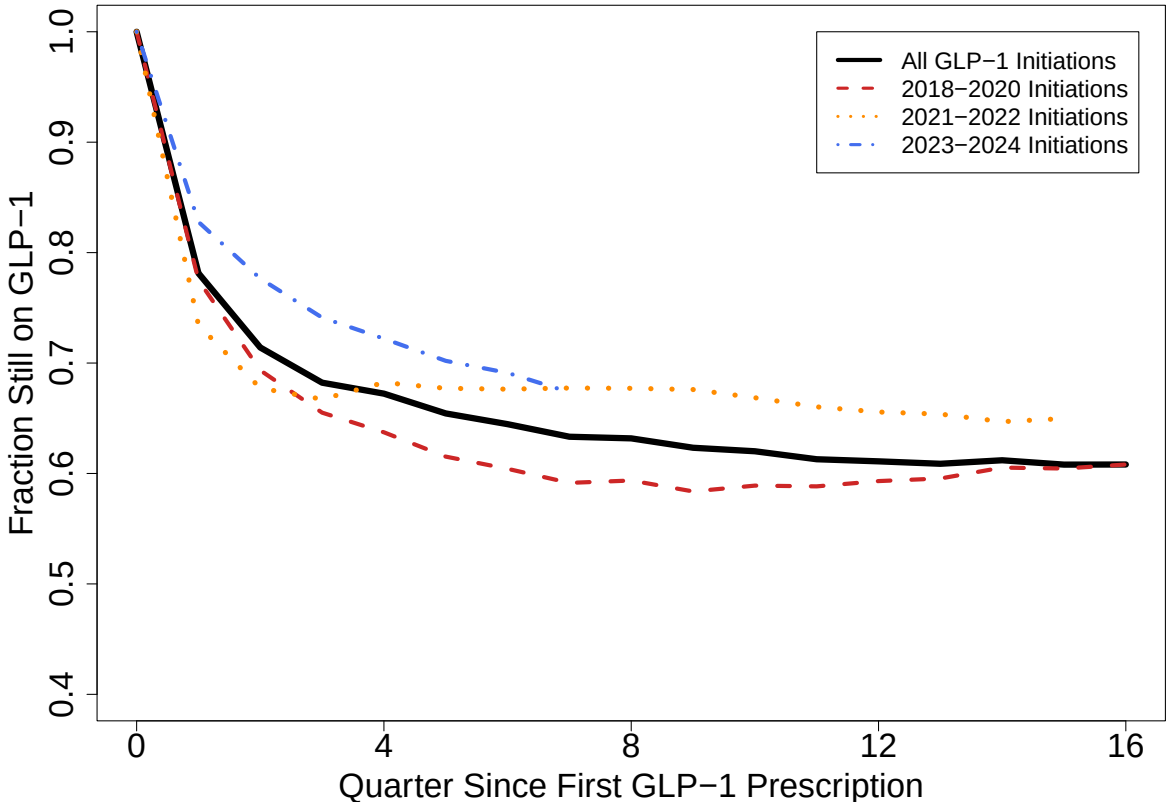
Notes: This figure shows the rise in GLP-1 among the entire VA population. Other GLP-1s include albiglutide, exenatide, lixisenatide, and tirzepatide. The dashed vertical line corresponds to November 2020, when semaglutide (Ozempic) was added to the VA national formulary.

Figure A.2: Distribution of PCP GLP-1 Propensity and First Stage



Notes: This figure presents a histogram of the leave-out PCP propensity to prescribe semaglutide to diabetic or obese patients in 2021–2022, as defined in Section 4. Overlaid on the histogram are first stage relationships between patients' 2018 PCP propensity and their cumulative GLP-1 use by the end of 2019 (gray) and by the end of 2022 (red). GLP-1 use rises between 2019 and 2020, and increases sharply among patients whose 2018 PCPs ultimately become high semaglutide adopters. The steeper slope in the 2022 cross-section corresponds to the kink in the event study estimates following the addition of semaglutide to the VA national formulary. The first stage relationships are local linear regressions estimated on the average of (residualized) $Z_{j(i)}$, over 50 equally spaced bins and weighted by the number of observations in each bin. 95% confidence bands are shown.

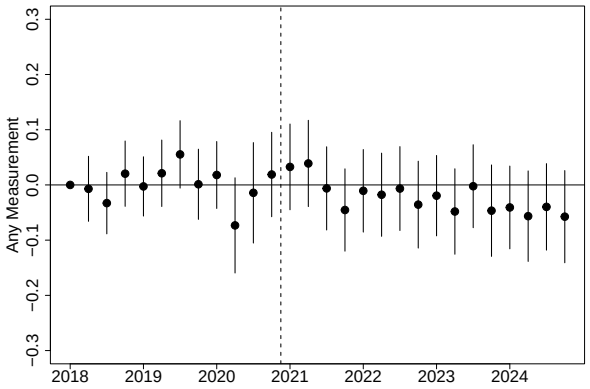
Figure A.3: GLP-1 Adherence Following Initiation



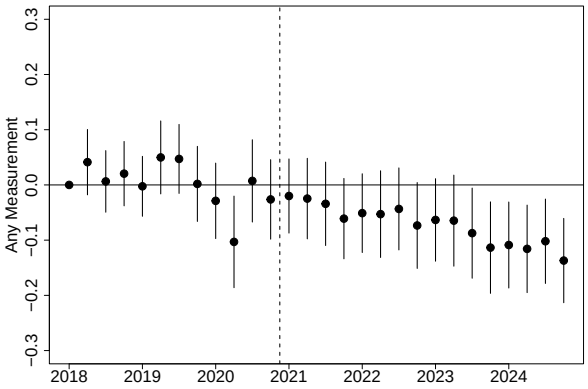
Notes: This figure presents the empirical surival function of GLP-1 use following first GLP-1 initiation for our baseline sample. The solid black line corresponds to all GLP-1 intiations (over all study years) and the other colored lines correspond to first initiation in specific years.

Figure A.4: Selection Into Measurement

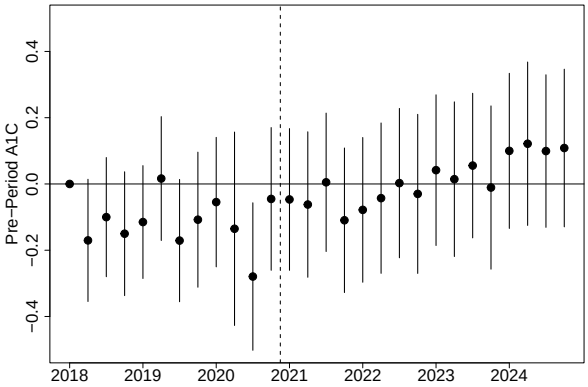
(a) Any Weight Measurement



(b) Any A1C Glucose Test



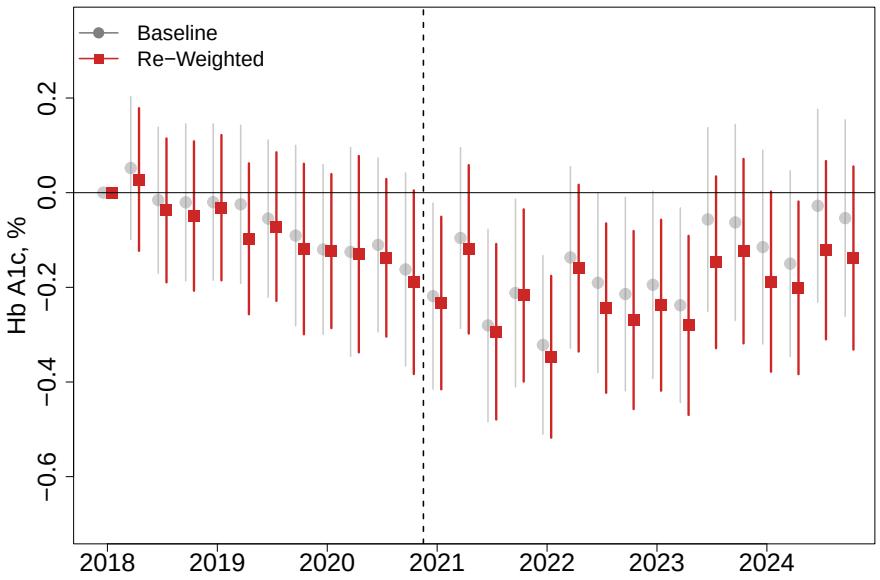
(c) Pre-2018 A1C Among Those Receiving A1C



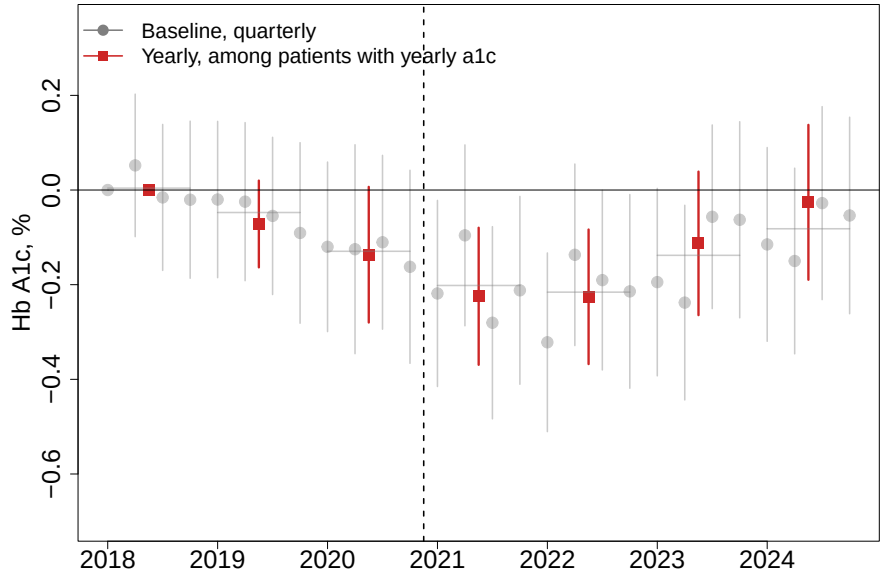
Notes: This figure shows the reduced form impact of PCP propensity to adopt GLP-1s on whether the patient has any weight measurement (panel a) or A1C glucose test (panel b) in a given year-quarter (Equation (1)). Means in the base period (2018 Q1) are 0.522 for any weight measurement and 0.324 for any A1C glucose measurement. Patient-year-quarter observations where the patient is dead are dropped. Panel c displays pre-period (2017) median A1C among patients that receive an A1C glucose test in a given year-quarter, conditional on receiving an A1C glucose test. This regression controls for PCP fixed effects instead of patient fixed effects since pre-period median A1C is constant within a patient. It shows that A1C glucose tests are more likely to be taken for ex ante more severe diabetes patients. Heteroskedastic-robust standard errors are clustered at the PCP-level, and 2018 Q1 is the omitted base period. The dashed vertical line corresponds to November 2020, when semaglutide (Ozempic) was added to the VA national formulary.

Figure A.5: Robustness: Accounting For Selection Into A1C Glucose Measurement

(a) Patient-Year-Quarter Re-Weighted to Match Pre-2018 A1C Distribution

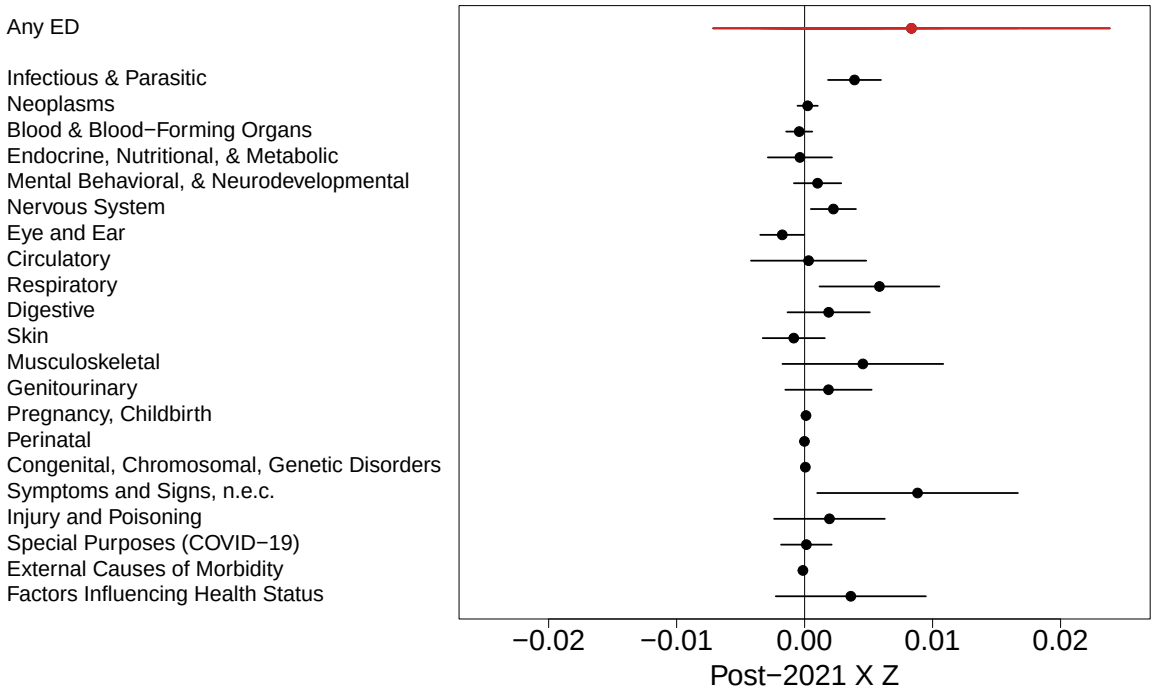


(b) Yearly A1C Among Patients with A1C Every Year



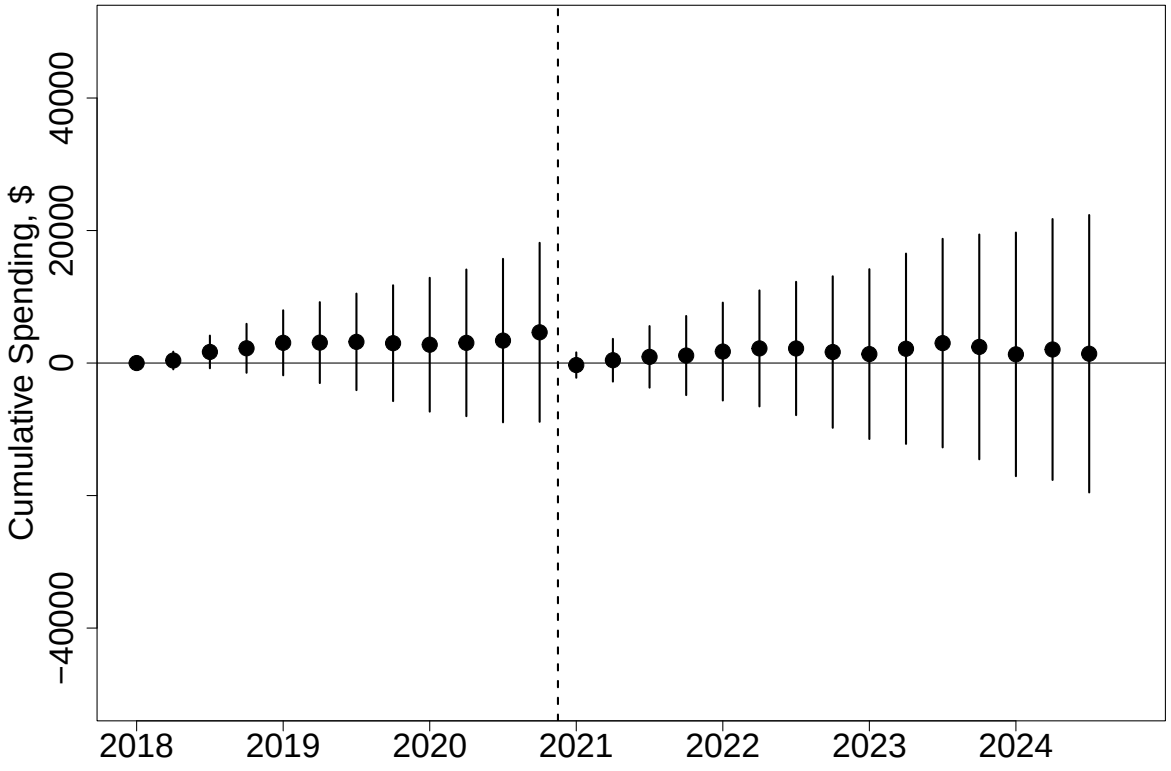
Notes: This figure displays two robustness exercises for A1C outcome. Following from Figure A.4 panel c, where A1C observations are increasingly taken from patients with ex ante higher A1C levels, panel a, re-weights A1C measurements in each year-quarter to match the ex ante pre-2018 distribution. Panel b collapses quarterly observations to yearly observations and estimates the event study on a sample of patients that have at least one A1C measurement per calendar year that they are alive (excluding pandemic year 2020). The sample size is 662,245 patients; among diabetic patients, 60% have annual A1C glucose measurements. The yearly event study coefficients among this sample is very similar to the quarterly means among the full sample (yearly averages displayed in grey horizontal lines). Heteroskedastic-robust standard errors are clustered at the PCP-level, and 2018 Q1 is the omitted base period. The dashed vertical line corresponds to November 2020, when semaglutide (Ozempic) was added to the VA national formulary.

Figure A.6: Emergency Department Visits by Major Diagnostic Category



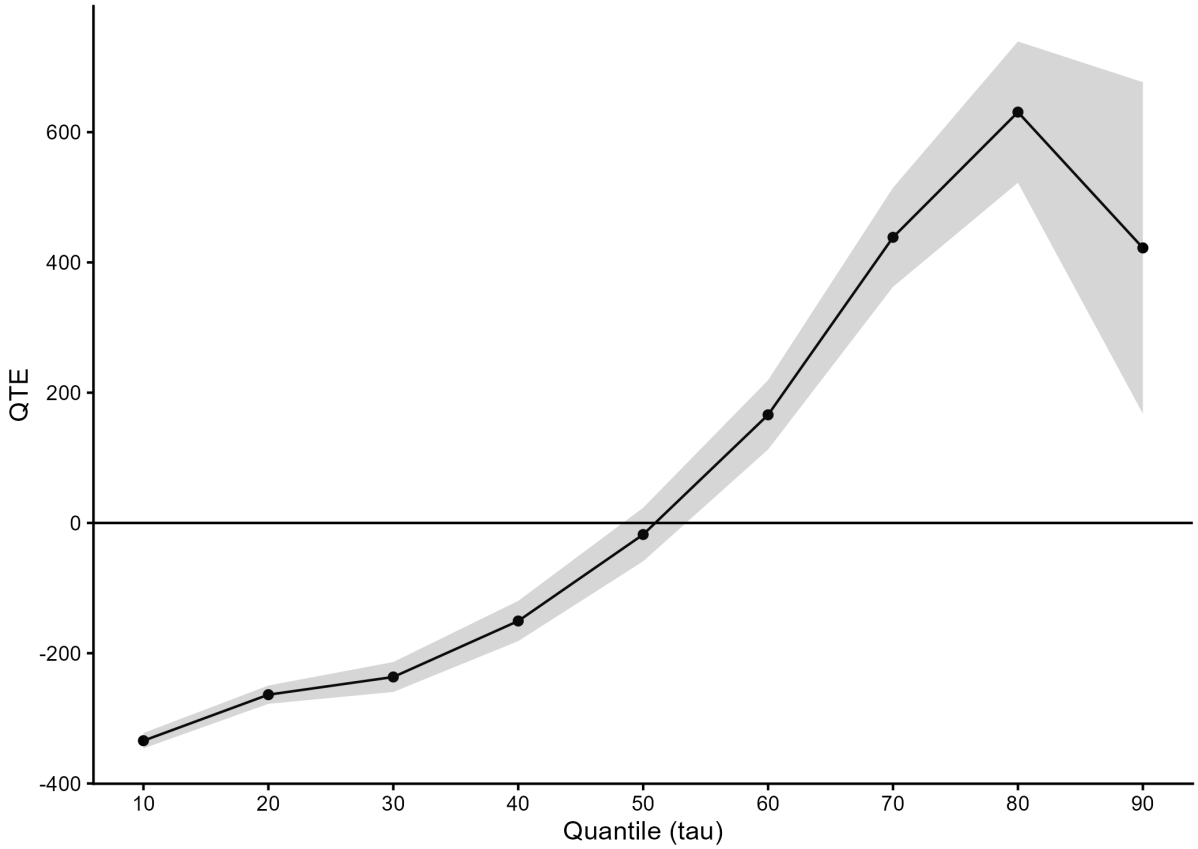
Notes: This figure displays the aggregate post-2021 vs pre-2021 coefficient on any emergency department (ED) visit and each major diagnostic category. Specifically, β from the following specification: $Y_{it} = \beta Z_{j(i)} \times \mathbb{1}\{t >= 2021\} + \alpha_i + \lambda_t + \theta_{s(i),t} + \varepsilon_{it}$. Heteroskedastic-robust standard errors are clustered at the PCP-level. N.e.c. stands for not elsewhere classified. The outcome variables are indicator for ED visits, thus the individual major diagnostic categories do not aggregate (average) to the "Any ED" outcome.

Figure A.7: Reduced Form Impacts on Cumulative Utilization



Notes: This figure displays shows the reduced form impact of PCP propensity to adopt GLP-1s on cumulative spending. The outcome variable is a cumulative version of Figure 5 panel d (contemporaneous spending): cumulative since 2018 until 2020, and cumulative since 2020 after 2020. Mean cumulative spending between 2021 and 2024 is $67,779. The specification is estimated on a sample of patient-year-quarter observations when the patient is alive. Heteroskedastic-robust standard errors are clustered at the PCP-level, and 2018 Q1 is the omitted base period. The dashed vertical line corresponds to November 2020, when semaglutide (Ozempic) was added to the VA national formulary.

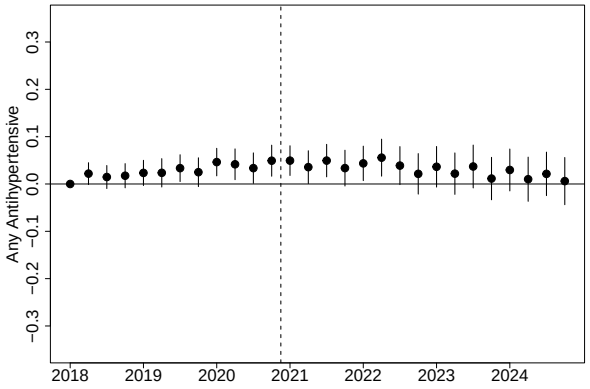
Figure A.8: Quantile Treatment Effect of High Propensity PCP on Total non-GLP-1 Medical Spending



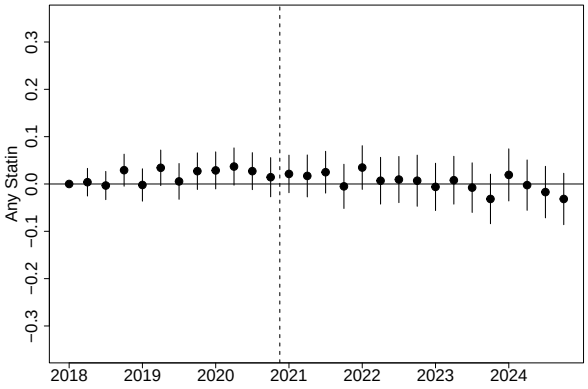
Notes: Figure reports treatment effect of having a high PCP for each decile. Treatment is defined as having a provider with propensity higher than the 80th percentile while control is defined as having a provider with propensity below the 20th percentile. See Appendix E for a description of how we implemented the Callaway and Li (2019) method. Standard errors are based on 100 empirical bootstrap iterations.

Figure A.9: Exclusion: Other PCP Practice Behaviors

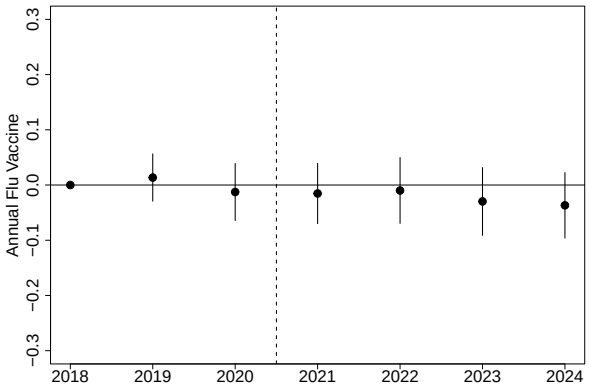
(a) Antihypertensives



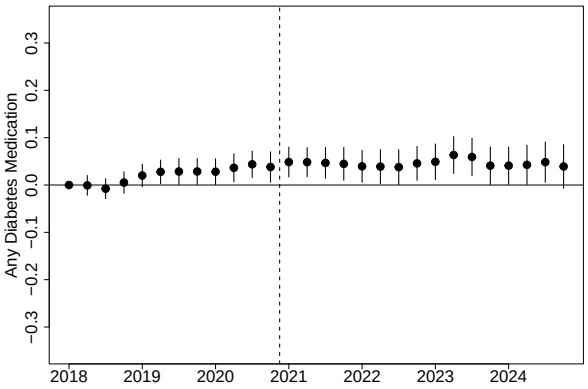
(b) Statins



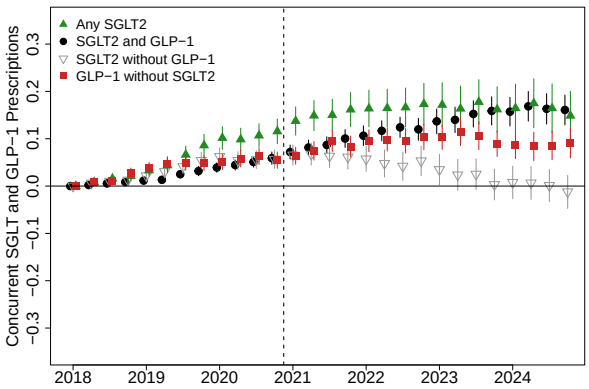
(c) Preventive Care: Flu Vaccines



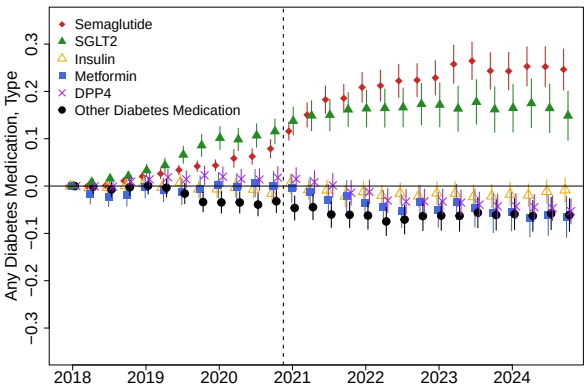
(d) Any Diabetes Medication



(e) SGLT2 Inhibitors



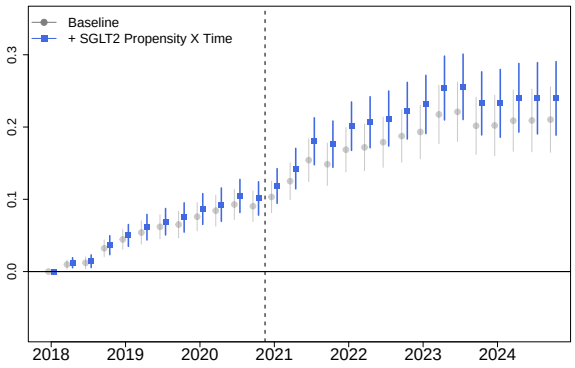
(f) Specific Diabetes Drugs



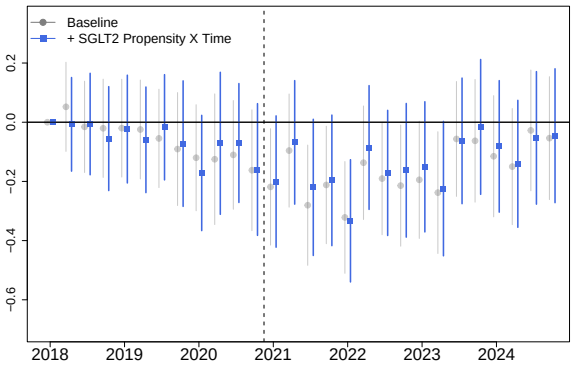
Notes: This figure shows the reduced form impact of PCP propensity to adopt GLP-1s on indicators for relevant medication drug use estimated from Equation (1). Panels (a) to (e) display any antihypertensives, and statins, annual seasonal flu vaccinations, any diabetes medication, and specific diabetes drugs. Panel (f) breaks down SGLT2 and GLP-1 use; see text for more details. Means in the base period (2018 Q1) are for antihypertensives: 0.554, statins: 0.408, flu vaccines: 0.559, any diabetes medication: 0.351, semaglutide: 0, metformin: 0.227, insulin: 0.140, SGLT2: 0.007: DPP-4: 0.028, and other diabetes medications: 0.118. Other diabetes medication includes sulfonylureas, meglitinides, thiazolidinediones, and alpha-glucosidase inhibitors. Patient-year-quarter observations where the patient is dead are dropped. Heteroskedastic-robust standard errors are clustered at the PCP-level, and 2018 Q1 is the omitted base period. The dashed vertical line corresponds to November 2020, when semaglutide (Ozempic) was added to the VA national formulary.

Figure A.10: Main Outcomes, Controlling for SGLT2 Adoption Propensity

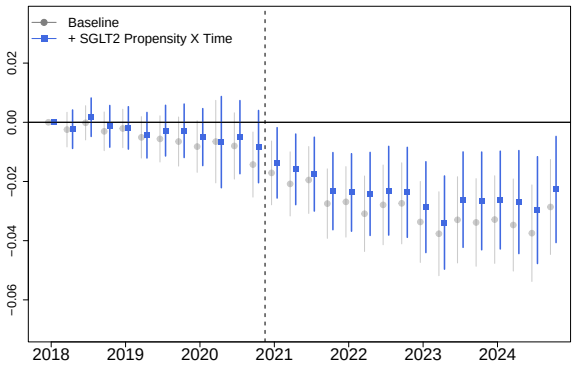
(a) GLP-1



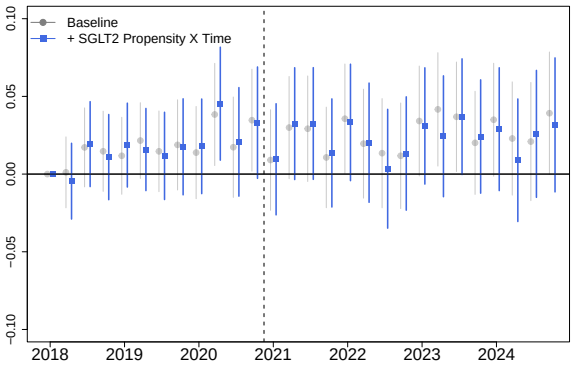
(b) A1C



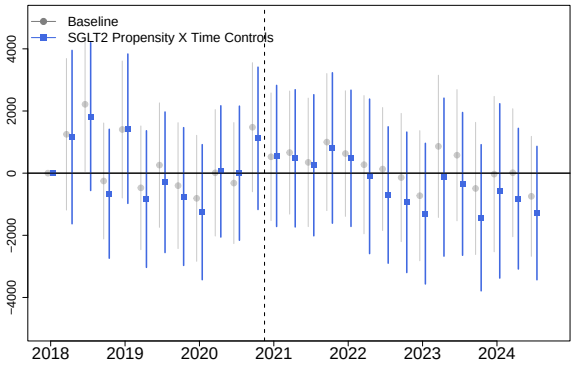
(c) Log Weight



(d) Any ED Visit



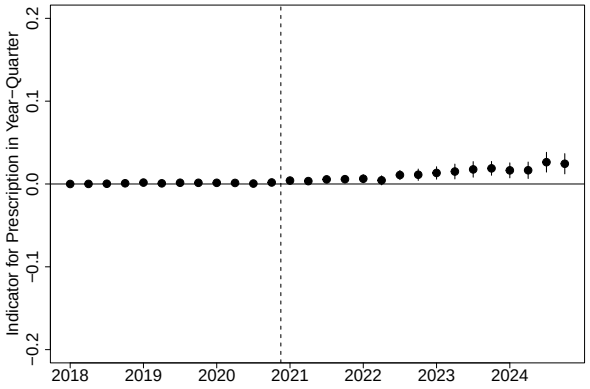
(e) Total Medical Spending



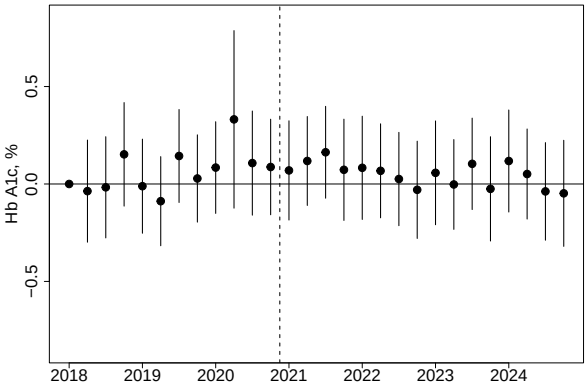
Notes: This figure displays robustness of the main outcomes after controlling for PCP SGLT2 inhibitor propensity interacted with year-quarter fixed effects. In Figure A.9, we see that high GLP-1 propensity PCPs are more likely to prescribe their patients SGLT2. SGLT2 propensity is constructed as the leave-out fraction of pre-2018 diabetic or obese patients who received SGLT2 in 2021–2022, $Z_{j(i)}^{SGLT2}$. The gray points show the baseline estimates from Equation (1) and the blue points show β_k^{GLP1} estimated from: $Y_{it} = \sum_{k=2018}^{2025} \beta_k^{GLP1} Z_{j(i)} \times \mathbf{1}\{t = k\} + \sum_{k=2018}^{2025} \beta_k^{SGLT2} Z_{j(i)}^{SGLT2} \times \mathbf{1}\{t = k\} + \alpha_i + \lambda_t + \theta_{s(i),t} + \varepsilon_{it}$. The correlation between PCPs' GLP-1 propensity and SGLT2 propensity is 0.10. Heteroskedastic-robust standard errors are clustered at the PCP-level, and 2018 Q1 is the omitted base period. The dashed vertical line corresponds to November 2020, when semaglutide (Ozempic) was added to the VA national formulary.

Figure A.11: Placebo Sample: Non-Diabetic, Non-Obese Patients Under Age 40

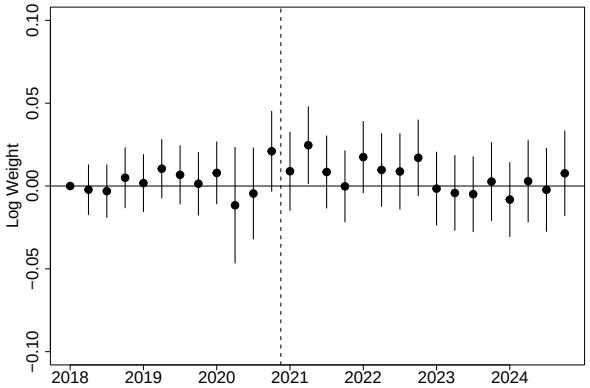
(a) GLP-1



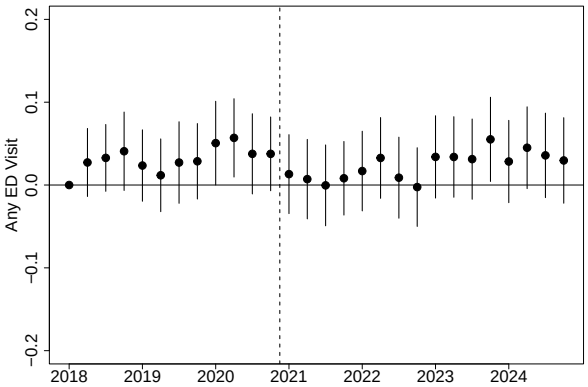
(b) A1C



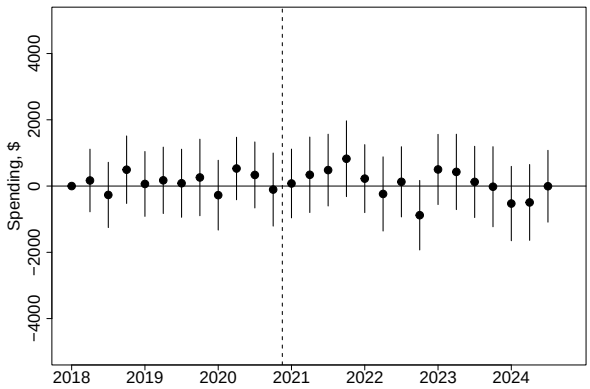
(c) Log Weight



(d) Any ED Visit

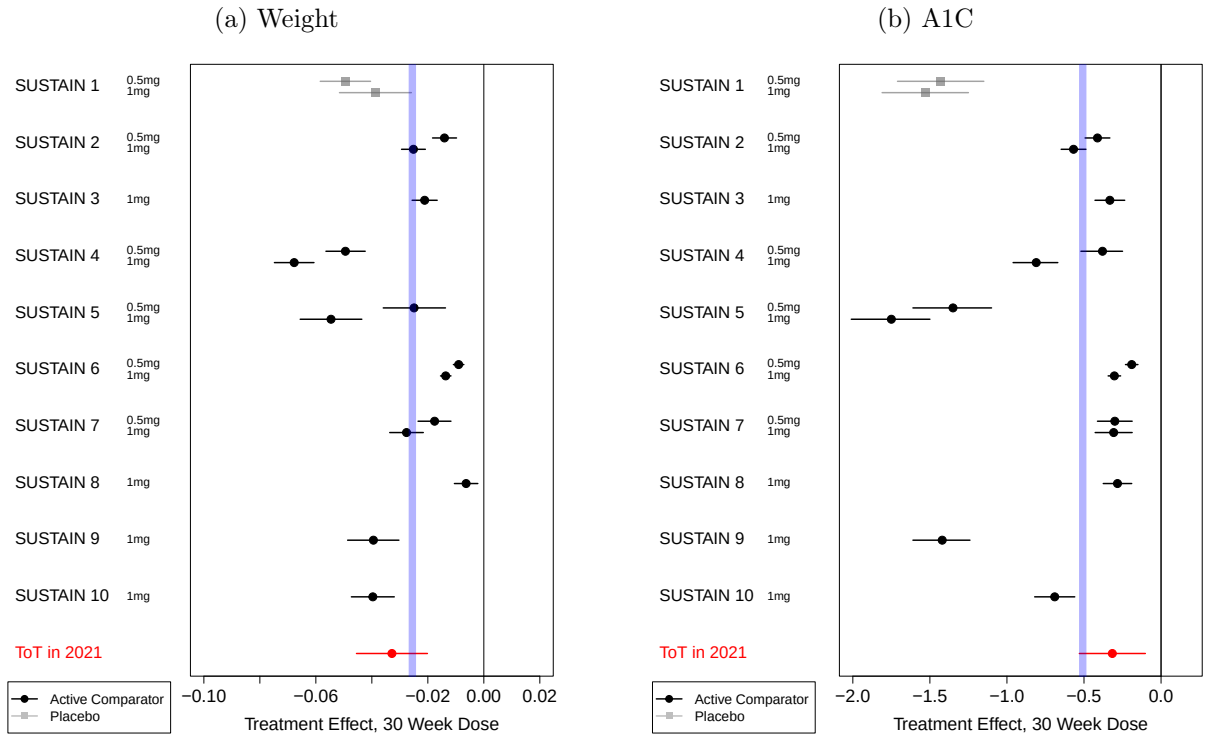


(e) Total Medical Spending



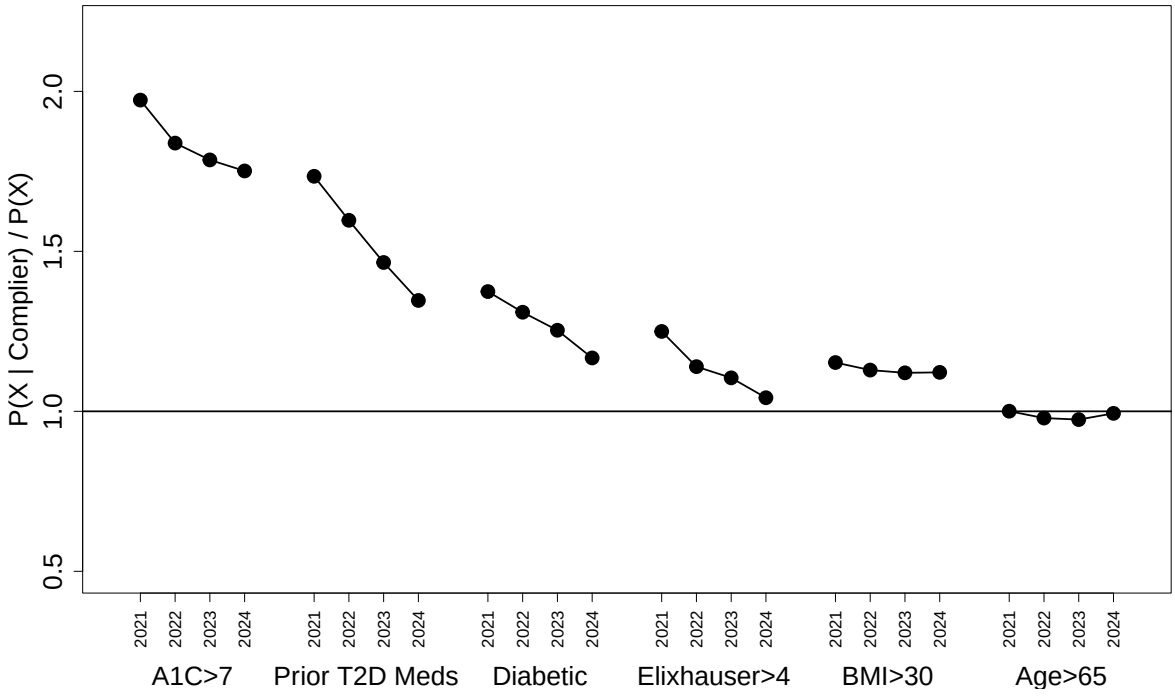
Notes: This figure shows the reduced form impact of PCP propensity to adopt GLP-1s on the main set of outcomes estimated from Equation (1) for a placebo falsification sample of non-diabetic and non-obese patients (prior to 2018) and were under age 40 in 2018. These patients are unlikely candidates for GLP-1. There is no overlap between these placebo patients and our baseline sample. There are no restriction on diabetes or obesity diagnoses after 2018. Spending is measured in 2025 dollars. Means in the base period (2018 Q1) are 0.000005 for any GLP1, 5.29 for median A1C, 5.24 for log weight (189 lbs), 0.075 for any ED visit, and $1,557 for total medical spending. Patient-year-quarter observations where the patient is dead are dropped. Heteroskedastic-robust standard errors are clustered at the PCP-level, and 2018 Q1 is the omitted base period. The dashed vertical line corresponds to November 2020, when semaglutide (Ozempic) was added to the VA national formulary.

Figure A.12: Comparing Quasi-Experimental Treatment on the Treated Effects to Clinical Trials



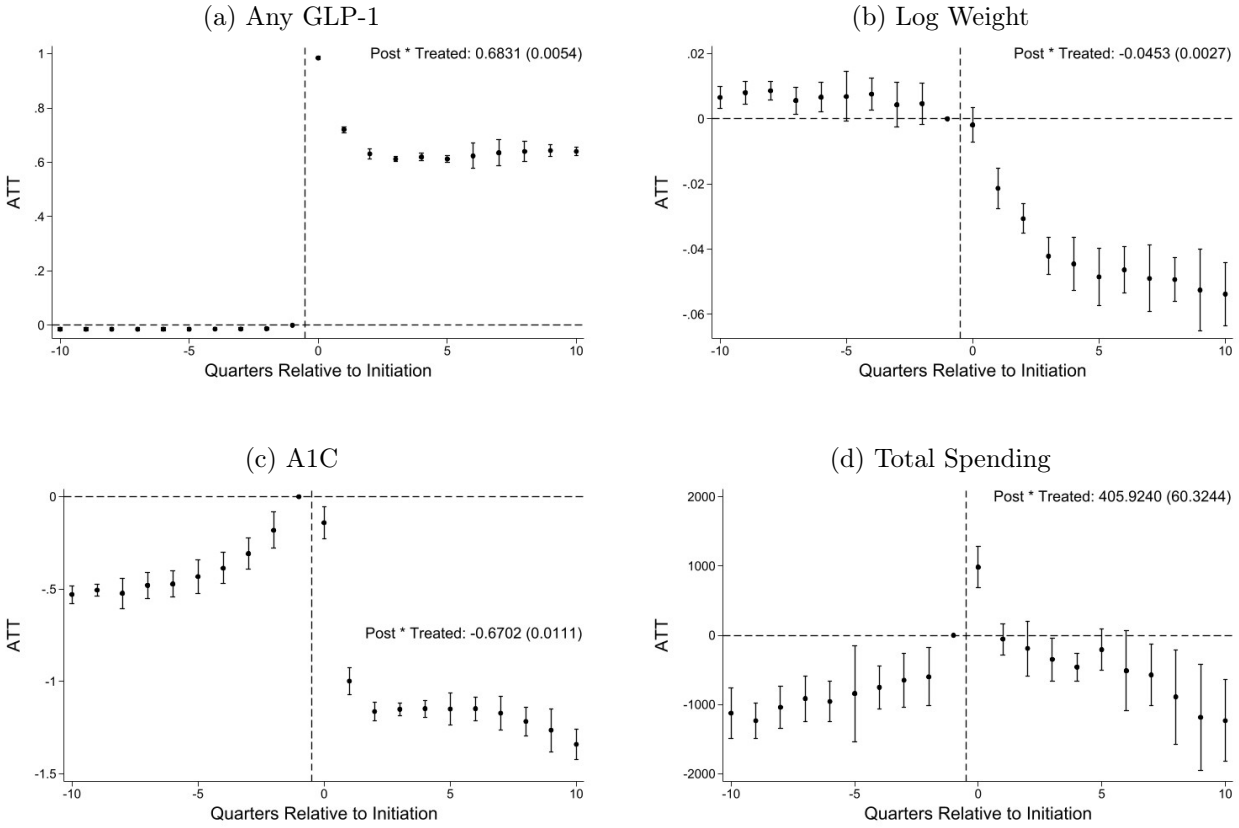
Notes: This figure compares our treatment effect on the treated (ToT) against RCT estimates from the SUSTAIN (Semaglutide Unabated Sustainability in Treatment of Type 2 Diabetes) trials on key clinical oucomes weight and A1C. Our quasi-experimental ToT estimates are for a 30-week dose of GLP-1 on outcomes in 2021. For ToT effects for other years, see Figure 6. Other than SUSTAIN 1, where the comparison group was the placebo, all other trials were against active comparators (e.g., vs insulin or GLP-1 + insulin vs only insulin, etc.). The vertical blue line corresponds to the treatment-sample-size-weighted average among the 9 comparator SUSTAIN trials. A summary of each of the SUSTAIN trials can be found in Table A.2. Estimates for the SUSTAIN trials are scaled to 30-weeks to match our ToT estimates; trial lengths for SUSTAIN 1 to 10 were 30, 56, 56, 30, 30, 104, 40, 52, 30, 30 weeks, respectively. Scaling of both the trial estimates and our dose response ToT estimates require an assumption that the treatment effect is linear in dosage. Standard errors (and 95% confidence intervals) around our ToT estimates are constructed via clustered bootstrap with 500 iterations.

Figure A.13: Complier Analysis by Year



Notes: This figure reports a complier analysis of baseline characteristics X for patients utilizing GLP-1s in a given year (2021–2024) is induced by their 2018 PCP's GLP-1 prescribing propensity. For each year, we plot the relative likelihood of characteristic X among compliers—patients whose treatment status changes with the instrument—compared to its unconditional prevalence in the population. All characteristics X are measured at baseline, prior to 2018. See Appendix D for more details of the complier analysis.

Figure A.14: Stacked Difference-in-Difference ATT



Notes: This figure plots average treatment on the treated estimates from a stacked difference in difference design described in Appendix F. Treated patients are those who initiated GLP-1 between Q4-2020 and Q4-2021 while control patients are those who initiated GLP-1 3 years later. For each period relative to the quarter of GLP-1 initiation, we plot the coefficient representing the average treatment effect on the treated. Standard errors are clustered at the initiation cohort level. See Appendix F for more details.

B Details on the Institutional Setting

In Section 2, we described the institutional setting. Ozempic is part of the VA National Formulary (VANF) and as such must generally be available for prescription at all VA facilities. By contrast, Wegovy is non-formulary, and is therefore subject to prior approval. Irrespective of inclusion in the VANF, the VA provides criteria for use (CFU) documents to assist providers in their clinical decision-making. The CFU for Ozempic and Wegovy are described below.

Ozempic The most recent CFU from March 2025 outlines the inclusion and exclusion criteria for prescribing Ozempic to Veterans (VA PBM, 2025). The exact criteria are as follows. Patients are in essence excluded from consideration if they 1) have a personal or family history of medullary thyroid carcinoma or with multiple endocrine neoplasia syndrome type 2, 2) have severe gastrointestinal dysmotility including gastroparesis, 3) have a history of pancreatitis, 4) are pregnant, 5) have a known proliferative diabetic retinopathy, severe nonproliferative diabetic retinopathy, clinically significant macular edema, or diabetic macular edema, 6) low body weight and/or presence of sarcopenia, or 7) are concomitantly using dipeptidyl peptidase 4 (DPP-4) inhibitors (e.g., sitagliptin, saxagliptin, linagliptin, etc.). Patients are included for consideration if they have type 2 diabetes and at least one of the following criteria are met: 1) the patient has established atherosclerotic cardiovascular disease (ASCVD) and/or chronic kidney disease (CKD) and is not a good candidate for empagliflozin, or 2) the patient's glycemic goals are not met with empagliflozin plus maximally tolerated doses of metformin, unless unable to use. We were able to locate an earlier version of the CFU for Ozempic, from 2022, using the Wayback Machine (VA PBM, 2022). The 2025 CFU is nearly identical to the 2022 CFU.

Wegovy The most recent CFU from August 2025 outlines the following inclusion and exclusion criteria for prescribing Wegovy to Veterans (VA PBM, 2025). Patients are in essence excluded from consideration if they are 1) pregnant, 2) lactating, 3) have type 1 diabetes, 4) have a personal or family history of medullary thyroid carcinoma or with multiple endocrine neoplasia syndrome type 2, 5) have severe gastrointestinal dysmotility including gastroparesis, 6) have a history of pancreatitis, or 7) have a known proliferative diabetic retinopathy, severe nonproliferative diabetic retinopathy, clinically significant macular edema, or diabetic macular edema. Patients are included for consideration if they have 1) documented participation in a comprehensive lifestyle intervention (CLI) that targets all three aspects of weight management: diet, physical activity, behavioral changes, and 2) a BMI greater than or equal to 30 kg/m^2 with at least one weight-related comorbidity.

C Details on Variable Definitions

In Section 3, we provided a brief description of the main outcome variables. Below, we describe the variables in more detail.

GLP-1 Prescriptions and Other Medications We construct patient-year-quarter indicators for any GLP-1 use from outpatient prescriptions, clinically administered initiation (e.g., hospital setting or supervised dose for educational purposes), and special authorization order requests. In addition to any use, we also measure intensive margin days supply, as well as total cumulative measures of ever use and acumulated days supply. We also construct a measure of any diabetes medication use and categorize individual diabetes drug classes : biguanides (metformin), insulin, SGLT2 inhibitors, DPP-4 inhibitors, sulfonylurea, semaglutide, and other non-semaglutide GLP-1s (e.g., tirzepatide, liraglutide, etc.), and others (thiazolidinediones, alpha-glucosidase inhibitors, etc.). To characterize the prescribing patterns of high semaglutide adopting PCPs, we also construct indicators for medications for statins and antihypertensives (e.g., ACE inhibitors, beta blockers, etc.).

Hemoglobin A1C Blood Glucose Levels We construct median A1C for each patient-year-quarter observation, measured in percent of hemoglobin in red blood cells that has glucose attached to it. In addition, we construct an indicator for median A1C below 7% (the typical target goal for T2D patients; American Diabetes Association, 2025).

Weight We take the logarithm of median weight for each patient-year-quarter, measured in pounds. We also report BMI as a secondary outcome, using a fixed, time-invariant median height for each patient based on measurements from 2015–2019. We use a time-invariant measure of height because unlike weight, height is unlikely to meaningfully change over time and changes are more likely to reflect measurement error.

Mental Health and Substance Use Outcomes We construct year–quarter indicators for overdose poisonings (ICD10 codes: T36-65) and alcohol intoxication or poisoning (ICD 10 codes: F10.12, F10.22, F10.92, T51) documented in VA and non-VA EDs. Because these are severe but relatively rare events, we complement these outcomes with measures of acute mental health derived from the Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001), which assesses depressive symptoms and psychological distress over the preceding two weeks. The PHQ-9 is a nine-item clinical screening tool that ranging from 0 to 27, with 27 being the most severe. Scores greater or equal to 5 potentially indicate mild symptoms, 10 indicating moderate, 15 indicating moderately severe, and 20 indicating severe.

From the Alcohol Use Disorders Identification Test (AUDIT-C; Bush et al., 1998), we construct average raw scores at the patient-year-quarter-level. The AUDIT-C is a 3-item questionnaire that asks about the patient's frequency and intensity of alcohol consumption over the past year. Scores

range from 0 to 12 with scores greater or equal to 4 potentially indicating moderate risk, 6 indicating high risk, and 8 indicating severe risk.

Healthcare Utilization and Spending We construct an outcome for any emergency department (ED) visit (VA and non-VA) in a given year-quarter and categorize ED visits into major diagnostic categories based on the primary diagnosis code. We also analyze healthcare spending via an average cost measure, based on Medicare relative value units (RVUs) and VA utilization data (VA Health Economics Resource Center, 2025; Wagner et al., 2003), which is available until 2024Q3. For non-VA spending, we use total charges from non-VA reimbursed claims, which represents the amount charged to and paid by the VA. We sum VA and non-VA spending to arrive at a total spending which can be disaggregated into outpatient and inpatient spending. Inpatient and outpatient spending are separately winsorized at the 99th percentile, conditional on positive spending each quarter ($327,336 and $77,999, respectively).

Non-VA community outpatient and inpatient utilization codes are classified via place of service codes, with POS=21, 31, 51, 61 as inpatient, and the remaining being outpatient. Note that the interpretation of total VA + non-VA spending is the cost incurred to the VA since VA average cost is meant to approximate VA operational cost and non-VA spending is measured via total charges.

In Table A.1 we also measure major adverse cardiovascular events (myocardial infarction and stroke) following Marso et al. (2016) using the following ICD-10 codes: I21, I22, I60, I61, I63, and I64.

D Complier Analysis

Following Abadie (2003), we non-parametrically estimate the characteristics of compliers relative to the baseline population (patients diagnosed with diabetes or obesity prior to 2018; see Section 3 for specifics), and we track how these complier characteristics change over time.

Specifically, for each year t and binary characteristic X, we estimate the probability of X given the patient is a complier divided by the probability of X in the baseline sample. This relative likelihood, $\gamma_{X,t}$, can be estimated using the following equivalence:

$$\gamma_{X,t} = \frac{\mathbb{P}(X|D_{t,1} > D_{t,0})}{\mathbb{P}(X)} = \frac{\mathbb{P}(D_t|Z = z_{high}, X = 1) - \mathbb{P}(D_t|Z = z_{low}, X = 1)}{\mathbb{P}(D_t|Z = z_{high}) - \mathbb{P}(D_t|Z = z_{low})}$$

where D_t is a binary indicator for whether the patient had GLP-1 treatment in year t. Therefore, our estimates are the ratio of two values: the probability of X among compliers and the probability of X among the all patients in our sample.

We estimate the numerator using a cross-sectional first degree local linear regression of D_t on residualized Z where we restrict to observations with $X = 1$ (where Z is residualized on station fixed effects). We then subtract the estimated conditional mean at the 1st percentile from the estimated conditional mean at the 99th percentile of residualized Z. The procedure for calculating the denominator is the same except we keep all observations regardless of their value of X. Importantly, the percentile values z_{high} and z_{low} are unchanged across years and samples.

Figure A.13 plots $\gamma_{X,t}$ for years 2021 through 2024 and for six health characteristics measured in 2018. In earlier periods, there is a higher incidence of adverse health conditions among compliers compared to the population. However, over time the population of compliers becomes healthier compared to the whole sample.

E Quantile Regressions

To conduct quantile regressions, we employ Callaway and Li (2019) which provides a method to calculate the quantile treatment effect of having a high propensity PCP in a difference-in-differences context. This method requires three periods of data where two periods are prior to treatment. We divide our sample into periods defined by Q1-2018 through Q2-2019, Q3-2019 through Q4-2020, and Q1-2021 through Q3-2024. We then calculate each patient's average quarterly spending in each period using only quarters where the patient was alive. Therefore, our outcome Y_{it} is the average quarterly medical spending for periods $t = 1, 2, 3$. We restrict this analysis to binary treatment. Thus, rather than include a continuous Z, we define a patient as treated ($D = 1$) if their provider's propensity is larger than the within-station 80th percentile and untreated ($D = 0$) if their provider propensity is below the within-station 20th percentile. We then only keep the observations with $D = 0$ or 1 (i.e., throw out patients with the middle 60% of Z). For computational reasons, we sample 50% of patients from this subsample, so our sample in this analysis is 20% of the full sample.

Two assumptions are required for this method. The first, the distributional difference in difference assumption, is analogous to the parallel trends assumption in difference in difference. Define Y_{0t} and Y_{1t} be the untreated and treated outcomes in period t, respectively. Then $\Delta Y_{0t} = Y_{0t} - Y_{0t-1}$ is the difference in untreated potential outcomes between period t and $t-1$. The distributional difference in difference assumption states:

$$\Delta Y_{0t} \perp D$$

In words, the trend for patients untreated potential outcomes is independent of whether the patient is actually treated. It is reasonable if you assume that the provider did not materially change their style of care when semaglutide was added to the formulary. It is analogous to the parallel trends assumption in difference in difference; we need this assumption to replace the unobserved trend for the treated group with the observed trend for the untreated group *at each quantile*.

The second is the Copula Stability Assumption. A copula function, $C_{X,Y}(\cdot,\cdot)$, captures the dependence between two marginal distributions such that $C_{X,Y}(F_x(\delta), F_y(\gamma) = F_{X,Y}(\delta,\gamma)$. Formally, the Copula Stability Assumption states:

$$C_{\Delta Y_{0t}, Y_{0t-1}|D=1}(\cdot,\cdot) = C_{\Delta Y_{0t-1}, Y_{0t-2}|D=1}(\cdot,\cdot)$$

Informally, this requires the dependence between $\Delta Y_{0t}|D = 1$ and $Y_{0t-1}|D = 1$ to be the same as the dependence between $\Delta Y_{0t-1}|D = 1$ and $Y_{0t-2}|D = 1$. This assumption states that if, between periods $t = 1, 2$, the largest increases in spending tend to go to those initially at the top end of spending in $t = 1$, then between periods $t = 2, 3$, the largest increases in spending will tend to go to those who start out at the top of the distribution in $t = 2$. This is plausible if the VA did not change other institutional features related to spending other then adding semaglutide to the formulary. Extrapolating the observed historical dependence between these untreated marginal distributions allows us to identify the unobserved joint distribution $F_{\Delta Y_{0t}, Y_{0t}|D=1}$.

We then use Callaway and Li (2019) to calculate the quantile treatment effects of having a high propensity PCP where untreated outcomes take the following form:

$$Y_{0it} = \theta_t + \eta_i + \nu_{it}$$

where θ_t is a time period trend, η_i is individual heterogeneity (which may be distributed differently for treated and untreated individuals), and ν_{it} is a time varying unobservable term. Under this baseline model of untreated patients, we calculate the quantile treatment effect for each decile of spending. The results are then plotted in figure Figure A.8 with standard errors based on 100 bootstrap iterations.

F Difference-in-Differences: Early- vs Later-Treated

While our treatment-on-the-treated (ToT) estimates recover the causal effect of GLP-1 exposure per unit of accumulated treatment, they do not correspond to a single initiation event and instead aggregate treatment histories across patients at different durations since adoption. To complement this approach, we implement a stacked difference-in-differences design (similar to Wing et al., 2025) around patients' first date of GLP-1 adoption.

Specifically, we define patients who initiated GLP-1 therapy between Q4-2020 and Q4-2021 as treated and use patients who initiated GLP-1 between Q4-2023 and Q4-2024 as our control. Since we track patients quarterly, this results in five initiation sub-experiments. For each sub-experiment a, we use patients who initiated GLP-1 in the analogous quarter three years later as control. For instance, sub-experiment a = Q2-2021 would have as its treated group patients who initiated GLP-1 in Q2-2021 and patients who initiated GLP-1 in Q2-2024 as its control. We combine these sub-experiments to yield an aggregated ATT parameter.

To do so, we move our data from calendar time to relative time. Let Y_{iae} represent patient i's outcome in sub-experiment a when they are e quarters away from initiation. Let $\gamma_{ia} = 1[A_i = a]$ be a dummy variable that equals 1 if the patient is *treated* in sub-experiment a and 0 otherwise. Then we run the following event study on this stacked dataset:

$$Y_{iae} = \gamma_{ia} + \sum_{h \neq -1} \beta_h \times \gamma_{ia} \times 1[e = h] + \tau_e + u_{iae} \tag{3}$$

where τ_e are relative time fixed effects. Under the assumption that treated and untreated patients would have moved in parallel absent semaglutide, the β_h parameters represent the aggregated ATT for being h periods away from initiation. In Figure A.14, we plot the β_h coefficients for four main outcomes: an indicator for if the patient is on GLP-1, the natural log of the patient's weight, the patient's A1C level, and quarterly spending on the patient. With each plot, we also include the two period difference-in-difference estimate where we interact treatment with the post period.

The event study exhibits modest pre-trends, suggesting that changes in patients' underlying health conditions contribute to the timing of GLP-1 initiation. In particular, treated patients experience rising A1C levels relative to untreated patients in the periods leading up to initiation, while spending spikes around the initiation quarter, consistent with endogenous treatment timing. We therefore interpret these estimates cautiously. Nonetheless, the stacked difference-in-differences results are qualitatively similar to our main findings in Figure 6 and Figure A.12. Initiation of GLP-1 therapy is associated with an approximate 4% reduction in weight and a 0.5 percentage point decline in A1C, alongside a modest increase in spending driven primarily by the quarter of initiation.