Performance Pay for Health Insurers: Evidence from Colombia's High-Cost Account*

Natalia Serna
Stanford University

Lin-Tung Tsai
UC Berkeley

September 10, 2026

Abstract

This paper shows that performance pay can dampen effort on rewarded tasks by exposing risk-averse agents to greater uncertainty in costs and in the returns to effort. We test this and other predictions using 13 years of individual-level health claims data from Colombia, exploiting the government's introduction of performance pay targeting HIV/AIDS patients in 2016 and cancer patients in 2019. Performance pay improved disease management for both conditions, increasing healthcare spending per patient by 25% and 32%, respectively. Heterogeneity analyses confirm our predictions: treatment effects are 50% larger for health conditions with lower cost variance, 50% larger among insurers with a higher ex-ante likelihood of receiving payments, and twice as large among insurers facing lower variance in that likelihood than their counterparts. A key lesson from the Colombian policy is that quality metrics need to target specific health conditions to concentrate insurer effort where moral hazard is most severe.

Keywords: Performance Pay, Health insurance, Moral hazard, Risk selection.

JEL codes: I11, I13, I18, L13.

*Serna: Department of Health Policy, Stanford University, nserna@stanford.edu. Tsai: Department of Economics, University of California–Berkeley, lintung_tsai@berkeley.edu. Natalia Serna acknowledges funding from the Stanford Discovery Innovation Fund. We thank Loren Baker, Maria Polyakova, Sherri Rose, and Maya Rossin-Slater for their valuable feedback.

1 Introduction

A central challenge in the design of contracts is inducing effort from agents whose actions are costly and imperfectly observable (Holmström, 1979).[1] Performance pay has emerged as a potential solution, tying compensation to measurable outcomes to align agent incentives more closely with principal objectives. Yet evidence on whether performance pay improves quality in healthcare markets is decidedly mixed (Burke et al., 2025; Gupta, 2021; Miller and Barbiarz, 2013), and particularly discouraging for health insurers (Fioretti and Wang, 2023; Markovitz et al., 2021; Layton and Ryan, 2015).

This paper asks why performance pay succeeds in some settings and fails in others, and shows that the answer lies in how the policy handles uncertainty. We develop a model in which a risk-averse insurer's optimal effort depends not only on the level of performance payments but on two sources of uncertainty: uncertainty in the costs of treating a health condition, and uncertainty in whether effort translates into a measurable, rewarded outcome. Both forms of uncertainty depress effort, even on the tasks the contract is designed to reward. This distortion is distinct from the standard multitasking concern that agents divert effort toward rewarded tasks at the expense of unrewarded ones (Tebaldi et al., 2024; Holmstrom and Milgrom, 1991). The model further shows that aggregating quality metrics across health conditions, as MA star ratings do, dilutes the link between an insurer's effort and its probability of being paid, whereas targeting metrics to specific, verifiable conditions strengthens it and increases the policy's effectiveness.

We test these predictions using Colombia's High-Cost Account (HCA), a performance pay mechanism introduced for HIV/AIDS patients in 2016 and cancer patients in 2019 (breast, cervical, colorectal, stomach, lung, leukemia, and non-Hodgkin lymphoma). Under this mechanism, insurers receive compensation for enrollees with targeted conditions only if they meet specified care quality standards.[2] Funds are redistributed from insurers with below-average disease prevalence to those with above-average prevalence who meet

[1] This problem is common in education markets (Barlevy and Neal, 2012; Goodman and Turner, 2010; Atkinson et al., 2009), managerial compensation (Griffith and Neely, 2009; Jensen and Murphy, 1990), and healthcare (Gupta, 2021; Miller and Barbiarz, 2013).

[2] The quality standards include thresholds for disease prevalence, disease incidence, viral load monitoring, medication adherence, etc.

the quality thresholds (Decree 2699 of 2007), so the mechanism is budget-neutral for the government.

Our data span 13 years of individual-level enrollment records and health claims from all individuals enrolled in Colombia's contributory health system—roughly 24 million people per year from 2011 to 2023.[3] We study the effects of the policy for HIV/AIDS and cancer separately given their distinct patterns of healthcare use. For both conditions we use a cohort event study design. The key observation motivating our cohort approach is that healthcare utilization is concentrated in the first three years following diagnosis. Comparing within-person trends in healthcare use would therefore confound the natural disease progression with the policy's effect. Instead, we compare cohorts of patients diagnosed with targeted diseases before and after the policy, and further compare them against cohorts diagnosed with untargeted diseases over the same period to net out contemporaneous trends unrelated to the policy.[4]

We find that performance pay for cancer increased healthcare spending per patient by 32.0% relative to the baseline of 23.1 million pesos ($7,800). The spending increase was concentrated in outpatient care services and consistent with improved disease management. Patients diagnosed with targeted cancers after the policy had 13.6% higher rates of chemotherapy, radiotherapy, and cancer-related surgeries relative to baseline, suggesting that, prior to the introduction of performance pay, some patients had gone without appropriate cancer treatment or faced significant delays in receiving it.

Heterogeneity analyses confirm the theoretical predictions. First, treatment effects on spending and care quality outcomes are 50% larger for health conditions with below-median variance in healthcare costs, suggesting that risk-averse insurers concentrate effort on more homogeneous diseases. Second, we construct a measure of each insurer's ex-

[3]The contributory system covers individuals who pay payroll taxes and their dependents. The low-income population unable to pay taxes are covered by the subsidized system, which is fully funded by the government. Insurance coverage in Colombia is near universal. We do not have health claims data for the subsidized system.

[4]These untargeted cohorts are specially important to net out the effects of the COVID-19 pandemic which was declared by the World Health Organization in March 2020. If targeted diseases were not disproportionately affected by the pandemic compared to untargeted ones, then the comparison allows use to identify the true policy effect. Because the HIV/AIDS policy was introduced before the pandemic, its estimated effects offer additional evidence on the impact of performance pay, independent of pandemic-related confounds.

ante likelihood of receiving performance payments, which is equal to the share of quality thresholds that insurers meet under the 2019 policy, computed using 2018 pre-policy claims data. Insurers with above-median likelihood of receiving payments exhibit 50% larger treatment effects, consistent with effort increasing in the ease of meeting the quality thresholds. Third, treatment effects are twice as large among insurers with below-median variance in their predicted payment likelihood as their counterparts, in line with effort decreasing in the uncertainty of being paid.

Our findings are qualitatively similar for the HIV/AIDS policy. Performance pay increased total healthcare spending among post-policy cohorts by 25.4% relative to the baseline of 12.3 million pesos ($4,200). As with cancer, this is consistent with improved disease management: the likelihood of claiming antiretroviral therapy (ART) medications rose by nearly 17.1%, and the likelihood of receiving vaccinations against bacterial and viral infections that are common complications of HIV/AIDS increased by 157.4%. Heterogeneity analyses confirm that these treatment effects are driven entirely by health conditions with below-median cost variance, consistent with our model's characterization of optimal effort and indicating that the results are not an artifact of any single policy.

Given that insurers in Colombia compete mainly on their network of covered providers, one way performance pay can affect health outcomes is by encouraging insurers to more actively manage which facilities they authorize care at, and to reduce wait times for medication dispensing, which occurs at insurer-owned facilities rather than at pharmacies. We examine this by comparing the characteristics of providers where patients receive care before and after the policy, focusing on the number of beds, and by using as an outcome an indicator for whether the patient received any disease-specific medication within the year of diagnosis. We find that patients with targeted cancers were more likely to receive care at larger, more specialized facilities after the policy, with no corresponding effect for HIV/AIDS patients, since disease management in that case depends more on timely medication dispensing than on facility-based care. Indeed, the likelihood of receiving disease-specific treatment (ART medications, in the case of HIV/AIDS) within the year of diagnosis was nearly twice as large as the corresponding effect across the full three-year post-diagnosis window, indicating that wait times for medication dispensing shortened.

Our paper contributes to the literature on performance pay in healthcare (Miller and Barbiarz, 2013; Eijkenaar, 2012), which has largely focused on provider-level contracts and found mixed results (e.g., Gupta et al., 2023; Einav et al., 2022; Gupta, 2021; Mullen et al., 2010). For instance, Burke et al. (2025) find that performance pay for skilled nursing facilities in the U.S. failed to improve targeted quality metrics such as 30-day readmission rates, while Gupta (2021) and Gupta et al. (2023) show that hospital performance pay under Medicare reduced readmissions and improved care coordination with downstream providers. In developing countries, performance pay has been studied in the context of HIV (De Walque et al., 2015), finding increased testing following its introduction, and in the context of maternal and infant care (Basinga et al., 2011), finding higher rates of preventive care visits for children but no effect on prenatal care visits. We argue that one reason for these opposing findings is uncertainty in returns to effort: if the link between effort and measurable outcomes is noisier in skilled nursing facilities than in hospitals, or in maternal care than in HIV management, then the resulting revenue uncertainty would dampen effort on rewarded tasks. More broadly, our results suggest that the effectiveness of performance pay depends on how precisely effort translates into measurable outcomes.

While most of the literature focuses on healthcare providers, a few exceptions study performance pay for insurers, including Fioretti and Wang (2023); Markovitz et al. (2021); Layton and Ryan (2015). These authors show that star quality ratings in Medicare Advantage, which determine risk-adjusted payments for health plans, have not improved plan quality, but have strengthened incentives for risk selection and, in some cases, spurred plan entry. Vatter (2025) states that one reason star ratings fail to improve outcomes is that scores are too granular: coarser ratings are easier for consumers to interpret, which promotes competition between plans and raises quality. Our paper complements this literature by showing that performance pay for insurers can improve patient outcomes when quality scores are disease-specific and not directly observable to enrollees, limiting the scope for gaming or selection.

The rest of this paper is structured as follows: section 2 presents our theoretical framework, section 3 introduces the Colombian health insurance system, section 4 describes our data, section 5 presents the empirical design, section 6 presents results for the cancer

policy and section 7 for the HIV/AIDS policy, section 8 explores the role of managed care, section 9 explains why the Colombian policy achieved its intended targets, and section 10 concludes.

2 Theoretical Framework

We begin by characterizing how performance pay influences insurer behavior across multiple dimensions of care quality. We propose a model of a health insurance market in which premiums are fixed and insurers choose a vector of efforts $\mathbf{e}_\theta = (e_{1\theta}, \ldots, e_{K\theta})$, where $e_{k\theta} \geq 0$ denotes effort along quality dimension k for consumer type $\theta \in \{0, 1, \ldots, I\} \equiv \Theta$, with $\theta = 0$ denoting healthy individuals. Each dimension k corresponds to a measurable quality indicator, such as ART prescription rates, vaccination rates, or early disease detection, that the regulator can contract on.

Demand and costs. Insurer demand from consumer type θ is $s_\theta(\mathbf{e}_\theta)$, which is increasing and concave in each effort dimension separately:

$$\frac{\partial s_\theta}{\partial e_{k\theta}} > 0, \quad \frac{\partial^2 s_\theta}{\partial e_{k\theta}^2} < 0, \quad \frac{\partial^2 s_\theta}{\partial e_{k\theta} \partial e_{k'\theta}} = 0 \quad \text{for } k \neq k'$$

The first condition captures the idea that consumers prefer insurers that invest more in each dimension of care quality. The concavity condition reflects diminishing returns to effort along each dimension. The zero cross-partial assumption means each quality dimension contributes independently to consumer demand, which keeps the model tractable while allowing for full heterogeneity in the type-specific derivatives $\partial s_\theta / \partial e_{k\theta}$.

The cost per enrollee of type θ is $C_\theta(\mathbf{e}_\theta) \sim \mathcal{N}(c_\theta(\mathbf{e}_\theta), \sigma_\theta^2)$, where the mean cost $c_\theta(\mathbf{e}_\theta)$ is increasing and convex in each effort dimension separately:

$$\frac{\partial c_\theta}{\partial e_{k\theta}} > 0, \quad \frac{\partial^2 c_\theta}{\partial e_{k\theta}^2} > 0, \quad \frac{\partial^2 c_\theta}{\partial e_{k\theta} \partial e_{k'\theta}} = 0 \quad \text{for } k \neq k'$$

Higher effort along any dimension raises costs (for instance, more intensive ART moni-

toring or more frequent vaccination outreach require greater expenditure) and these costs are convex, reflecting the increasing difficulty of improving quality along any given dimension. The derivatives $\partial c_\theta / \partial e_{k\theta}$ are consumer type-specific to capture the fact that improving a given quality indicator may be more resource-intensive for some patients than others.

Quality thresholds and performance payments. The regulator sets a vector of quality thresholds $\mathbf{I}^* = (I_1^*, \ldots, I_K^*)$ and a vector of weights $\mathbf{w} = (w_1, \ldots, w_K)$ such that $w_k \geq 0$ and $\sum_k w_k = 1$. The probability that the insurer meets threshold I_k^* for indicator k depends on its effort along that dimension:

$$\tau_{k\theta}(e_{k\theta}, I_k^*) = \Pr\Big(I_{kj}(e_{k\theta}) \geq I_k^*\Big)$$

where $\partial \tau_{k\theta} / \partial e_{k\theta} > 0$, $\partial^2 \tau_{k\theta} / \partial e_{k\theta}^2 < 0$, and $\partial \tau_{k\theta} / \partial I_k^* < 0$. The overall probability of receiving performance payments aggregates across indicators as follows:

$$\tau_\theta(\mathbf{e}_\theta, \mathbf{I}^*, \mathbf{w}) = \sum_k w_k \cdot \tau_{k\theta}(e_{k\theta}, I_k^*)$$

Performance payments are then $B_\theta \sim \mathcal{N}(p_\theta \tau_\theta(\mathbf{e}_\theta, \mathbf{I}^*, \mathbf{w}), \sigma_B^2)$, where p_θ is the per-capita payment conditional on meeting the quality thresholds.

Insurer's problem. The insurer profit function is $\Pi(\mathbf{e}_\theta) = -e^{-2\rho\pi(\mathbf{e}_\theta)}$. This profit function follows a Constant Absolute Risk Aversion (CARA) specification to reflect risk aversion. With performance payments and assuming separability across consumer types, the insurer's expected profit is:

$$\mathbb{E}[\Pi(\mathbf{e}_\theta, p_\theta)] = \sum_{\theta \in \Theta} \left[\Big(R_\theta - c_\theta(\mathbf{e}_\theta) + \tau_\theta(\mathbf{e}_\theta, \mathbf{I}^*, \mathbf{w})\, p_\theta \Big) s_\theta(\mathbf{e}_\theta) - \frac{\rho}{2} \Big(\sigma_\theta^2\, c_\theta(\mathbf{e}_\theta)^2\, s_\theta^2(\mathbf{e}_\theta) + \sigma_B^2\, s_\theta^2(\mathbf{e}_\theta) \Big) \right]$$

The insurer maximizes this profit over the full effort vector $\mathbf{e}_\theta$. Given the separability assumptions, the first-order condition for each effort dimension k is:

$$FOC(e_{k\theta}, p_\theta) = \left(R_\theta - c_\theta + \tau_\theta p_\theta\right) \frac{\partial s_\theta}{\partial e_{k\theta}} - \left(\frac{\partial c_\theta}{\partial e_{k\theta}} - \frac{\partial \tau_{k\theta}}{\partial e_{k\theta}} w_k p_\theta\right) s_\theta$$
$$- \rho \left(\sigma_\theta^2 \, c_\theta \, \frac{\partial c_\theta}{\partial e_{k\theta}} \, s_\theta^2 + \sigma_B^2 \, s_\theta \, \frac{\partial s_\theta}{\partial e_{k\theta}}\right) = 0 \tag{1}$$

The term $(\partial \tau_{k\theta}/\partial e_{k\theta})\, w_k \, p_\theta$ is the key object in this expression: it represents the marginal increase in the probability of clearing indicator k's threshold with respect to effort, scaled by the weight the regulator places on that indicator and the per-capita payment. This term reduces the effective marginal cost of effort along dimension k from $\partial c_\theta/\partial e_{k\theta}$ to $\partial c_\theta/\partial e_{k\theta} - (\partial \tau_{k\theta}/\partial e_{k\theta})\, w_k \, p_\theta$. Based on the insurer's FOC we derive the following results.

Proposition 1. *The insurer's optimal effort $e^*_{k\theta}$ along dimension k for consumer type θ is: (i) increasing in the per-capita performance payment p_θ, (ii) increasing in the weight w_k placed on indicator k, (iii) decreasing in the threshold I^*_k, (iv) decreasing in the cost variance σ^2_θ, (v) decreasing in the payment variance σ^2_B.*

The proof follows from applying the implicit function theorem and Topkis' theorem as seen in Appendix 2. Proposition 1 indicates that the insurer's optimal effort will be higher for tasks that are better rewarded, relatively easier to achieve, and lead to more certain expected profits.

Proposition 2 (Multitasking)**.** *An increase in the weight w_k placed on indicator k raises optimal effort $e^*_{k\theta}$ along dimension k but has no effect on optimal effort $e^*_{k'\theta}$ along any other dimension $k' \neq k$.*

Proof. Given the zero cross-partial assumptions on demand and costs, $FOC(e_{k'\theta}, p_\theta)$ for $k' \neq k$ does not depend on w_k. Hence $\partial e^*_{k'\theta}/\partial w_k = 0$ for all $k' \neq k$. The effect on $e^*_{k\theta}$ follows directly from part (ii) of Proposition 1. □

The multitasking result in Proposition 2 relies on the assumption that the impact of effort on one dimension does not depend on the level of effort on any other dimension. This is consistent with the disease-specific quality metrics under the HCA: the indicators used to evaluate insurer performance for HIV/AIDS patients, such as ART utilization rates and viral load monitoring, are clinically distinct from those used for cancer patients,

such as breast cancer screening rates or chemotherapy adherence. Hence, effort directed at improving care for HIV/AIDS patients has no effect on the quality indicators for cancer patients, and vice versa. If performance pay diverts effort away from unrewarded health conditions, then comparing outcomes between targeted and untargeted conditions would violate the Stable Unit Treatment Value Assumption (SUTVA), since the policy would worsen outcomes in the control group.

Proposition 3 (Moral Hazard)**.** *For any two types $\theta > \theta'$, performance payments $p_\theta > 0$ increase the effort gap $e^*_{k\theta} - e^*_{k\theta'}$ relative to the no-payment counterfactual. That is:*

$$\frac{\partial}{\partial p_\theta}\left(e^*_{k\theta} - e^*_{k\theta'}\right) > 0$$

The proof is in Appendix 2. Proposition 3 states that without performance payments, the insurer has weak incentives to invest in high-risk types precisely because they are more costly: the terms c_θ and $\partial c_\theta / \partial e_{k\theta}$ are larger for $\theta > \theta'$, reducing optimal effort relative to low-risk types (a moral hazard problem on the supply side). The introduction of $p_\theta > 0$ corrects this distortion by raising both the expected revenue from type θ through $\tau_{k\theta}\, p_\theta$ and the marginal return to effort through $(\partial \tau_{k\theta} / \partial e_{k\theta})\, w_k\, p_\theta$. Since neither of these terms appears in the first-order condition for θ', the payment exclusively lifts effort for the high type, widening the effort gap.

Proposition 4 (Disease targeting)**.** *Optimal effort $e^*_{k\theta}$ on disease-specific quality dimension k is higher under a payment scheme that targets specific diseases than under a scheme that aggregates across diseases, as long as the per-capita payment is not scaled up sufficiently to offset the reduction of marginal incentives across diseases.*

Proof. Suppose quality indicators, thresholds, and payments are aggregated across diseases into a single score $\overline{\tau}_k = \sum_\theta \omega_\theta \tau_{k\theta}$, where ω_θ are population weights, and that insurers receive a single per-capita payment $\overline{p}$ when meeting the aggregate score threshold. Then, the marginal return to effort in the FOC in equation (1) becomes:

$$\frac{\partial \overline{\tau}_k}{\partial e_{k\theta}} w_k \overline{p} = \omega_\theta \frac{\partial \tau_{k\theta}}{\partial e_{k\theta}} w_k \overline{p} < \frac{\partial \tau_{k\theta}}{\partial e_{k\theta}} w_k p_\theta$$

where the inequality follows from the fact that $\omega_\theta < 1$ and the assumption that the single per-capita payment $\overline{p}$ does not exceed the reduction in marginal incentives. □

Proposition 4 establishes that a performance payment scheme with disease-specific quality measures, thresholds, and per-capita payments induces strictly higher insurer effort than an untargeted scheme. The intuition is that targeting concentrates incentives: by tying payments directly to performance on a specific condition, the marginal return to effort on that condition is larger than under an aggregate scheme that dilutes incentives across diseases. However, this comes with a trade-off. A targeted scheme generates larger effort gains but only for patients with the targeted condition, while an untargeted scheme, though inducing lower effort on any given condition, spreads incentives across a broader patient population. The optimal design therefore depends on whether the social value of improvements in disease-specific care outweighs the gains from broader but weaker incentives across health conditions.

Another margin along which insurers may respond to performance pay is risk selection (Fioretti and Wang, 2023). For example, insurers may deny coverage to patients with targeted diseases that have high severity to increase the likelihood of meeting the quality thresholds. However, the HCA was designed to limit this response. While selecting less severe enrollees would increase the likelihood τ_θ of meeting quality thresholds, it would simultaneously reduce the per-capita payment p_θ, which is calculated as the average healthcare cost of condition θ. We elaborate on the policy's design in the next section. Selecting less severely ill enrollees therefore does not necessarily increase insurer revenues, which minimizes the incentive to engage in risk selection.

3 Institutional Background

To test the theoretical predictions, we use data from Colombia's healthcare system. This system was established in 1993 and is organized around two schemes. The contributory scheme covers the population who pay payroll taxes and their dependents (roughly half of the country's population), while the subsidized scheme is fully government-funded and

covers low-income populations (the other half of the population). Together they achieve near-universal coverage.

Within each scheme, private and public insurers offer a single standardized national health plan. Premiums are zero, cost-sharing rules are indexed to enrollees' monthly income, and the set of covered services is determined by the government uniformly across insurers. Competition therefore does not occur along the price or benefit dimensions but along the formation of provider networks and the negotiation of service prices with in-network providers. Individuals may enroll with any insurer operating in their municipality of residence and, while there is no formal open enrollment period, they must remain with their chosen insurer for at least 12 non-consecutive months before being allowed to switch.

Instead of premiums, insurers receive per capita risk-adjusted transfers from the government that control for the enrollees' sex, age, and municipality of residence. Because the formula is relatively coarse (Riascos et al., 2014), it fails to account for the higher costs associated with chronic conditions, making these enrollees predictably unprofitable (Serna, 2024).

Performance Pay Policy. To address the limitations of the coarse risk adjustment formula, the Colombian government introduced a complementary mechanism in 2007 known as the High-Cost Account (HCA). The HCA compensates insurers for the costs associated with specific high-burden diseases, initially covering end-stage renal disease without specifying quality metrics and expanding to HIV/AIDS in 2016 and to cancer in 2019 (breast, cervical, colon, stomach, prostate, leukemia, and non-Hodgkin lymphoma) while at the same time introducing care quality metrics for these conditions.[5]

Unlike the ex-ante risk adjustment formula, the HCA is an ex-post mechanism. Rather than compensating insurers based on predicted costs prior to enrollment, it redistributes capitated payments across insurers after enrollment decisions and health claims are made, conditioning these transfers on realized quality of care. This timing distinction is conse-

[5]See Resolution 1912 of 2015, Ministry of Health and Social Protection. Hemophilia has been covered under the HCA since 2016, though payments for this condition are not tied to quality metrics.

quential for insurer incentives. The ex-ante formula compensates insurers for consumers' private information about their health status but cannot contract on insurer behavior. The HCA contracts directly on insurer actions, potentially solving a moral hazard problem whereby insurers might otherwise underinvest in the care of their sickest enrollees.

The HCA operates as follows. All insurers contribute to a common pool of funds in proportion to their national market shares, regardless of whether they enroll patients with targeted diseases. To determine the size of this pool of funds, the government first computes the prevalence of disease θ at insurer j and age group a as:

$$\rho_{\theta aj} = 100{,}000 \cdot \frac{D_{\theta aj}}{N_{aj}}$$

where $D_{\theta aj}$ is the number of patients with disease θ in age group a enrolled with insurer j, and N_{aj} is the total number of enrollees in age group a at insurer j. Population-level prevalence for age group a is then $\rho_{\theta a} = 100{,}000 \cdot \frac{\sum_j D_{\theta aj}}{\sum_j N_{aj}}$.

The differential prevalence of disease θ at insurer j and age group a is $f_{\theta aj} = (\rho_{\theta aj} - \rho_{\theta a}) \cdot N_{aj}/100{,}000$, and total differential prevalence at the insurer is $f_{\theta j} = \sum_a f_{\theta aj}$. Letting p_θ denote the per-capita payment for disease θ (computed as the average healthcare cost in the population), the common pool of funds is:

$$CF_\theta = \sum_j \mathbf{1}\{f_{\theta j} > 0\} \cdot f_{\theta j} \cdot p_\theta$$

and each insurer contributes $b_\theta = CF_\theta \cdot N_j / \sum_j N_j$, where $N_j = \sum_a N_{aj}$. Intuitively, the common pool of funds is the expected amount of money needed to equalize disease prevalence across insurers. By making *all* insurers contribute to the pool of funds, the HCA implicitly transfers resources from low- to high-prevalence insurers.

Funds are then redistributed exclusively to insurers that meet government-specified quality thresholds for each targeted disease. Let $I^*_{\theta k}$ denote the threshold for the k-th quality indicator for disease θ, and let $I_{\theta kj}$ denote insurer j's measured value of that indicator. Population-weighted compliance with the k-th indicator at insurer j is defined

as:

$$q_{\theta kj} = \mathbf{1}\{I_{\theta kj} > I^*_{\theta k}\}(I_{\theta kj} - I^*_{\theta k})N_j$$

and the share of compliance attributed to indicator k is $\tau_{\theta kj} = q_{\theta kj}/\sum_j q_{\theta kj}$. Insurer j then receives a payment equal to:

$$Y_\theta = \sum_k w_{\theta k} \cdot \tau_{\theta kj} \cdot CF_\theta \tag{2}$$

where $w_{\theta k}$ is the weight the government assigns to the k-th quality indicator. Hence, only insurers with high disease prevalence that meet the quality thresholds get paid by the HCA in proportion to how far they are from the thresholds.

For HIV/AIDS, the quality thresholds include targets for the share of pregnant women screened for HIV/AIDS, the share of people with HIV/AIDS on ART with an adequate recent viral load, and the share of patients with early detection of the disease, among others. For cancer, they include the share of women screened for breast cancer, the share of men screened for prostate cancer, and the share of cancer patients in remission, among others.

For HIV/AIDS, contributions to the common pool were collected between November and December 2015, with funds redistributed to qualifying insurers in January 2016. For cancer, contributions were collected in December 2018 and redistributed in January 2019.[6]

The HCA is a zero-sum mechanism that redistributes resources from insurers with below-average disease prevalence to those with above-average prevalence who also meet the quality thresholds. This design imposes no fiscal burden on the government or the health system, but it introduces a potential downside. Insurers that do not expect to meet the quality thresholds may reduce care quality or engage more strongly in risk selection in order to minimize costs, since they will contribute to the pool of funds regardless of their performance. Thus, we will examine heterogeneity across insurers based on their predicted probability of receiving payments, to test whether those with low expectations reduce effort.

[6] See Resolution 5036 of 2015 and Resolution 3339 of 2019.

4 Data

We use individual-level enrollment and health claims data for the roughly 24 million people covered by the contributory system between 2011 and 2023. The enrollment files are snapshots of enrollment in June of every year, reporting the individual's insurer, municipality of residence, sex, and age. The health claims data report all the healthcare services that an individual received every year (outpatient, inpatient, and prescriptions) and include information such as the insurer, provider, negotiated service price, procedure code, and diagnoses following the International Classification of Diseases (ICD) nomenclature.

The enrollment records have information from all insurers including those in the subsidized system. That is, we observe individuals who are ever enrolled in the contributory system including for years in which they move into the subsidized system. In the health claims data, we only have information for insurers in the contributory system. So, if an individual moves to the subsidized system in a given year we do not observe their healthcare use that year.

A key challenge for our empirical design is defining the treatment group and treatment timing when individuals may be diagnosed with multiple diseases in different years. For example, an individual diagnosed with breast cancer in 2013 and lung cancer in 2019 would belong to the pre-policy cohort based on the first diagnosis but to the post-policy cohort based on the second. To avoid this ambiguity, we restrict the cancer treatment group to individuals diagnosed with exactly one targeted cancer type who have no HIV/AIDS diagnosis during the sample period.

For the HIV/AIDS analysis, the treatment group comprises individuals diagnosed with HIV/AIDS with no concurrent cancer diagnosis, but allowing these patients to potentially be diagnosed with any other chronic health condition. We exclude HIV/AIDS patients ever enrolled with insurers EPS013 and EPS003, which were terminated by the government around the time the HIV/AIDS policy was introduced; excluding them ensures that our outcomes are not confounded by the care disruptions associated with these terminations (Buitrago et al., 2025). Appendix 1 describes our data cleaning process and the procedure for classifying patients as diagnosed with a targeted condition using the

claims data.

We construct the control group for the cancer policy from conditions that are clinically similar to the targeted cancers but excluded from the HCA. For instance, stomach cancer (ICD-10 code C16) is a targeted condition, so the control group includes other cancers of the digestive system, such as malignant neoplasms of the esophagus (C15), small intestine (C17), and liver and intrahepatic bile ducts (C22). Appendix Table 1 lists the ICD codes defining the treatment and control groups for cancer. To validate the construction of diagnoses from ICD codes, Appendix Figure 1 compares the prevalence of targeted cancers in our sample against statistics from the contributory system reported to the HCA, finding that the two are closely aligned.

Unlike the cancer setting, where the HCA generates within-cancer type variation and thus a natural comparison group, the HIV/AIDS policy applies uniformly across all HIV/AIDS patients, leaving no untreated subgroup within the same disease category. Thus, to construct a control group we focus on diseases that share the chronic-care profile of HIV/AIDS (which requires continuous medication and viral load monitoring): tuberculosis, asthma, hypertension, diabetes, and epilepsy.[7] Nevertheless, even among these plausible comparators, there is no reason to expect each disease to track HIV/AIDS's pre-policy spending trend equally well, and a simple average across them implicitly assumes equal relevance. To address this, we use a synthetic control method to construct a weighted combination of these control diseases with weights chosen specifically to reproduce HIV/AIDS's pre-period total healthcare spending trajectory, age at diagnosis, and share of males.

We implement this matching at the level of 3-digit ICD codes. For each 3-digit code, we compute the average healthcare spending across all patients in that diagnosis per year from 2011 to 2015, the average age, and the share of males. Weights on the control diseases are chosen by solving a constrained quadratic program that minimizes the distance between the characteristics of treated 3-digit codes and a weighted combination of the control diseases' characteristics, subject to the weights being non-negative and summing to one, and with a ridge penalty added to discourage the solution from concentrating weight on

[7] Appendix Table 1 lists the ICD codes used to identify HIV/AIDS patients and its control conditions.

a single control disease (Abadie and L'hour, 2021).

5 Empirical Design

After defining treatment and control groups, we examine pre-policy outcome trends to motivate our empirical design. The top panel of Appendix Figure 2 shows that healthcare spending among both treated and control cancer patients is low in the years preceding diagnosis and trends upward gradually. At the time of diagnosis, spending spikes sharply before tapering off and returning to near pre-diagnosis levels by the third year post-diagnosis. Healthcare utilization for cancer patients is therefore heavily concentrated in the first three years following diagnosis. The bottom panel of Appendix Figure 2 reports corresponding spending trends for HIV/AIDS patients and their control conditions. Consistent with their chronic nature, most diseases exhibit relatively flat post-diagnosis spending trajectories, with asthma and epilepsy as exceptions, where spending declines one year after diagnosis.

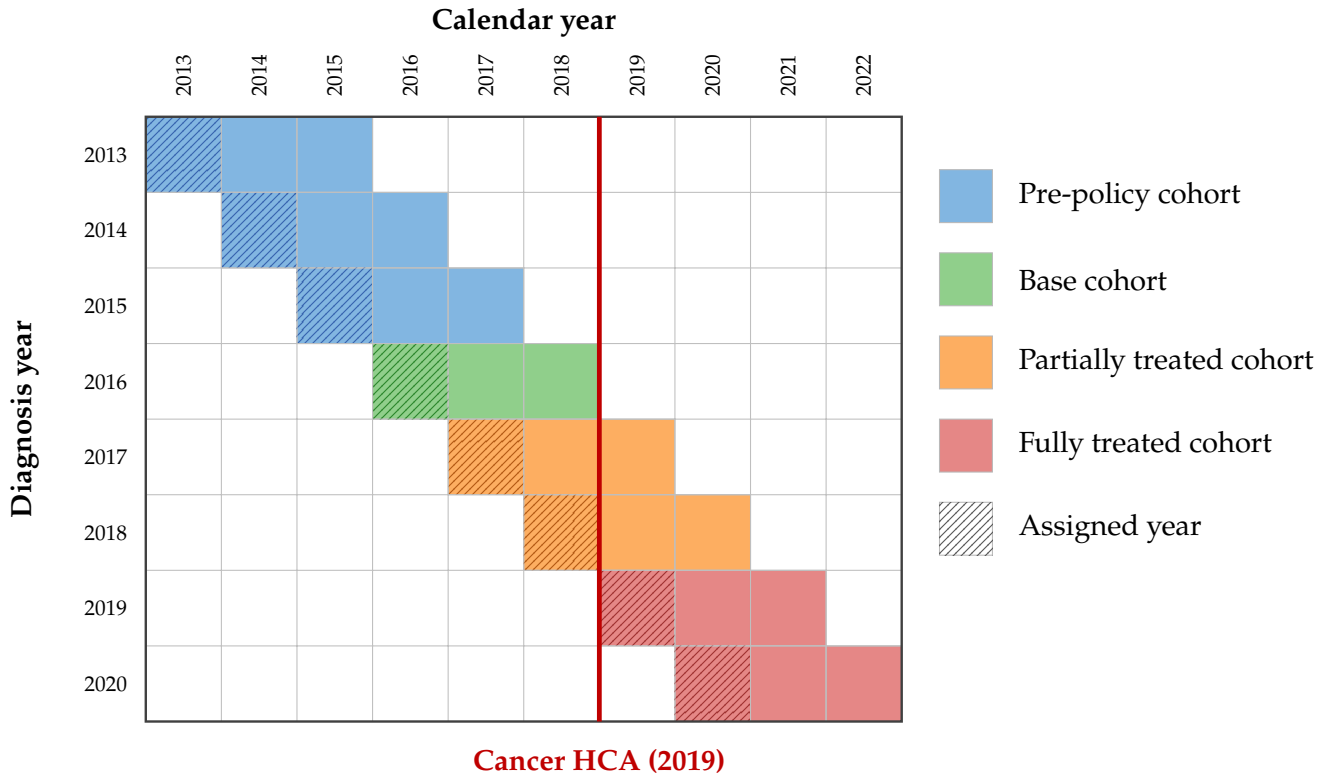
These patterns inform our empirical design. We use a cohort event study that compares cohorts of individuals diagnosed with a targeted disease before the policy against those diagnosed after, using at most three years of post-diagnosis data, where healthcare use is concentrated. A simple comparison of pre- and post-policy cohorts would be equivalent to an interrupted time series, which is vulnerable to confounding from contemporaneous changes unrelated to the HCA. Hence, we introduce the control group of untargeted diseases. Our estimator computes the difference in outcomes between targeted and untargeted cohorts diagnosed after the policy, subtracts the analogous difference from cohorts diagnosed before the policy, and conditions these comparisons on the calendar year relative to the policy. This approach yields a classic event study design, formally in the equation below:

$$y_i = \sum_{k=-5, k\neq -1}^{4} \beta_k \mathbf{1}\{t - t^* = k\} T_{d(i)} + \gamma_{d(i)} + \delta_{t(i)} + \lambda_{j(i)} + \varepsilon_i \tag{3}$$

where i denotes an individual, $t(i)$ denotes their diagnosis year, t^* is the year of the HCA

policy, $T_{d(i)}$ is an indicator for whether individual i has a diagnosis targeted by the HCA, $\gamma_{d(i)}$ is a diagnosis dummy (defined by the 3-digit ICD code), $\delta_{t(i)}$ is a cohort fixed effect, and $\lambda_{j(i)}$ is an insurer fixed effect. We cluster standard errors at the level of 3-digit ICD code, since this determines treatment by the policy, and estimate the regression separately for the cancer and HIV/AIDS policies.

FIGURE 1: Cohort Event Study Design



Note: Figure presents the cohort design for our event study specification. Each row corresponds to a diagnosis cohort and column to a calendar year. Outcomes are averaged over the three-year window following diagnosis and assigned to the diagnosis year. The base cohort (2016) is the last cohort whose outcome window falls entirely in the pre-policy period. The 2017 and 2018 cohorts are partially treated. Cohorts diagnosed in 2019 and later are fully treated. The hatched cell in each row indicates the assigned (diagnosis) year.

Figure 1 illustrates the construction of the cohort design for the cancer HCA. The design for HIV/AIDS follows the same approach. Each row corresponds to a diagnosis cohort and each column to a calendar year. For each patient, we compute the sum of outcomes over the three-year window following diagnosis and assign the result to the year of diagnosis. Cohorts whose outcome windows fall entirely before the policy (blue) serve as the pre-policy comparison group. The 2016 cohort (green) serves as the base year ($k = -1$), as it is the last cohort whose outcome window falls entirely in the pre-period.

The 2017 and 2018 cohorts are partially treated (orange), since portions of their outcome windows fall within the post-policy period. Cohorts diagnosed in 2019 and later (red) are fully treated.

For HIV/AIDS, the pre-policy comparison group consists of cohorts diagnosed in 2011–2013. Cohorts diagnosed in 2014 and 2015 are partially treated, and 2016 marks the first fully treated cohort. For both analyses, we exclude patients diagnosed in 2022 and 2023 for whom outcomes are censored.

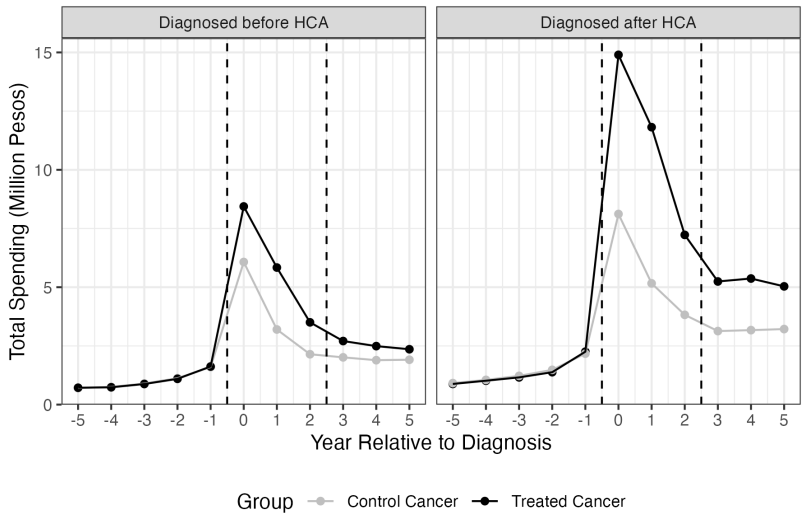
Appendix Table 4 presents summary statistics of treatment and control groups in the pre-policy period for the cancer analysis in Panel A and the HIV/AIDS analysis in Panel B. For cancer, the two groups are balanced in terms of the fraction of males and the pre-diagnosis healthcare spending. The treatment group gets diagnosed at slightly older ages and has higher spending on outpatient care and prescription medications. HIV/AIDS patients are disproportionately male, have an average age of 36 years, and represent substantially higher healthcare spending than their comparator diseases.

5.1 Identification

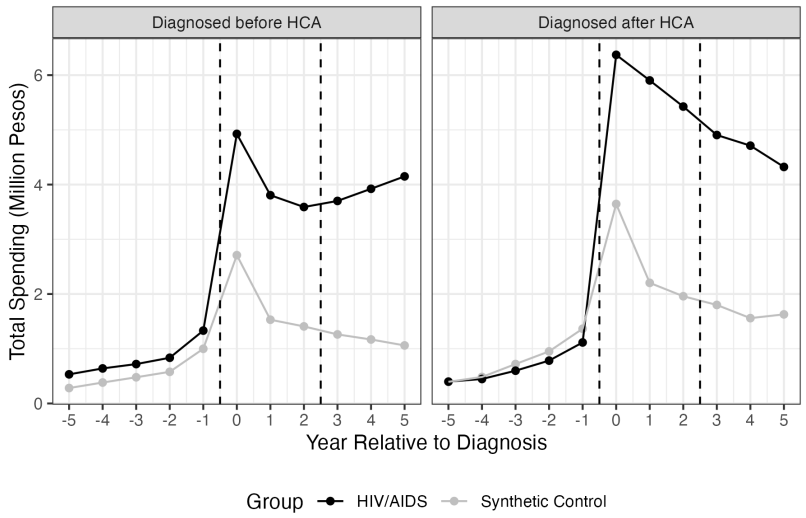
Valid identification of the policy's effects rests on the following assumptions. The first assumption is the absence of contemporaneous policy confounders, such as changes in sample composition or the COVID-19 pandemic. Sample composition is particularly worrisome for a cohort analysis which naturally has different sets of patients in the pre- and post-policy periods. The pandemic confounds effects only if targeted diseases were disproportionately more likely to be affected by COVID-19. The second is assumption SUTVA. This assumption may be violated if insurers improve dimensions of care quality that apply broadly to all enrollees, in which case treatment effects would be underestimated. Conversely, if insurers divert effort away from untargeted conditions, control group outcomes would worsen and treatment effects would be overestimated. The third assumption is parallel pre-trends, requiring that treated and control individuals exhibit parallel outcome trends before the HCA. We assess the plausibility of each of these assumptions below.

FIGURE 2: Healthcare Spending Trends Between Treatment and Control Groups

(A) Cancer Groups



(B) HIV/AIDS Groups



Note: Figure presents the trends of medical spending around diagnosis of cancer patients by whether the patient is diagnosed before the HCA policy. The dashed lines represent the aggregation window used in the main DiD analysis. Panel A presents treatment and control group trends for the cancer HCA. Panel B presents treatment and control group trends for the HIV/AIDS HCA. Cancers affected by HCA are defined as treated and others as controls. See Appendix Table 1 for the ICD codes used.

The first concern related to sample composition is that performance pay may alter diagnosis rates, changing the composition of post-policy cohorts. If the policy incentivizes earlier or more frequent screening, post-policy cohorts may systematically contain patients diagnosed at earlier disease stages relative to pre-policy cohorts, which would lead us to attribute differences to true improvements in care quality rather than differences in disease

severity. We assess this potential source of bias by examining whether diagnosis rates and patient characteristics change around the policy introduction. Appendix Figures 3 and 4 compare age at diagnosis, number of diagnosed patients, share of male patients, and total pre-diagnosis healthcare spending between treatment and control groups for the cancer and HIV/AIDS policies respectively. We find no significant changes in these placebo outcomes after the HCA, with the exception of a small reduction in the male share following the cancer policy, suggesting that our main results are unlikely to be driven by compositional changes in the patient pool.

Next, we assess the validity of the parallel pre-trends and SUTVA assumptions by plotting raw outcome trends for treated and control groups. Figure 2 presents the trend in total healthcare spending by year relative to diagnosis for cohorts diagnosed before the policy in the left panels and cohorts diagnosed after the policy in the right panels. Panel A presents trends for the cancer policy and Panel B for the HIV/AIDS policy. The solid black line represents the treatment group and the solid gray line represents the control group. The dashed vertical lines represent the three years after diagnosis that we use in our cohort design.

Several patterns are worth noting. First, pre-policy, post-diagnosis spending trends are similar between treated and control groups for both cancer and HIV/AIDS, consistent with parallel pre-trends. Second, spending spikes sharply at the time of diagnosis for both treated and control patients and then declines rapidly for cancer, confirming the front-loading of healthcare utilization that motivates our use of at most three years of post-diagnosis data. Third, the gap between treated and control spending is visibly larger in the post-policy cohorts than in the pre-policy cohorts. In the pre-policy cohorts, treated cancer patients spend roughly 2 million pesos ($680) more than controls at the time of diagnosis; in the post-policy cohorts, this gap widens to approximately 7 million pesos ($2,400), suggesting that the cancer policy induced a meaningful increase in healthcare spending among targeted patients. Similarly, HIV/AIDS patients spend 2.5 million pesos ($845) more than control diseases at diagnosis before the policy, widening to nearly 4 million pesos ($1,350) afterward. Fourth, spending patterns among control cancer patients and control diseases for HIV/AIDS are strikingly similar across the pre- and post-policy

cohorts in every year relative to diagnosis, suggesting that the policy did not spill over to untargeted patients and that SUTVA holds in our analysis. Control groups experience a slight increase in spending in the post-policy period perhaps reflecting the impact of the COVID-19 pandemic.

5.2 Outcomes

We consider two classes of outcomes: healthcare spending and care quality. For spending, we examine total healthcare spending, spending by service type (outpatient care and prescription medications), and spending on disease-specific treatments—chemotherapy, radiotherapy, and cancer related surgery for all cancer patients; prescriptions of antiretroviral therapy (ART) for HIV/AIDS; antimycobacterials for tuberculosis; antiepileptics for epilepsy; drugs for obstructive airway diseases for asthma; insulins and blood glucose lowering drugs for diabetes; and antihypertensives, diuretics, beta blockers, calcium channel blockers, and renin-angiotensin agents for hypertension. Appendix Table 2 lists the services and ATC codes used for constructing these outcomes.

For care quality, the cancer analysis uses indicators for hospital admission, cancer complications such as metastasis, and receipt of the cancer-specific treatments listed above.[8] The HIV/AIDS analysis uses indicators for vaccination against bacterial and viral diseases (e.g., pneumococcus and flu), and hospital admissions. Appendix Table 3 lists the ICD codes and services used to identify HIV/AIDS complications and vaccinations.

6 Effects of Performance Pay for Cancer

Figure 3, Panel I presents results on spending outcomes for the cancer HCA. Panel A shows that total healthcare spending per patient evolved in parallel across targeted and control cancers prior to the policy, then rose by 7.4 million pesos ($2,500) on average after its introduction, or 32.0% relative to the baseline treated mean of 23.1 million pesos ($7,800).

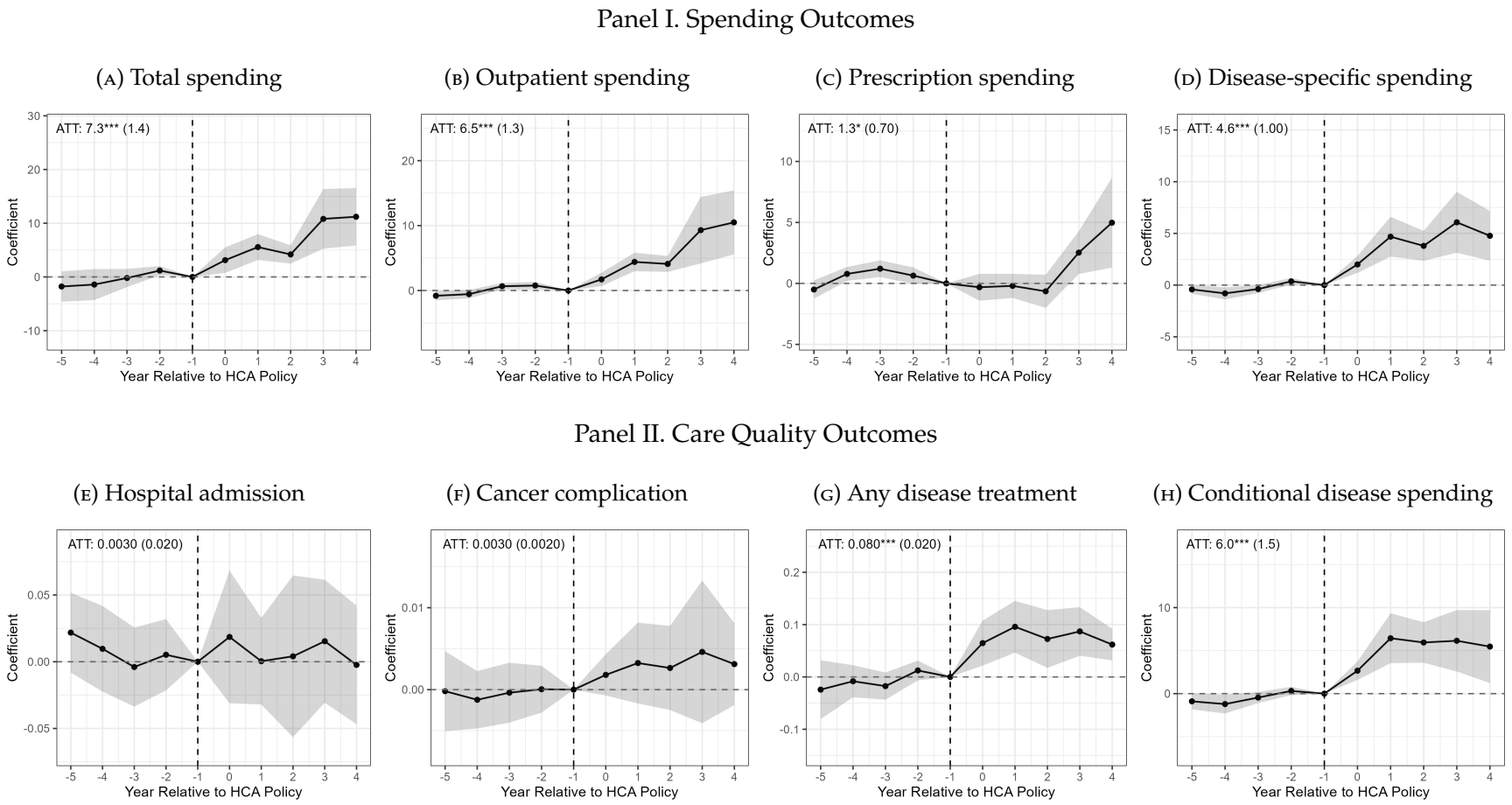
[8] We do not classify metastasis as a distinct cancer type; individuals experiencing this complication alongside their primary cancer diagnosis therefore remain in our sample under the restriction that patients be diagnosed with only one condition over the sample period.

The smaller effects in the first two post-policy years reflect the inclusion of partially treated cohorts, whose outcome windows only partially overlap with the post-policy period; the larger effects in later years reflect fully treated cohorts only. The spending increases we document are not a mechanical consequence of performance payments themselves, since our outcome measures spending on healthcare services delivered to patients, not insurer revenues. The results are also not attributable to the COVID-19 pandemic since we identify effects even in the years before the pandemic.

The increase in total spending is concentrated in outpatient services, which increase by 6.5 million pesos ($2,200) (Panel B), and prescription medications, which increase by a similar amount in the last year of the sample period where only fully treated cohorts contribute to the effect (Panel C). Within outpatient care, the gains are largest for services directly related to cancer management, such as chemotherapy, radiotherapy, and cancer-related surgeries, which increase 106.3% relative to the baseline treated mean of 4.4 million pesos ($1,500) (Panel D). These patterns suggest that the cancer HCA expanded access to cancer treatment rather than driving broad increases in healthcare utilization.

Figure 3, Panel II presents event study results on care quality outcomes. Panels E and F show no significant change in hospital admissions or cancer complications such as metastasis, suggesting the spending increases documented above reflect better care rather than deteriorating health. The rise in disease-specific spending reflects: (i) a 7.9 percentage point (p.p.) increase in service utilization equivalent to 13.6% of the baseline treated mean (Panel G), and (ii) a 6.0 million pesos ($2,000) increase in spending conditional on claiming at least one cancer treatment or 80.1% relative to the baseline treated mean (Panel H). Overall, these results indicate that performance pay for insurers increased effort among targeted cancers in a way that improved disease management, both by driving previously un-managed patients into treatment and by providing more intensive services to patients already receiving care. This is consistent with Proposition 1, which predicts that performance pay raises insurer effort on rewarded tasks.

FIGURE 3: Impact of Cancer HCA on Healthcare Spending and Care Quality



Note: Figure presents coefficients and 95% confidence intervals of the cohort DiD design for the effect of HCA policy on medical utilization of cancer patients (Equation 3). Cancer complication is defined as having metastasis or other secondary cancer. Cancer treatment includes chemotherapy, radiotherapy, and cancer-related surgeries (see Appendix Table 2 for details). 'Conditional disease spending' measures the effect on treatment spending conditional on having any cancer treatment. Utilization outcomes are measured in the year of the diagnosis and the two years after, where cancer-induced utilization are most concentrated. The text on the top-left reports the coefficient and standard deviation in parentheses of the effect aggregated in the post-periods. Standard errors are clustered at the ICD-3 code level.

6.1 Test of Propositions

Healthcare spending variance. Proposition 1(iv) predicts that optimal insurer effort is decreasing in the uncertainty of costs across health conditions. To test this, we compute the pre-policy variance of healthcare costs for each 3-digit ICD code and split the sample into patients with above- and below-median cost variance (with the median across 3-digit ICD codes weighted by patient counts).

Table 1 presents the results. Treatment effects are substantially and statistically larger for low-variance conditions across all spending outcomes. For total healthcare spending, the difference between low- and high-variance groups equals 5.1 million pesos ($1,740) or 22% relative to the pre-policy treated mean. This pattern holds for outpatient spending, prescription medications, and disease-specific spending, where low-variance conditions drive most of the spending increase documented in the main results.

These differential effects are unlikely to reflect reversion to the mean. If low-variance cancers had lower pre-policy spending and were simply reverting toward a long-run level, we would expect to see increasing trends in the pre-policy period for this group. However, full event study plots in Appendix Figure 5 show flat and parallel pre-policy trends for both low- and high-variance conditions, suggesting the divergence after the policy reflects a true treatment effect.

TABLE 1: Heterogeneity of Cancer HCA by Spending Variance

Outcome	Low Variance	High Variance	p-value (Difference)
Total spending	7.56*** (0.63)	2.43 (2.03)	0.02**
Outpatient spending	6.96*** (0.77)	2.59 (1.62)	0.01**
Prescription spending	3.26*** (0.70)	-0.23 (0.99)	0.00***
Conditional disease spending	4.73*** (0.70)	1.80* (0.97)	0.01**
Any cancer treatment	0.09*** (0.01)	0.02* (0.01)	0.00***
Cancer complication	0.00 (0.00)	0.00 (0.00)	0.68
Hospital admission	-0.01 (0.03)	0.02** (0.01)	0.30

Note: Table presents coefficients and standard errors of a 2x2 cohort difference-in-differences design for the effect of HCA policy on cancer patients (Equation 3) by variance of healthcare spending. For each type of cancer, we measure the variance of the healthcare costs among its patients in the pre-policy period. We then split the sample by whether the patient's cancer has a cost variance above the median, weighted by patient counts. Cancer complication is defined as having metastasis or other secondary cancer. Cancer treatment includes chemotherapy, radiotherapy, and cancer-related surgeries (see Appendix Table 2 for details). 'Conditional Cancer treatment spending' measures the effect on treatment spending conditional on having any cancer treatment. Utilization outcomes are measured in the year of the diagnosis and the two years after, where cancer-induced utilization are most concentrated. Standard errors are clustered at the ICD-3 code level.All spending variables are measured in millions of COP. * p<0.1, ** p<0.05, *** p<0.01

For low-variance cancers, we also find a meaningful increase in the likelihood of receiving cancer-related treatments, while the high-variance group exhibits a smaller and less precise response. The difference between the two groups in this outcome equals 6.5 p.p, which is significant at the 1% level. In contrast, we find no heterogeneity in hospital admissions and cancer complications. Hence, the results support the theoretical prediction that insurers exert less effort when costs for a particular health condition are highly uncertain.

Likelihood of meeting quality thresholds. Another prediction of our model is that optimal effort under performance pay is higher the easier it is to meet the quality thresholds and the lower is the variance of the payment (Proposition 1(iii), (v)). To test the first prediction we take advantage of the fact that 2019 values of the quality indicators for each insurer were calculated by the government using 2018 health claims data and made publicly available by the HCA institute.[9] Using these pre-policy values, we construct for each insurer-3 digit ICD code combination a measure of the predicted likelihood of receiving HCA payments. This likelihood is defined as the share of quality indicators that the insurer meets. Specifically, we approximate equation (2) as:

$$\hat{y}_{\theta j} = \sum_{k} w_{\theta k} \hat{\tau}_{\theta k j}$$

where

$$\hat{\tau}_{\theta k j} = \frac{\mathbf{1}\{\hat{I}_{\theta k j} > I^{*}_{\theta k}\}}{\sum_{j} \mathbf{1}\{\hat{I}_{\theta k j} > I^{*}_{\theta k}\}}$$

and where $\hat{I}_{\theta k j}$ is the value of indicator k for insurer j and 3-digit ICD code θ. Then, we aggregate across health conditions for each insurer weighting by patient counts $\hat{y}_j = \sum_\theta w_\theta \hat{y}_{\theta j}$, and split the treatment group into patients whose insurer falls above or below the median of $\hat{y}_j$.

This approach is valid as long as $\hat{y}_{\theta j}$ is a good predictor of both future quality threshold

[9] Measured values of the quality indicators and HCA contributions and payments per insurer-year are available at https://www.minsalud.gov.co/salud/pos/paginas/unidad-de-pago-por-capitacion-upc.aspx, section "Ajuste Ex-Post."

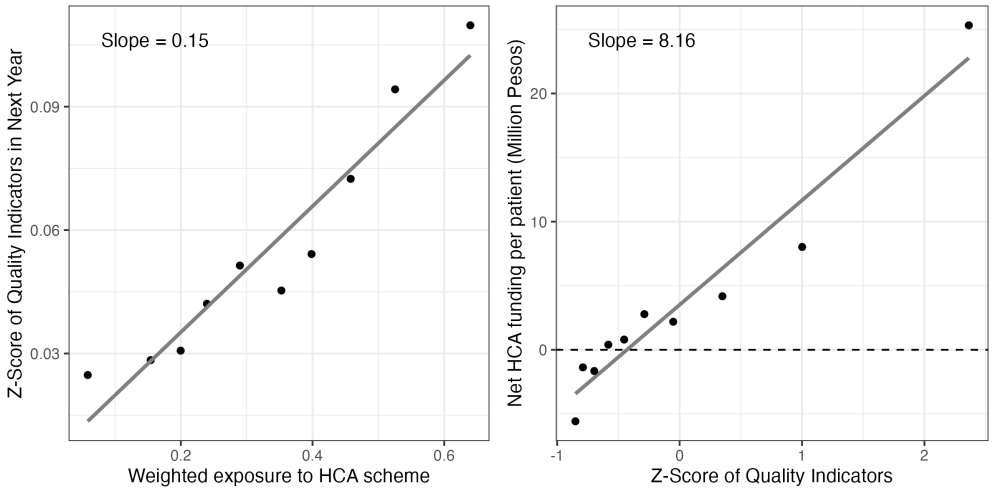
compliance and the magnitude of HCA payments. Figure 4 provides support for these conditions. Panel A plots $\hat{y}_{\theta j}$ against its 2020 counterpart and shows a strong positive correlation, indicating that an insurer's likelihood of compliance is persistent over time. This persistence is partly mechanical because disease prevalence tends to be stable across years, but reflects a strong signal of insurer behavior more broadly.

Panel B plots $\hat{y}_{\theta j}$ against the net per-capita HCA transfer, defined as payments received minus contributions made ($Y_\theta - b_\theta$ from section 3), divided by the number of patients with disease θ. The strong positive relationship confirms that insurers with a higher predicted likelihood of compliance receive larger net transfers, validating $\hat{y}_{\theta j}$ as a measure of expected financial gain from the HCA.

FIGURE 4: Predicting Payments and Likelihood of Meeting Quality Thresholds

(A) Intertemporal Correlation of Quality Scores

(B) HCA Payments and Quality Scores



Note: Panel (A) presents the correlation between likelihood of receiving performance payments (as defined in Equation 2) and the standardized value of the quality indicators in the next year. Panel (B) presents the correlation between standardized quality indicator and net HCA funding per patient in the year. An observation is an insurer-year-cancer. The dots report the conditional means in each bin, the line is a linear fit.

Table 2 presents the results. Treatment effects for total healthcare spending are 50% larger among high-compliance likelihood insurers. These insurers increase total spending by 5.0 million pesos more ($1,700) and outpatient spending by 2.5 million pesos more ($857) than their low-compliance counterparts. Spending on cancer-related treatments follows the same pattern, and it is driven entirely by utilization rather than prices: the

TABLE 2: Heterogeneity of Cancer HCA by Likelihood of Compliance

Outcome	Low Compliance	High Compliance	p-value (Difference)
Total spending	5.53*** (1.25)	10.55*** (2.61)	0.08*
Outpatient spending	5.64*** (1.18)	8.17*** (2.04)	0.28
Prescription spending	1.84** (0.77)	0.85 (0.61)	0.32
Conditional disease spending	4.47*** (0.94)	8.07*** (2.76)	0.22
Any cancer treatment	0.05*** (0.02)	0.15*** (0.04)	0.03**
Cancer complication	0.00** (0.00)	0.00 (0.00)	0.41
Hospital admission	0.00 (0.02)	0.01 (0.02)	0.89

Note: Table presents coefficients and standard errors of a 2x2 cohort difference-in-differences design for the effect of HCA policy on cancer patients (Equation 3) by likelihood of compliance. For each insurer, we measure the exposure to the HCA policy by the average policy weighted sure of quality indicators (Equation 2). We then split the sample by whether the patient's insurer has a exposure above the median, weighted by patient counts. Cancer complication is defined as having metastasis or other secondary cancer. Cancer treatment includes chemotherapy, radiotherapy, and cancer-related surgeries (see Appendix Table 2 for details). 'Conditional Cancer treatment spending' measures the effect on treatment spending conditional on having any cancer treatment. Utilization outcomes are measured in the year of the diagnosis and the two years after, where cancer-induced utilization are most concentrated. Standard errors are clustered at the ICD-3 code level. All spending variables are measured in millions of COP. * p<0.1, ** p<0.05, *** p<0.01

likelihood of claiming cancer-related treatments is 10 p.p. higher among high-compliance insurers than low-compliance ones. We also find no differences in hospital admissions or cancer complications across groups.

Performance pay uncertainty. To test the second prediction that effort is decreasing in payment uncertainty, Table 3 splits the sample by predicted payment variance, defined as $\hat{y}_j(1-\hat{y}_j)$, above or below the patient-weighted median. Insurers with below-median payment variance exhibit treatment effects that are twice as large as their counterparts for outcomes such as total healthcare spending, outpatient spending, and spending on cancer-related procedures. For instance, the difference between the two groups in the latter outcome equals 5.3 million pesos ($1,800) or 23% relative to the baseline treated mean. We find no corresponding differences in hospital admissions or cancer complications, mirroring the pattern from the cost-variance and compliance-likelihood heterogeneity results. The findings thus suggest that whether performance pay translates into better care depends not only on the level of expected payments but on how confidently insurers can predict receiving them.

TABLE 3: Heterogeneity of Cancer HCA by Variance of Compliance Likelihood

Outcome	Low Variance	High Variance	p-value (Difference)
Total spending	9.64*** (2.01)	4.06*** (1.39)	0.02**
Outpatient spending	8.48*** (1.62)	3.89*** (1.22)	0.02**
Prescription spending	1.11* (0.65)	1.29 (0.85)	0.87
Conditional disease spending	8.42*** (2.21)	3.08*** (0.86)	0.02**
Any cancer treatment	0.10*** (0.03)	0.05* (0.02)	0.18
Cancer complication	0.00* (0.00)	0.00 (0.00)	0.59
Hospital admission	0.02 (0.02)	-0.02 (0.04)	0.34

Note: Table presents coefficients and standard errors of a 2x2 cohort difference-in-differences design for the effect of HCA policy on cancer patients (Equation 3) by variance of the compliance likelihood. For each insurer, we measure the exposure to the HCA policy by the average policy weighted sure of quality indicators (Equation 2). We then split the sample by whether the patient's insurer has a exposure above the median, weighted by patient counts. Cancer complication is defined as having metastasis or other secondary cancer. Cancer treatment includes chemotherapy, radiotherapy, and cancer-related surgeries (see Appendix Table 2 for details). 'Conditional Cancer treatment spending' measures the effect on treatment spending conditional on having any cancer treatment. Utilization outcomes are measured in the year of the diagnosis and the two years after, where cancer-induced utilization are most concentrated. Standard errors are clustered at the ICD-3 code level. All spending variables are measured in millions of COP. * p<0.1, ** p<0.05, *** p<0.01

6.2 Robustness Checks

We conduct several robustness checks to assess the validity of our cancer HCA results. These exercises are reported in Appendix Figure 6. For each of the outcomes, we report estimates under five alternative specifications. First, we include diagnosis-age-gender fixed effects and insurer-year fixed effects as additional controls in equation (3). Second, we exclude individuals who are ever enrolled with insurers SaludCoop (EPS013) or Cafesalud (EPS003), which were terminated in 2015 and 2019, respectively (Serna, 2025). This exercise ensures our results are not driven by care disruptions associated with these forced transitions. Fourth, we exclude individuals who were ever enrolled with insurers Nueva EPS (EPS037), Salud Total (EPS002), and Sanitas (EPS005) which exhibit anomalous jumps in cancer diagnoses in certain years, to verify that our results are not sensitive to these idiosyncratic data patterns. Fifth, we exclude individuals who die during the sample period. Our main results are remarkably robust across these specifications. Finally, Appendix Figure 7 shows that treatment effects are not driven by any single cancer type but are present across all targeted conditions.

7 Effects of Performance Pay for HIV/AIDS

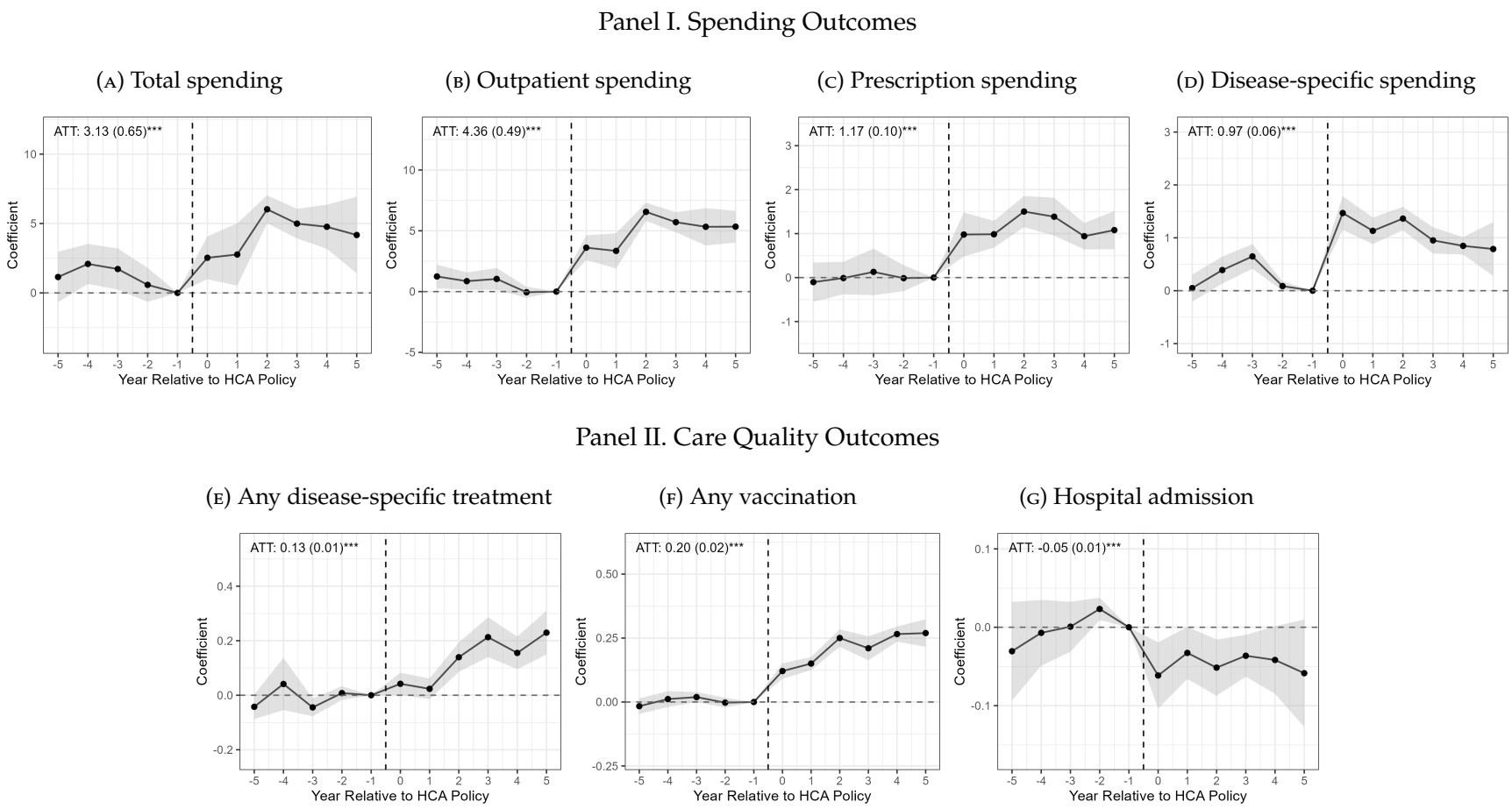
Figure 5, Panel I presents results on spending outcomes for the HIV/AIDS policy. Across all outcomes we find strong evidence of parallel pre-trends, with the exception of disease-specific spending, which drifts modestly upward before the policy, a small movement relative to the sharp post-policy increase (Panel D). Panel A shows that performance pay increased total healthcare spending by 3.1 million pesos ($982), or 25.4% relative to baseline. This increase is driven by outpatient care services (Panel B), which rise by 58.8%, and by prescription spending (Panel C), driven primarily by ART prescriptions. ART spending alone increased by nearly 1 million pesos ($338) after the policy— almost a third of the annual per-patient cost of ART in the pre-period.[10]

Figure 5, Panel II presents results on care quality outcomes. Panel E shows that the likelihood of claiming disease-specific treatments (corresponding to ART prescriptions for HIV/AIDS) increased by 13.0 p.p. after the policy or 17.1% relative to baseline. This increase suggests that the policy induced better adherence to clinical guidelines by softening insurers' financial exposure to funding expensive ART prescriptions.

Because HIV compromises the immune system's ability to fight infection, vaccination against infectious diseases is a cornerstone of appropriate HIV/AIDS management. We find that performance pay increased the likelihood of patients being vaccinated against bacterial and viral infection by 20 p.p. or 157.4% relative to baseline. Panel G likewise shows significant reductions in the likelihood of hospital admission (16.9%). Together, this evidence suggests that performance pay improved the quality of HIV/AIDS care: patients were both better protected against the infections that commonly complicate the disease and less likely to experience the downstream health consequences of such complications.

[10] See https://hivpreventioncoalition.unaids.org/es/news/colombia-issues-first-ever-compulsory-license-landmark-step-expanding-access-affordable-hiv for details on per-capita ART treatment costs in Colombia.

FIGURE 5: Impact of HIV HCA on Spending and Care Quality



Note: Figure presents results of the cihort event study design for the effect of HCA policy on HIV patients. The text on the top-left reports the coefficient and standard deviation in parentheses of effect on being diagnosed after the policy (Equation 3). Antiretroviral therapy (ART) is defined as claims with an ATC code of J05AF/G/E/J/R.

7.1 Test of Propositions

Unlike the cancer analysis, we cannot reconstruct the insurer-level quality scores for HIV/AIDS because several of the HCA thresholds involve laboratory results, such as viral load measurements, that are not reported in the claims data. This limits our ability to test the predictions of our model related to compliance likelihood and payment uncertainty. The design does, however, allow us to test the prediction that impacts on effort are larger when cost uncertainty is low.

TABLE 4: Heterogeneity of HIV/AIDS HCA by Spending Variance

Outcome	Low Variance	High Variance	p-value (Difference)
Total spending	4.67*** (0.54)	-0.78 (1.87)	0.01***
Outpatient spending	5.06*** (0.40)	-0.34 (0.74)	0.00***
Prescription spending	1.18*** (0.08)	0.09 (0.35)	0.00***
Chronic disease drug spending	1.32*** (0.18)	0.32*** (0.12)	0.00***
Any chronic disease drug	0.12*** (0.03)	-0.09 (0.06)	0.00***
Any HIV complication	-0.02*** (0.00)	-0.05*** (0.01)	0.03**
Hospital admission	-0.06*** (0.01)	-0.00 (0.05)	0.26

Note: Table presents results of the event study design for the effect of HCA policy on the care quality of HIV/AIDS patients by spending variance above or below the median. The table reports the coefficient and standard error in parenthesis of the post-policy indicator from Equation 3. Antiretroviral therapy (ART) is defined as claims with an ATC code of J05AF/G/E/J/R. Vaccination is defined as vaccination against bacterial and viral diseases. HIV complication is AIDS-defining conditions. See Appendix Table 3 for details.

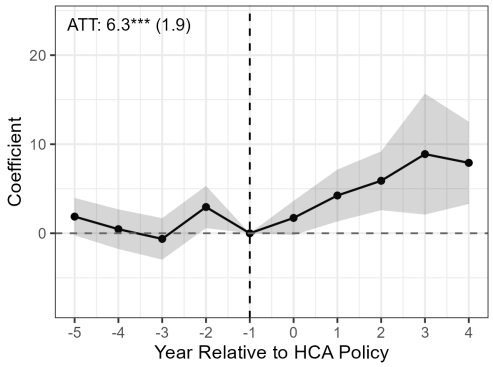
We split 3-digit ICD codes into above- and below-median pre-period spending variance (weighted by patient counts) and estimate the event study regression separately within each subsample. Table 4 presents the results, which confirm that treatment effects across all outcomes are driven by low-variance ICD codes related to HIV/AIDS. Strikingly, among high-variance conditions, we find no significant policy effects on outpatient spending, prescription spending, the likelihood of claiming ART prescriptions, or hospital admissions. The fact that we obtain qualitatively similar heterogeneity patterns in both the cancer and HIV/AIDS analyses suggests that our results are not an artifact of the particular targeted disease chosen for study.

8 The Role of Managed Care

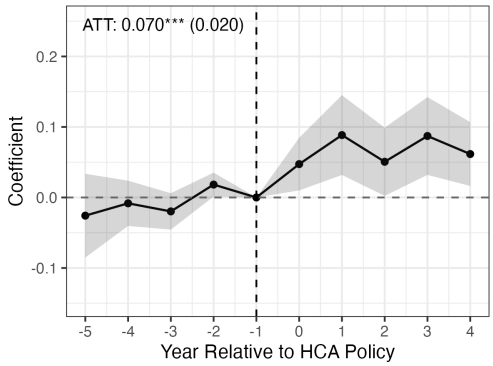
Our results show that performance pay worked as intended for both cancer and HIV/AIDS patients, increasing healthcare spending among targeted conditions and, more importantly, directing insurer effort toward the specific services that determine disease management. Understanding how insurers translate stronger financial incentives into better patient care is therefore a central question for the design of performance pay schemes in managed care health systems.

FIGURE 6: Mechanisms for Performance Pay Effects

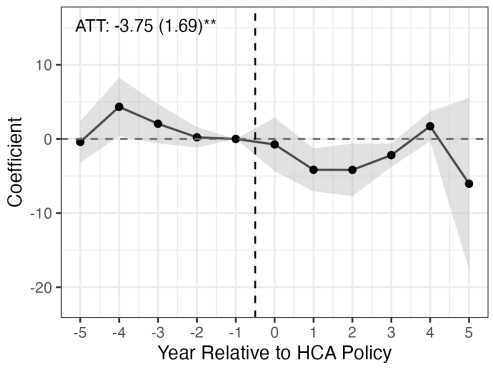
(A) Cancer: Provider Number of Beds



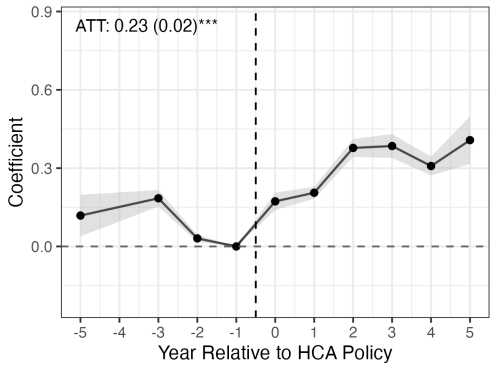
(B) Cancer: Any disease-specific drug within a year



(C) HIV/AIDS: Provider Number of Beds



(D) HIV/AIDS: Any disease-specific drug within a year



Note: Panels (A) and (B) report the effect of HCA policy on the quality of provider visited by cancer patients (see Figure 7 for details). Panels (C) and (D) report the effect of HCA policy on HIV patients (see Figure 5 for details). For each provider, we define whether it has an ICU by having any ICU claims in the sample period (see Table 2) and the number of inpatient beds from the *Registro Especial de Prestadores de Salud* file from the Ministry of Health and Social Protection. We then calculate for each patient the average measure across visited providers, weighted by the number of claims.

We investigate these mechanisms by examining care management practices at insurers'

networks of covered providers, which is the primary dimension of insurer competition in Colombia. Specifically, we use provider characteristics and a measure of wait time as outcomes to examine where patient care is being delivered and authorized by the insurer and how long it takes for the patient to receive it. For provider characteristics we use on the average number of beds during the sample period and for wait time we use an indicator for whether the patient received a disease-specific medication within the year of diagnosis.

For this analysis, we impose the additional sample restriction that patients remain enrolled with the same insurer three years after diagnosis and exclude patients who die during the sample period. This allows the patient to have sufficient interaction with their insurer's network of providers and avoids potential confounding from insurer switching.

Figure 6, Panel A shows that patients with targeted cancers visited providers with approximately 6.3 more beds on average than patients with untargeted cancers. The corresponding outcome for HIV/AIDS patients (Panel C) shows no change after the introduction of performance pay, perhaps because HIV/AIDS management relies more heavily on timely access to ART medications than on frequent contact with healthcare facilities (NCBI Advisory Council, 2021). Consistent with this, Panels B and D show that receipt of disease-specific treatment within the first year increased by 7 p.p. for targeted cancer patients and by 23 p.p. for HIV/AIDS patients. The cancer effect is similar in magnitude to the effect on receiving any disease-specific treatment within the full three-year post-diagnosis window, whereas the HIV/AIDS effect in the first year alone is twice as large as the effect across all three years combined, suggesting that the policy shortened wait times for medication dispensing among HIV/AIDS patients.

9 Discussion

Performance pay for health insurers has been studied mainly in the context of MA in the U.S., where research has found that payments tied to plan-level star quality ratings have not improved patient outcomes and, in some cases, have strengthened incentives for risk selection (Fioretti and Wang, 2023; Layton and Ryan, 2015). The Colombian

HCA targets insurer incentives through a similar logic but, as we show, the policy did generate meaningful improvements in disease management. One potential explanation for these opposing findings across the two countries is plan pricing: if Colombian insurers face stronger competitive pressure to attract patients because there are no premiums, performance pay might translate more readily into care improvements. However, both settings feature fixed or zero premiums, ruling out price competition as a driver of the difference.

Another likely explanation is how performance payments are designed in each setting. Proposition 4 shows that performance pay improves insurer effort when payments are tied to specific, verifiable quality metrics for well-defined health conditions, or those for which there is little scope for diagnosis discretion. The Colombian HCA satisfies these conditions since it targets disease-specific metrics for cancer and HIV/AIDS that are clinically precise and do not easily spill over onto other health conditions.

In contrast, MA star ratings are designed around population-level metrics. As shown in Appendix Table 5, these measures include blood pressure control, plan-level all-cause readmissions, transitions of care, and access to needed care, all of which apply broadly across the enrollee population rather than to specific high-burden diseases. The breadth of the performance pay policy in MA dilutes the marginal return to effort on any particular dimension, weakening the relation between insurer effort and measured quality, and leading to lower effort as our theoretical framework predicts. The Colombian HCA's focus on disease-specific thresholds may therefore be a key feature of its effectiveness.

10 Conclusions

This paper studies performance pay for health insurers, combining a theoretical framework with empirical evidence from Colombia. The theory characterizes how quality thresholds, the uncertainty of healthcare costs, and the level and uncertainty of per-capita payments jointly determine whether performance pay induces insurer effort. The theory predicts that insurer effort is decreasing in the uncertainty of costs and per-capita performance payments, but increasing in the ease of meeting the quality thresholds. We test

these predictions using individual-level claims data from Colombia between 2011 and 2023, where the government introduced performance pay for HIV/AIDS in 2016 and for cancer in 2019.

We find that performance pay improved disease management for both targeted conditions. For cancer, the policy increased healthcare spending by 7.4 million pesos per patient ($2,500), driven by a meaningful increase in the utilization of cancer-specific treatments such as chemotherapy, radiotherapy, and cancer-related surgeries. For HIV/AIDS, the policy improved treatment adherence, with a larger share of patients claiming ART medications and receiving preventive vaccinations after its introduction.

Heterogeneity analyses confirm all three theoretical predictions: treatment effects are 50% larger among insurers with a higher ex-ante likelihood of receiving payments, consistent with effort increasing in the ease of meeting quality thresholds; twice as large among insurers with below-median payment variance, consistent with effort decreasing in the uncertainty of being paid; and 50% larger for health conditions with below-median cost variance, consistent with risk-averse insurers concentrating effort on more homogeneous conditions. To achieve these gains, insurers directed cancer patients to larger, more specialized facilities within their networks and shortened wait times for medication dispensing for HIV/AIDS patients.

These findings have important implications for the design of performance pay in healthcare markets. Our results suggest that targeting payments to specific, verifiable conditions is a key determinant of the policy's effectiveness.

References

Abadie, A. and J. L'hour (2021): "A penalized synthetic control estimator for disaggregated data," *Journal of the American Statistical Association*, 116, 1817–1834.

Atkinson, A., S. Burgess, B. Croxson, P. Gregg, C. Propper, H. Slater, and D. Wilson (2009): "Evaluating the impact of performance-related pay for teachers in England," *Labour Economics*, 16, 251–261.

Barlevy, G. and D. Neal (2012): "Pay for percentile," *American Economic Review*, 102, 1805–1831.

Basinga, P., P. J. Gertler, A. Binagwaho, A. L. Soucat, J. Sturdy, and C. M. Vermeersch

(2011): "Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation," *The Lancet*, 377, 1421–1428.

Buitrago, G., P. Rodriguez-Lesmes, N. Serna, and M. Vera-Hernandez (2025): "Do Health Insurers Matter for Health Outcomes? Evidence from a Large Natural Experiment in Colombia," *Working Paper*.

Burke, R. E., F. Hutchins, J. Heintz, S. Appel, J. Norman, S. Patel, A. Gupta, L. Rose, and R. M. Werner (2025): "Skilled Nursing Facility Value-Based Purchasing Failed To Achieve Hospital Readmission Reductions And Other Targets: Article examines Medicare's Skilled Nursing Facility Value-Based Purchasing program," *Health Affairs*, 44, 722–730.

De Walque, D., P. J. Gertler, S. Bautista-Arredondo, A. Kwan, C. Vermeersch, J. de Dieu Bizimana, A. Binagwaho, and J. Condo (2015): "Using provider performance incentives to increase HIV testing and counseling services in Rwanda," *Journal of Health Economics*, 40, 1–9.

Eijkenaar, F. (2012): "Pay for performance in health care: an international overview of initiatives," *Medical Care Research and Review*, 69, 251–276.

Einav, L., A. Finkelstein, Y. Ji, and N. Mahoney (2022): "Voluntary Regulation: Evidence from Medicare Payment Reform," *The Quarterly Journal of Economics*, 137, 565–618.

Fioretti, M. and H. Wang (2023): "Performance Pay in Insurance Markets: Evidence from Medicare," *Review of Economics and Statistics*, 105, 1128–1144.

Goodman, S. F. and L. J. Turner (2010): "Teacher incentive pay and educational outcomes: Evidence from the New York City bonus program," *Journal of Labor Economics, Forthcoming*.

Griffith, R. and A. Neely (2009): "Performance pay and managerial experience in multitask teams: evidence from within a firm," *Journal of Labor Economics*, 27, 49–82.

Gupta, A. (2021): "Impacts of Performance Pay for Hospitals: The Readmissions Reduction Program," *American Economic Review*, 111, 1241–1283.

Gupta, A., G. David, and L. Kim (2023): "The Effect of Performance Pay Incentives on Market Frictions: Evidence from Medicare," *International Journal of Health Economics and Management*, 23, 27–57.

Holmström, B. (1979): "Moral Hazard and Observability," *The Bell journal of economics*, 74–91.

Holmstrom, B. and P. Milgrom (1991): "Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design," *The Journal of Law, Economics, and Organization*, 7, 24–52.

Jensen, M. C. and K. J. Murphy (1990): "Performance pay and top-management incentives," *Journal of Political Economy*, 98, 225–264.

Layton, T. J. and A. M. Ryan (2015): "Higher Incentive Payments in Medicare Advantage's Pay-for-Performance Program Did Not Improve Quality But Did Increase Plan Offerings," *Health Services Research*, 50, 1810–1828.

Markovitz, A. A., J. Z. Ayanian, D. Sukul, and A. M. Ryan (2021): "The Medicare Advantage Quality Bonus Program Has Not Improved Plan Quality," *Health Affairs*, 40, 1918–1925.

Miller, G. and K. Barbiarz (2013): "Pay-for-performance incentives in low-and middle-income country health programs," *Encyclopedia of Health Economics*.

Mullen, K. J., R. G. Frank, and M. B. Rosenthal (2010): "Can you get what you pay for? Pay-for-performance and the quality of healthcare providers," *The RAND Journal of Economics*, 41, 64–91.

NCBI Advisory Council (2021): "Guidelines for the Use of Antiretroviral Agents in Adults and Adolescents with HIV," *HHS Panel on Antiretroviral Guidelines for Adults and Adolescents*.

Riascos, A., E. Alfonso, and M. Romero (2014): "The Performance of Risk Adjustment Models in Colombian Competitive Health Insurance Market," *Documentos CEDE*.

Schneider, E., S. Whitmore, K. M. Glynn, K. Dominguez, A. Mitsch, M. T. McKenna, C. for Disease Control, P. (CDC), et al. (2008): "Revised surveillance case definitions for HIV infection among adults, adolescents, and children aged< 18 months and for HIV infection and AIDS among children aged 18 months to< 13 years–United States, 2008," *MMWR Recomm Rep*, 57, 1–12.

Serna, N. (2024): "Determinants of Provider Networks: Risk Selection vs. Administrative Costs," *Working paper*.

——— (2025): "Exogenous Exits, Market Structure, and Equilibrium Contracts in Health Care," *American Economic Review: Insights*.

Tebaldi, P., F. Paternollo, and A. Veiga (2024): "Designing Incentives for Multitasking Agents: Evidence from Payments to Physicians in England," *Working Paper*.

Vatter, B. (2025): "Quality disclosure and regulation: Scoring design in medicare advantage," *Econometrica*, 93, 959–1001.

Appendix 1 Data Cleaning Details

Cancer sample. The main cancer analysis uses a cross-sectional person-level dataset in which each observation is a cancer patient's first three years of care following diagnosis. The data sources are the *Base de Suficiencia* (individual health records) from 2013–2023, provided by each insurer. Insurers that appear in the health claims data are those that pass the Ministry of Health's data quality filters. Enrollment records supply demographics information such as age and sex.

We assign each person their earliest cancer diagnosis (if any) using three-digit ICD-10 codes. A cancer diagnosis is recognized only if the person accumulates more than one claim with that diagnosis code that is neither a prescription nor a consultation service, across the full sample period. See Appendix Table 1 for the ICD codes that define a cancer diagnosis.

The sample is then restricted to individuals with a single cancer diagnosis, non-missing sex and age at diagnosis, no concurrent HIV diagnosis, and an insurer whose claims passed the data quality filters in all three follow-up years. Individuals with zero total spending over the window are dropped (this dropped fewer than five observations). Insurer–diagnosis–cohort cells whose patient count or mean spending exceeds three times the average of all other years for that insurer–diagnosis combination are excluded (see Appendix Figure 8). Because the HCA policy took effect in 2019 and the follow-up window is two years, the first treated diagnosis cohort is 2017.

Outcome variables. Spending outcomes cover total spending and itemized categories (outpatient, prescriptions, surgery, chemotherapy, radiotherapy), all summed over the three-year follow-up window. We additionally construct hospital admission counts and average provider-level characteristics (number of beds and ICU availability) weighted by the number of claims. Provider level characteristics are obtained from December snapshots from 2011 to 2016 of the *Registro Especial de Prestadores de Salud* available at https://prestadores.minsalud.gov.co/habilitacion/.

HIV sample. The HIV analysis uses the same data source and diagnosis identification procedure, but differs in several respects. The outcome window covers only the diagnosis year itself rather than three subsequent years. Unlike cancer, the diagnosis threshold counts all claims bearing the HIV diagnosis code. Additional outcomes include ART uptake, vaccines, and HIV complications. Individuals with a concurrent cancer diagnosis are excluded, as are those ever enrolled in two insurers (EPS013 and EPS003) that were terminated in 2015 and 2019, respectively. The same outlier cell exclusion applies.

Appendix Table 1: ICD Code List for Treated and Control Diseases

Disease	ICD-3 Codes
Panel A: Cancer	
Treated	
Breast Cancer	C50
Cervical Cancer	C53
Colorectal Cancer	C18, C19, C20, C21
Gastric Cancer	C16
Lung Cancer	C33, C34
Lymphoma & Leukemia	C82, C83, C84, C85, C91, C92
Prostate Cancer	C61
Control	
Bone Cancer	C40, C41
CNS Cancer	C69, C70, C71, C72
Endocrine Cancer	C73, C74, C75
Oral & Pharyngeal Cancer	C00, C01, C02, C03, C04, C05, C06, C07, C08, C09, C10, C11, C12, C13, C14
Other Digestive Cancer	C15, C17, C22, C23, C24, C25, C26
Other Female Genital Cancer	C51, C52, C54, C55, C56, C57, C58
Other Hematologic Cancer	C81, C86, C87, C88, C89, C90, C93, C94, C95, C96
Other Male Genital Cancer	C60, C62, C63
Other Respiratory Cancer	C30, C31, C32, C35, C36, C37, C38, C39
Skin Cancer	C43, C44
Soft Tissue Cancer	C45, C46, C47, C48, C49
Urinary Tract Cancer	C64, C65, C66, C67, C68
Panel B: HIV/AIDS	
Treated	
HIV	B20, B21, B22, B23, B24
Control	
Asthma	J45, J46, Y55
Diabetes	E10, E11, E12, E13, E14, E15, G59, G63, G99, H28, H36, M14, N08, T38, Y42, Z13, Z83
Epilepsy	G40, G41, R56, T42, Y46
Hypertension	I10, I11, I12, I13, I15, I27, I67
Tuberculosis	A15, A16, A17, A18, A19, B90, J65, M01, N74, Z03, Z11, Z20, Z23

Note: This table present the ICD codes used to identify cancer and HIV patients in the main sample. ICD codes are shorten to the 3-digits level for brevity.

Appendix Table 2: Outcome Definitions for Cancer and Chronic Condition Patients

Variable	Category	Definition
Cancer Treatment	Chemotherapy	Induction chemotherapy; intrathecal chemotherapy; mono-chemotherapy (single cycle); poly-chemotherapy (low and high risk); intra-arterial and intracavitary administration.
Cancer Treatment	Radiotherapy	Teletherapy with radioisotopes, photons, or electrons; brachytherapy; conformal radiotherapy; stereotactic radiosurgery with linear accelerator.
Cancer Treatment	Surgery	Excision, resection, and ablation procedures from the surgical section of CUPS by cancer site: breast; male and female genital organs; digestive organs; respiratory tract; urinary tract, endocrine glands, hematologic system, CNS, skin, and oral/pharyngeal cavity. Excludes cosmetic, reconstructive, transplant, and non-oncologic procedures.
Admission	Admission	Low, medium, and high complexity inpatient stays; intensive care; intermediate care; partial hospitalization.
Provider with ICU	ICU	Neonatal, pediatric, and adult intensive care units (including burn units).
Vaccine	Vaccine	Prophylactic vaccinations against bacterial and viral diseases; antitoxin administration (tetanus, botulism, diphtheria).
Cancer Complication	Cancer Complication	Secondary and unspecified malignant neoplasm of lymph nodes, Secondary malignant neoplasm of respiratory and digestive organs, Secondary malignant neoplasm of other and unspecified sites (based on ICD codes).
Chronic Disease Drug	ART (HIV/AIDS)	Antiretroviral therapy: nucleoside and nucleotide reverse transcriptase inhibitors, non-nucleoside reverse transcriptase inhibitors, protease inhibitors, integrase inhibitors, and fixed-dose antiviral combinations (ATC J05AE, J05AF, J05AG, J05AJ, J05AR).
Chronic Disease Drug	Antimycobacterials (Tuberculosis)	Drugs for treatment of tuberculosis (ATC J04A).
Chronic Disease Drug	Antiepileptics (Epilepsy)	Antiepileptics (ATC N03A).
Chronic Disease Drug	Obstructive airway drugs (Asthma)	Drugs for obstructive airway diseases: inhaled and systemic adrenergics, glucocorticoids, anticholinergics, and other agents (ATC R03).
Chronic Disease Drug	Diabetes drugs (Diabetes)	Insulins and blood glucose lowering drugs (ATC A10).
Chronic Disease Drug	Antihypertensives (Hypertension)	Antihypertensives, diuretics, beta blocking agents, calcium channel blockers, and agents acting on the renin-angiotensin system (ATC C02, C03, C07, C08, C09).

Note: This table present the description of the service and ATC codes used to identify medical utilization outcomes in the main sample.

Appendix Table 3: ICD Codes and Services for HIV Complications and Vaccinations

Complication	ICD Codes
Bacterial infections, multiple or recurrent	A40, A41, G00, J13, J14, J15, J16, J17, J18, M00, M86
Candidiasis	B371, B378
Coccidioidomycosis	B38
Cryptococcosis, extrapulmonary	B45
Cryptosporidiosis, chronic intestinal	A072
Cytomegalovirus	B25
Encephalopathy, HIV related	B220
Herpes simplex: chronic ulcers, bronchitis, pneumonitis, or esophagitis	B00
Histoplasmosis, disseminated or extrapulmonary	B39
Isosporiasis, chronic intestinal	A073
Lymphoid interstitial pneumonia or pulmonary lymphoid hyperplasia complex	B221
Mycobacterium infections	A15, A16, A17, A18, A19, A31
Pneumocystis jirovecii pneumonia	B59
Progressive multifocal leukoencephalopathy	A812
Salmonella septicemia, recurrent	A021
Toxoplasmosis of brain	B58
Wasting syndrome attributed to HIV	B222

Note: Table reports the ICD code used to identify HIV Complications based on AIDS-defining conditions excluding cancers. Definition of AIDS is from CDC (Schneider et al., 2008).

Appendix 2 Proofs

Proposition 1. *The insurer's optimal effort $e^*_{k\theta}$ along dimension k for consumer type θ is:*

(i) Increasing in the per-capita performance payment p_θ

(ii) Increasing in the weight w_k placed on indicator k

(iii) Decreasing in the threshold I^*_k

(iv) Decreasing in the cost variance σ^2_θ

(v) Decreasing in the payment variance σ^2_B

Proof. Given separability, each first-order condition in (1) depends only on $e_{k\theta}$, so we apply the implicit function theorem and Topkis' theorem dimension by dimension. Taking the cross-partial of $FOC(e_{k\theta}, p_\theta)$ with respect to p_θ gives:

$$\frac{\partial\, FOC(e_{k\theta};\, p_\theta)}{\partial p_\theta} = \left(\tau_{k\theta} + \frac{\partial \tau_{k\theta}}{\partial e_{k\theta}} w_k s_\theta\right) \frac{\partial s_\theta}{\partial e_{k\theta}} \geq 0$$

establishing (i). The cross-partial with respect to w_k gives:

$$\frac{\partial\, FOC(e_{k\theta};\, p_\theta)}{\partial w_k} = \frac{\partial \tau_{k\theta}}{\partial e_{k\theta}}\, p_\theta\, s_\theta > 0$$

establishing (ii). The cross-partial with respect to I^*_k gives:

$$\frac{\partial\, FOC(e_{k\theta};\, p_\theta)}{\partial I^*_k} = \frac{\partial \tau_{k\theta}}{\partial I^*_k}\, w_k\, p_\theta\, s_\theta < 0$$

since $\partial \tau_{k\theta} / \partial I^*_k < 0$, establishing (iii). The cross-partial with respect to $-\sigma^2_\theta$ gives:

$$\frac{\partial\, FOC(e_{k\theta};\, p_\theta)}{\partial(-\sigma^2_\theta)} = \rho\, c_\theta\, \frac{\partial c_\theta}{\partial e_{k\theta}}\, s^2_\theta \geq 0$$

establishing (iv). Finally,the cross-partial with respect to $-\sigma^2_B$ gives:

$$\frac{\partial\, FOC(e_{k\theta};\, p_\theta)}{\partial(-\sigma^2_B)} = \rho s_\theta \frac{\partial s_\theta}{\partial e_{k\theta}} \geq 0$$

establishing (v). □

Proposition 3. *For any two types $\theta > \theta'$, performance payments $p_\theta > 0$ increase the effort gap $e^*_{k\theta} - e^*_{k\theta'}$ relative to the no-payment counterfactual. That is:*

$$\frac{\partial}{\partial p_\theta}\left(e^*_{k\theta} - e^*_{k\theta'}\right) > 0$$

Proof. Consider the first-order conditions for types θ and θ' separately. For the high type θ, the FOC from equation (1) is:

$$\begin{aligned}\left(R_\theta - c_\theta(\mathbf{e}_\theta) + \tau_\theta(\mathbf{e}_\theta, \mathbf{I}^*, \mathbf{w})\, p_\theta\right) \frac{\partial s_\theta}{\partial e_{k\theta}} - \left(\frac{\partial c_\theta}{\partial e_{k\theta}} - \frac{\partial \tau_{k\theta}}{\partial e_{k\theta}}\, w_k\, p_\theta\right) s_\theta(\mathbf{e}_\theta) \\ - \rho\left(\sigma^2_\theta\, c_\theta(\mathbf{e}_\theta)\, \frac{\partial c_\theta}{\partial e_{k\theta}}\, s^2_\theta(\mathbf{e}_\theta) + \sigma^2_B\, s_\theta(\mathbf{e}_\theta)\, \frac{\partial s_\theta}{\partial e_{k\theta}}\right) = 0\end{aligned}$$

For the low type θ', since $p_{\theta'} = 0$, the FOC reduces to:

$$\left(R_{\theta'} - c_{\theta'}(\mathbf{e}_{\theta'})\right) \frac{\partial s_{\theta'}}{\partial e_{k\theta'}} - \frac{\partial c_{\theta'}}{\partial e_{k\theta'}}\, s_{\theta'}(\mathbf{e}_{\theta'}) \quad - \rho\left(\sigma^2_{\theta'}\, c_{\theta'}(\mathbf{e}_{\theta'})\, \frac{\partial c_{\theta'}}{\partial e_{k\theta'}}\, s^2_{\theta'}(\mathbf{e}_{\theta'}) + \sigma^2_B\, s_{\theta'}(\mathbf{e}_{\theta'})\, \frac{\partial s_{\theta'}}{\partial e_{k\theta'}}\right) = 0$$

Since $p_{\theta'} = 0$, the first-order condition for type θ' does not depend on p_θ, and therefore:

$$\frac{\partial e^*_{k\theta'}}{\partial p_\theta} = 0$$

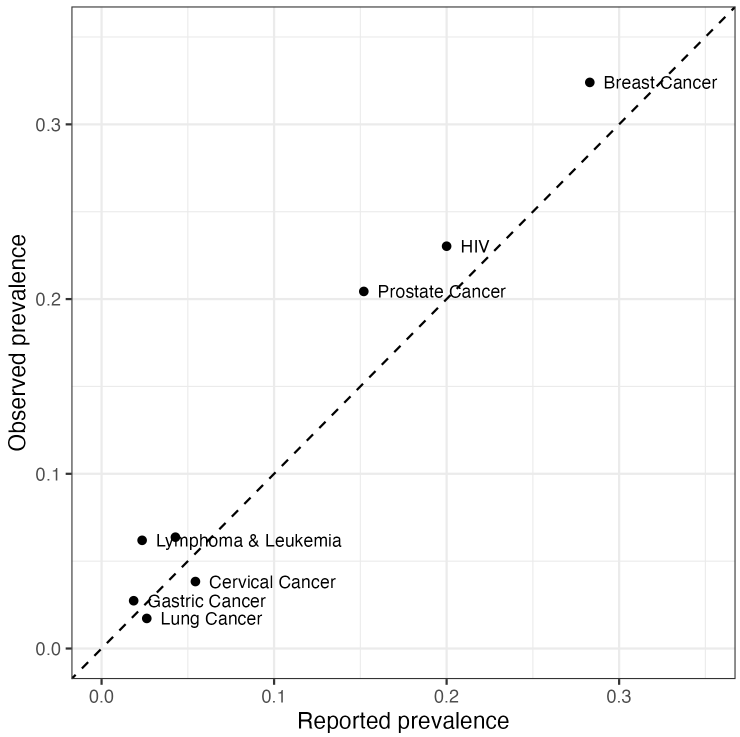
It follows that:

$$\frac{\partial}{\partial p_\theta}\left(e^*_{k\theta} - e^*_{k\theta'}\right) = \frac{\partial e^*_{k\theta}}{\partial p_\theta} - \frac{\partial e^*_{k\theta'}}{\partial p_\theta} = \frac{\partial e^*_{k\theta}}{\partial p_\theta}$$

It remains to show that $\partial e^*_{k\theta}/\partial p_\theta > 0$, which follows from proposition 1. □

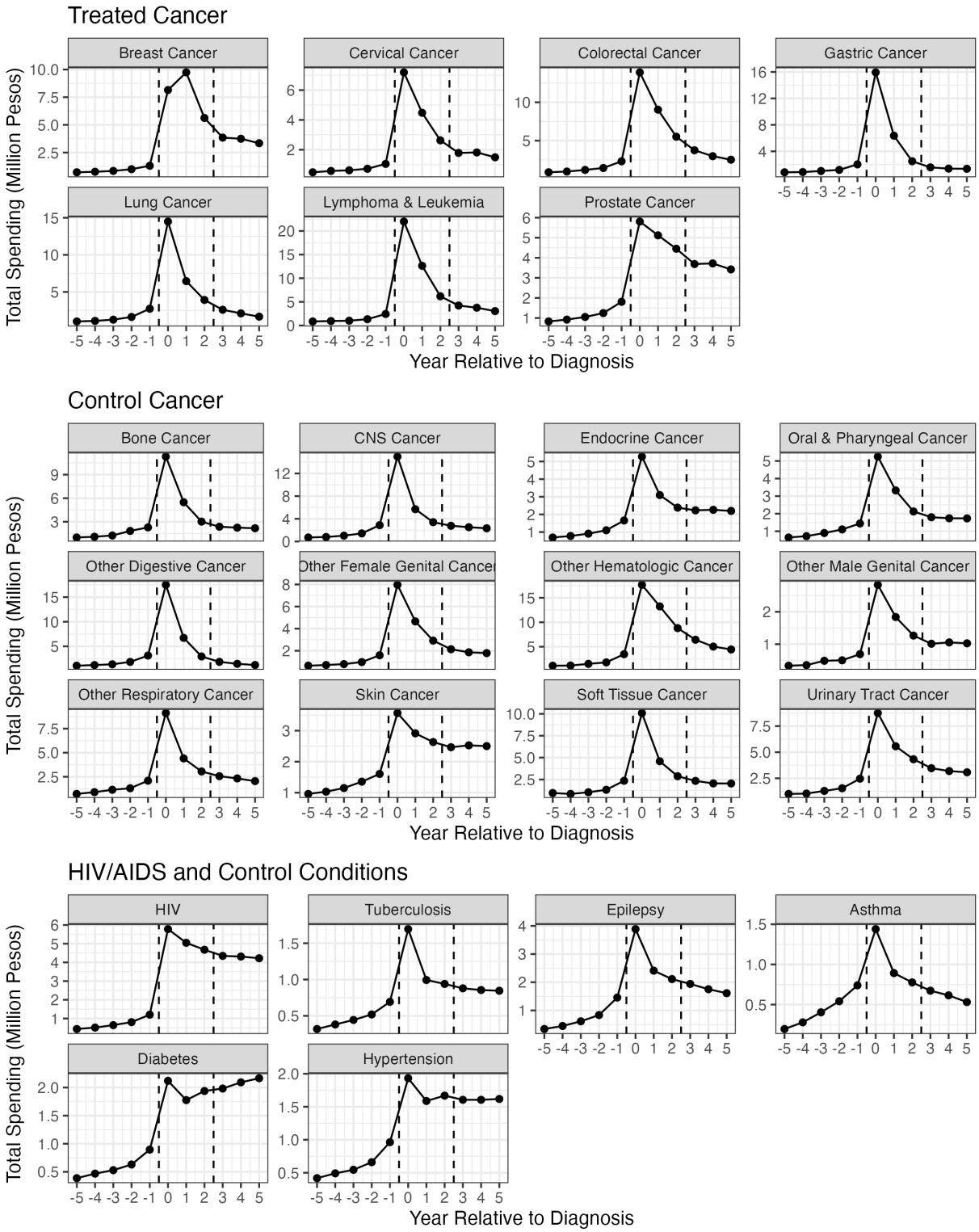
Appendix 3 Additional Descriptives and Robustness Checks

Appendix Figure 1: Comparison of Prevalence



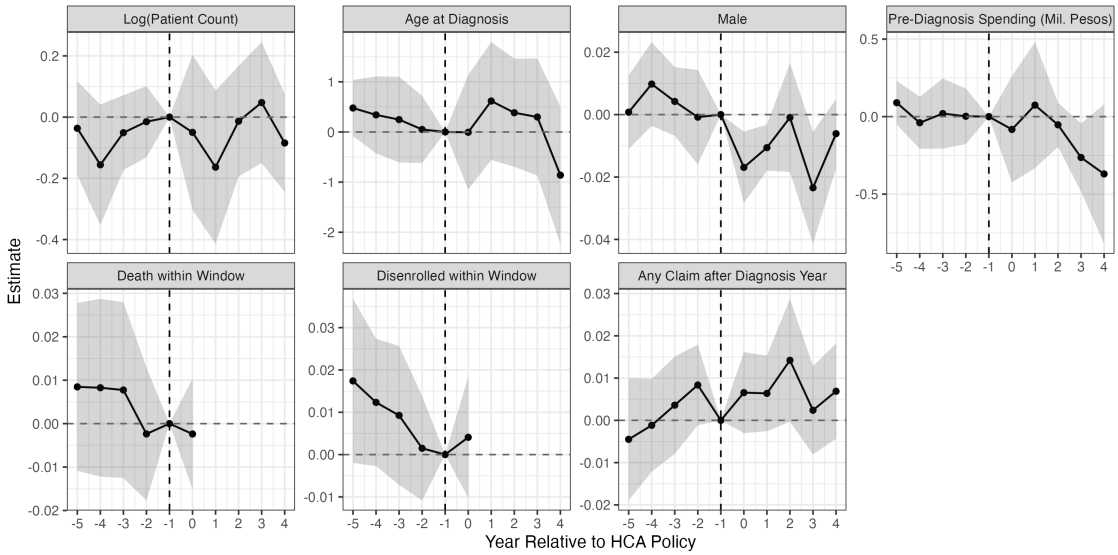
Note: Figure compares the prevalence observed in the sample against the official statistics reported by the HCA in 2016. Observed prevalence is defined as the ratio between patients with any diagnosing claims associated with the cancer, and the total number of enrollees in percentage points. Reported prevalence are obtained by official reports published by the HCA. Only prevalence of the treated diseases are reported.

Appendix Figure 2: Healthcare Spending by Year Relative to Diagnosis



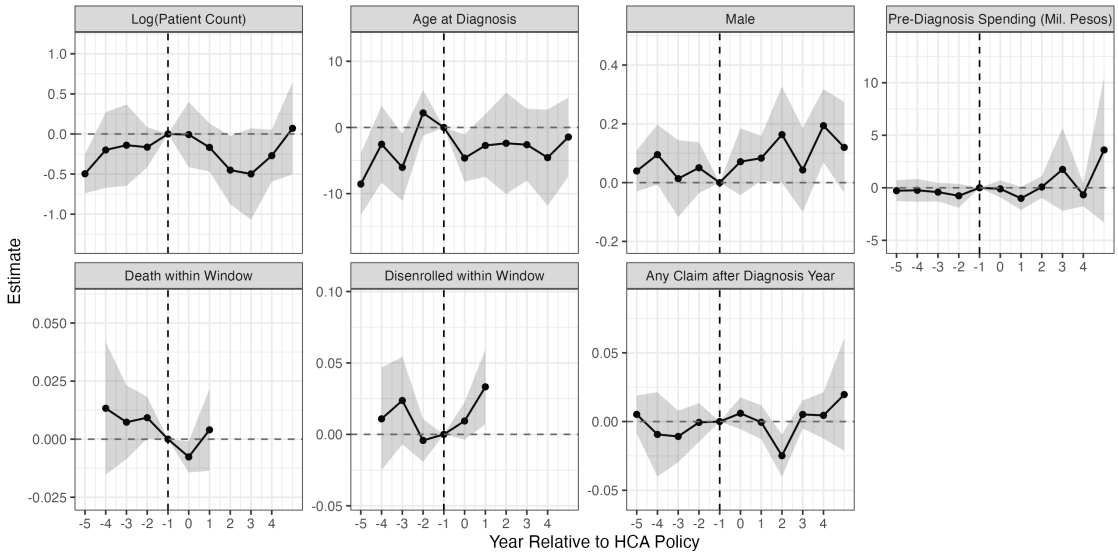
Note: Figure presents the trends of medical spending around diagnosis of cancer, HIV, and control conditions by disease. The dashed lines represent the aggregation window used in the analysis. Cancers affected by HCA are defined as treated and others as controls. See Appendix Table 1 for the ICD codes used.

Appendix Figure 3: Placebo Tests for Cancer HCA



Note: Figure reports the placebo results for the Cancer DiD design. Pre-Diagnosis Spending is defined as the total spending in the year before the diagnosis. For age, male, and pre-diagnosis spending, the results are obtained with the same specification as in Figure 3. Log(Patient Count) is constructed by aggregating the data to the cohort level, and running the regression with log of patient count as the outcome.

Appendix Figure 4: Placebo Tests for HIV/AIDS HCA



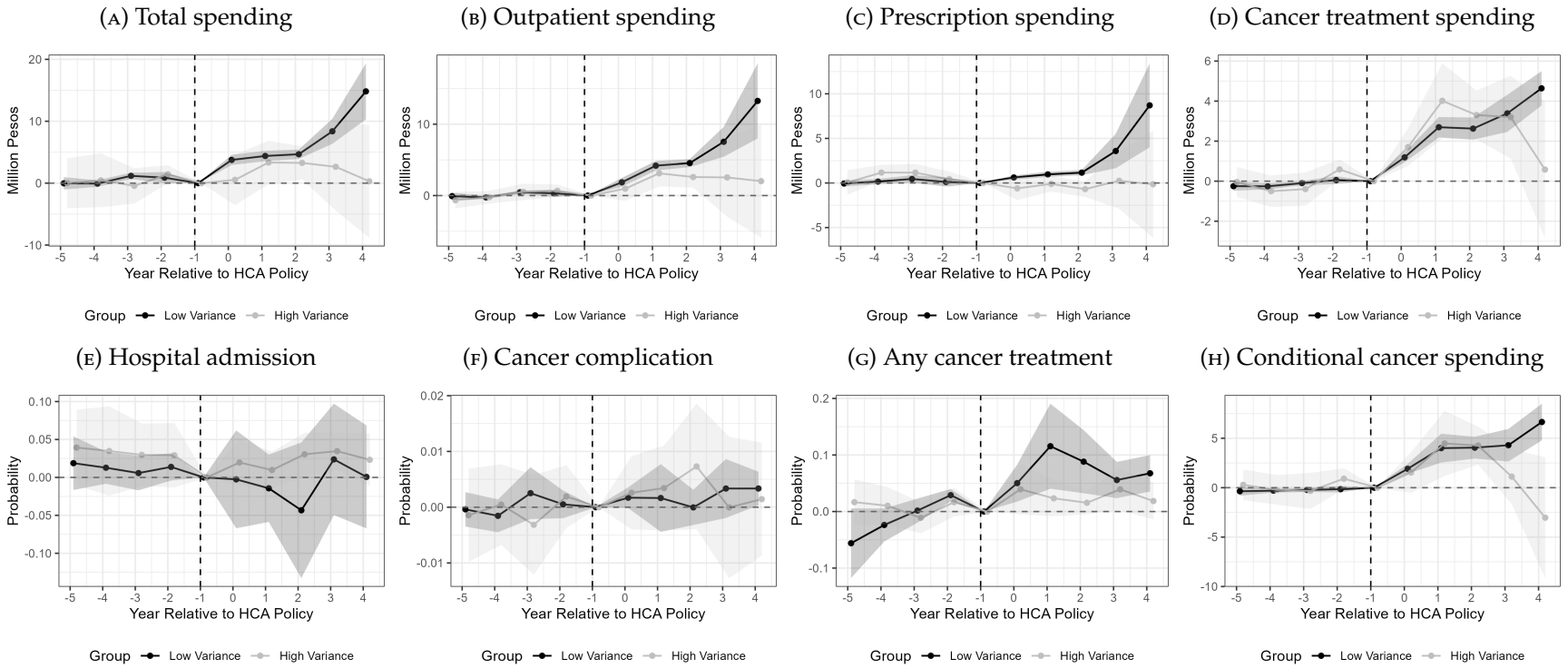
Note: Figure reports the placebo results for the HIV/AIDS DiD design. Pre-Diagnosis Spending is defined as the total spending in the year before the diagnosis. For age, male, and pre-diagnosis spending, the results are obtained with the same specification as in Figure 3. Log(Patient Count) is constructed by aggregating the data to the cohort level, and running the regression with log of patient count as the outcome.

Appendix Table 4: Summary Statistics of Treated and Control Diseases

Panel A: Cancer		
	Treated Cancer	Control Cancer
Male	0.407 (0.491)	0.404 (0.491)
Age at Diagnosis	58.50 (17.80)	53.90 (20.70)
Pre-Diagnosis Spending (Mil. Pesos)	1.302 (5.256)	1.437 (5.586)
Total Spending (Mil. Pesos)	18.10 (30.30)	11.70 (22.20)
Any Cancer Treatment	0.575 (0.494)	0.512 (0.500)
Outpatient Spending (Mil. Pesos)	9.105 (17.50)	4.824 (10.30)
Prescription Spending (Mil. Pesos)	5.835 (16.10)	2.711 (9.533)
Provider Herfindahl Index	0.437 (0.167)	0.453 (0.167)
Any Admission	0.100 (0.300)	0.080 (0.271)
Patients	138,652	154,615
Unique ICD Codes	16	66
Panel B: HIV/AIDS		
	HIV/AIDS	Synthetic Control
Male	0.746 (0.435)	0.508 (0.500)
Age at Diagnosis	36.60 (15.10)	26.50 (23.90)
Pre-Diagnosis Spending (Mil. Pesos)	1.030 (4.626)	0.869 (4.525)
Total Spending (Mil. Pesos)	12.90 (18.70)	5.490 (16.90)
Any Disease Drug	0.702 (0.458)	0.710 (0.454)
Outpatient Spending (Mil. Pesos)	7.663 (10.00)	2.201 (7.671)
Prescription Spending (Mil. Pesos)	3.061 (6.494)	1.138 (6.409)
Disease Drug Spending (Mil. Pesos)	1.149 (3.248)	0.141 (1.424)
Any Admission	0.304 (0.460)	0.327 (0.469)
Any Vaccine	0.126 (0.331)	0.080 (0.271)
Patients	8,475	86,428
Unique ICD Codes	5	32

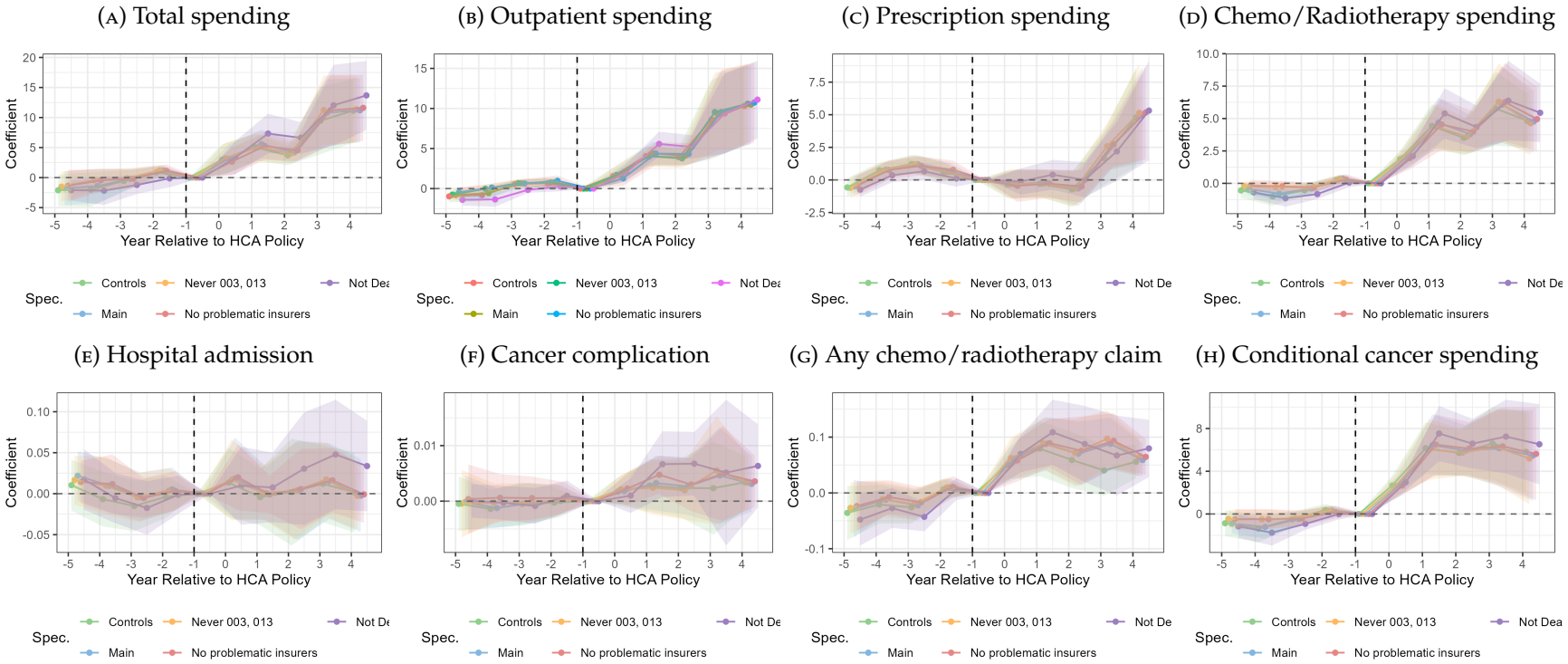
Note: Table presents mean and standard deviations in parenthesis of cancer patients diagnosed before the enactment of HCA policy. Cancers affected by HCA are defined as treated and others as controls. See Appendix Table 1 for the ICD codes used. 'Pre-Diagnosis Spending' measures the total spending in the year before the diagnosis. Utilization outcomes are measured in the year of the diagnosis and the two years after, where cancer-induced utilization are most concentrated.

Appendix Figure 5: Heterogeneity by Variance of Healthcare Spending



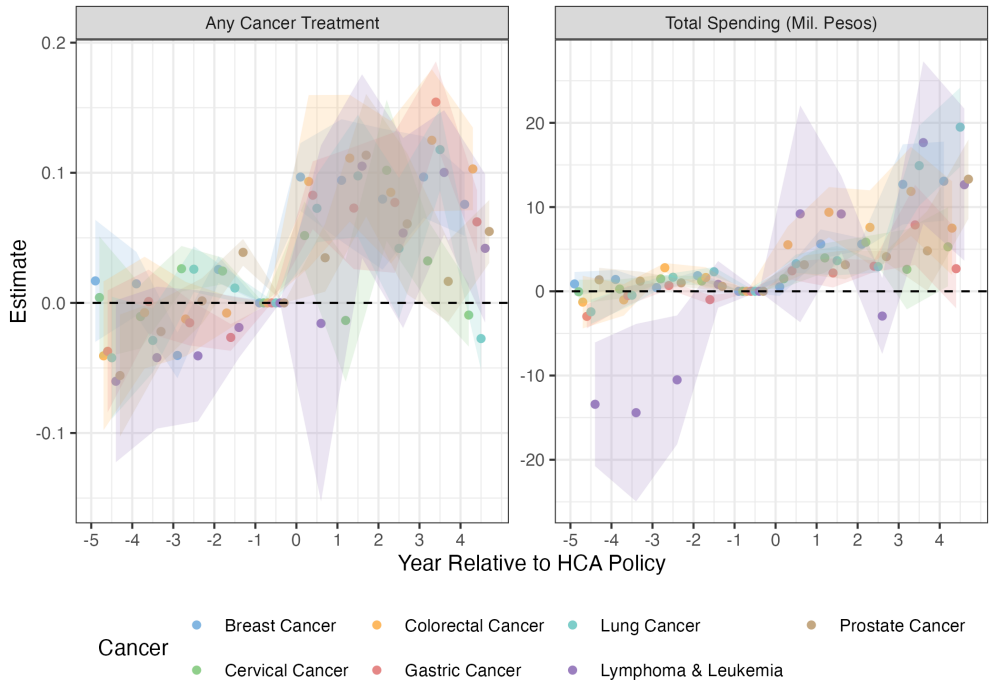
Note: Figure presents coefficients and 95% confidence intervals of the cohort DiD design for the effect of HCA policy on cancer patients (Equation 3) by variance of healthcare spending. For each cancer, we measure the variance of the healthcare costs among its patients in the pre-policy period. We then split the sample by whether the patient's cancer has a cost variance above the median, weighted by patient counts. Cancer complication is defined as having metastasis or other secondary cancer. Cancer treatment includes chemotherapy, radiotherapy, and cancer-related surgeries (see Appendix Table 2 for details). 'Cancer treatment spending + Any claims' measures the effect on treatment spending conditional on having any cancer treatment. Utilization outcomes are measured in the year of the diagnosis and the two years after, where cancer-induced utilization are most concentrated. The text on the top-left reports the coefficient and standard deviation in parentheses of the effect aggregated in the post-periods. Standard errors are clustered at the ICD-3 code level.

Appendix Figure 6: Robustness Checks on Cancer HCA Impact



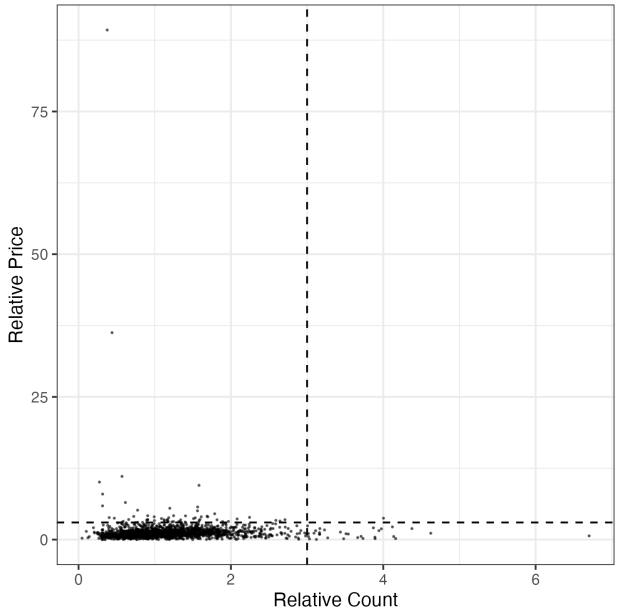
Note: Figure reports the robustness result from DiD design for cancer patients. 'Controls' controls for insurer-year and diagnosis-age-gender fixed effect. 'Never 003, 013' removes patient who ever enrolled with EPS003, EPS013, the two insurers terminated in 2016. 'Not Dead' removes patient who dies in the sample period. 'No problematic insurers' removes patient enrolled with EPS003, EPS013, EPS008, three insurers with a high chance of containing outliers. See note for Figure 3 for details.

Appendix Figure 7: Robustness Checks by Type of Cancer



Note: Figure reports the heterogeintiy result from DiD design for cancer patients by cancer type. We estimate the baseline DiD with the treatment indicator interacted with the type of cancer. See note for Figure 3 for details.

Appendix Figure 8: Outlier Cohorts in the Cancer Sample



Note: Figure reports the relative patient count and cohort cells by Insurer–diagnosis–cohort cells compared to the average of all other years for that insurer–diagnosis combination. The dashed line mark the ratio threshold where observations exceeding the 3 times the relative mean are excluded.

Appendix Table 5: Medicare Advantage Star Quality Rating Measure

	Measure
	Breast Cancer Screening
	Colorectal Cancer Screening
	Annual Flu Vaccine
	Improving or Maintaining Physical Health
	Improving or Maintaining Mental Health
	Monitoring Physical Activity
	Special Needs Plan (SNP) Care Management
	Care for Older Adults – Medication Review
	Care for Older Adults – Pain Assessment
	Osteoporosis Management in Women who had a Fracture
	Diabetes Care – Eye Exam
	Diabetes Care – Blood Sugar Controlled
	Kidney Health Evaluation for Patients with Diabetes
	Controlling Blood Pressure
	Reducing the Risk of Falling
	Improving Bladder Control
	Medication Reconciliation Post-Discharge
	Plan All-Cause Readmissions
	Statin Therapy for Patients with Cardiovascular Disease
	Transitions of Care
	Follow-up after ED Visit for People with Multiple High-Risk Chronic Conditions
	Getting Needed Care
	Getting Appointments and Care Quickly
	Customer Service
	Rating of Health Care Quality
	Rating of Health Plan
	Care Coordination
	Complaints about the Plan
	Members Choosing to Leave the Plan
	Health Plan Quality Improvement
	Plan Makes Timely Decisions about Appeals
	Reviewing Appeals Decisions
	Call Center – Foreign Language Interpreter and TTY Availability

Note: Table reports the quality measures for Star Ratings Part C in Medicare, accessed from https://www.cms.gov/files/document/2027-star-ratings-measures.pdf.