Health Insurance Markets with Endogenous Risks*

Vitor Luz Farinha [†] Álvaro Parra[‡]

January 10, 2023

[WORK IN PROGRESS]

Abstract

We study insurance markets where health treatments affect the evolution of individuals' health. Considering perfect competition as a benchmark, we analyze the impact of market power on insurance policies, treatment levels, and the distribution of health outcomes throughout the individual's life. The impact of market power on treatment levels and health outcomes is non-trivial and operates through two novel channels. First, in the presence of market power, firms internalize the impact of current treatment decisions on future rents to be extracted from the consumer. We refer to a market as *aligned* if rents extracted from healthy individuals are higher than those from unhealthy ones. Second, while the presence of market power leads to lower utility to consumers, its impact on the *health-premium* of consumers, i.e., the increase in their utility from being healthy, is ambiguous. The presence of market power leads to better treatment if the market is aligned and the consumer's health premium is higher. We also compare outcomes under competition and monopoly to those of autarky (no insurance) and find that the comparison of treatment levels across these environments is ambiguous.

Keywords: Health Insurance, Market Competition, Dynamic Contracting

*The authors thank workshop participants at UBC. All errors are our own. Vitor Luz Farinha and Álvaro Parra are supported in part by funding from the Social Sciences and Humanities Research Council of Canada.

[†]Vancouver School of Economics, University of British Columbia, 6000 Iona Dr, Vancouver, BC, V6T 1L4, Canada. vitor.farinhaluz@ubc.ca

[‡]Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, BC, V6T1Z2, Canada. alvaro.parra@sauder.ubc.ca

1 Introduction

One's choice of health insurance at a given moment not only affects current outcomes, such as financial security and well-being, but also has consequences for our future health through the treatment choices it induces. For instance, proper treatment of a broken leg may require surgery, immobilization, and physiotherapy. If a lower-cost improper treatment is pursued, such as simple immobilization, patients will have reduced mobility and an increased likelihood of future trauma. Crucially, an individual's ability to obtain proper treatment depends on their insurance coverage. Consequently, the impact of market conditions on offered insurance policies must indirectly affect treatment provision and the distribution of health outcomes in a market. Surprisingly, the long-term effects of endogenous treatment choices on individuals' health have been completely absent from the literature on insurance markets.

In this paper, we study how market competition affects the provision of health insurance when future health outcomes are endogenous to the level of treatment individuals undertake. In particular, we study equilibrium insurance and treatment choices throughout an individual's lifetime and their implied distribution over health outcomes. At the beginning of each period, each individual has a publicly observed health status. Each health status induces a probability distribution over health shocks (such as a broken leg). Following the realization of the health shock, the consumer undertakes a certain treatment level, which affects her current consumption level and the distribution over the next period's health status.

We characterize equilibrium contracts and treatment choices in three canonical cases: perfect competition, monopoly, and autarky. We find that different levels of competition generate different insurance contracts, affecting treatment choices and, consequently, the distribution of health outcomes throughout the individual's lifetime. Moreover, we flesh out the key economic forces behind the connection between competition and treatment provision.

In a competitive market, in each period, firms observe a particular consumer's health status and make insurance contract offers. We assume that treatment choices are contractible and the health shocks are observable; hence, an insurance contract can directly prescribe treatment levels conditional on the health shock realization. Due to the presence of competition, firms anticipate zero profits in future periods and the linkage across different periods, as well as the incentive for treatment provision, occurs solely through the consumer's participation constraint: a contract allowing better treatment is more attractive to the consumer as it increases the likelihood of being healthy in the future.

In the presence of market power, however, the monopolist insurer anticipates future rents from trading with the consumer in the future. Additionally, since the monopolist is able to fully extract rents from the consumer, her continuation utility, in each period, is identical to her continuation no-insurance, or autarky, utility which constitutes her outside option. When deciding treatment choices the monopolist considers both the individual value of treatment (higher probability of being health in the future) and also their prospects of future rents. From an insurer's perspective, a higher quality treatment can be understood as an investment in future (more profitable) health statuses. The monopolist has incentives to steer individuals to the most profitable health status. The most profitable health status, however, may not coincide with being healthy. We refer to a market as *aligned* when healthy consumers are more profitable. In aligned environments, the monopolist tends to induce better treatments. In contrast, monopolists induce worse treatments in misaligned situations, leading to worse health outcomes throughout the individual's life. One of our contributions is to highlight the role of market alignment in treatment provision and characterize when firms have aligned or misaligned incentives.

Our analysis also provides a deeper understanding of what drives individuals' treatment choices. From an individual's perspective, treatment is a sacrifice of today's consumption to improve the prospects of future (health) outcomes. The cost of treatment is thus inherently linked to the individual's income and consumption level. All else constant, a decreasing marginal utility of consumption (risk aversion) implies that poorer individuals chose lower treatment levels because of their higher opportunity costs of giving up consumption. This lower level of treatment implies that health shocks can act as poverty traps. Individuals receiving a shock involving an expensive minimal treatment choose to pursue lower quality treatments as the cost of treatment increases with their reduced consumption. In turn, lower treatment leads to worsts distribution of health statuses and poorer future health outcomes.

This work builds upon several strands of literature. In the health insurance literature, the work closest to us is Ghili *et al.* (forthcoming). They analyze competition through short- and long-term contracts in competitive insurance markets. We build upon their model by introducing the option of having different treatment levels affecting the evolution of the individual's health status. In addition, we also examine the role that market competition has in the contract's terms.

The work on competition in health markets is extensive, focusing on several aspects of competition, such as the choice of providers (Capps *et al.*, 2003) and insurance plans (Ho and Lee, 2017), negotiation between insurers and providers (Gowrisankaran *et al.*, 2015), competition and the quality of provisions (Cooper *et al.*, 2011; Gaynor *et al.*, 2013), the impact of moral hazard (Einav *et al.*, 2013), and adverse selection in health insurance markets (Handel, 2013). Handel and Ho (2021) presents an in-depth survey of the literature in the industrial organization of health markets.

This paper proceeds as follows. Section two introduces the most general model of health insurance and endogenous health risks. Section three presents and discusses our findings in a two-period model, with two health statuses and a single health shock. Section 4 generalizes our characterization into the general model. Finally, Section 5 concludes, setting guidelines for future work.

2 Model

Health status and care. Consider the health decisions of an individual living $t = 1, \ldots, T$ periods, receiving a deterministic per-period income y_t, and a stochastic health shock $m_t \in M$. The individual has a health status $s_t \in S$, evolving stochastically according to her health care decisions. The level of treatment $\alpha_t \in \mathbb{R}_+$ is decided after the health shock arrives. We interpret the shock m_t as the cost of the minimal medical treatment prescribed and α_t as any treatment undertaken beyond the minimal level of care. Given health status, s_t, the individual's health shock m_t distributes according to $G(\cdot \mid s_t)$. Next period's health status, s_{t+1}, distributes according to $F(\cdot \mid s_t, \alpha_t, m_t)$; capturing that the health status evolves depending on current status, health shock, and level of care.

Insurers and contracts. There are $j = 1, \ldots, J$ insurers offering one-period or long-term contracts. A contract specifies a *net* transfer $\tau(s, \alpha, m_t)$ to the individual as a function of her health status, health shock, and, when contractible, level of care. The net transfer accounts for the premium paid by the consumer and any reimbursements made by the insurance firm. For example, if the consumer paid premium ρ and had no claims, we have $\tau = -\rho$.

A one-period contract $\tau_t(s_t, \alpha_t, m_t)$ depends on the period's health status, shock, and care undertaken. A long-term contract potentially depends on the individual's medical *history*. For a given variable x_t, let $x_k^t = (x_k, \ldots, x_t)$ be its t period history. A long-term contract starting in period k can be represented by $\mathcal{T}_k = \{\tau_t\}_{t=k}^{T}$ with $\tau_t(s_k^t, \alpha_k^t, m_k^t)$ representing the transfer in period t which depends on the history of health statuses s_k^t, heath shocks m_k^t, and treatment α_k^t. We assume that insurers are committed to the long-term contracts they offer, but the individual may switch contracts at the beginning of every period.

Payoffs. The individual has a Bernoulli utility over net consumption, which is income minus health expenses, given by $u(\cdot)$ satisfying $u', -u'' > 0$. Future consumption is discounted at rate δ. For a given health shock m, health care decisions have a per-period cost $m + k(\alpha; m)$. The cost function k increases in each dimension and is strictly convex in treatment. We further assume $k(0;m) = k'(0;m) = 0$ and $\lim_{\alpha\to 1} k'(\alpha;m) = \infty$ for every m. The consumer's total utility is given by

$$\sum_{t=1}^{T} \delta^{t-1} u[y_t - m_t - k(\alpha_t; m_t) + \tau_t(s_t, \alpha_t, m_t)].$$

A firm's profit, if it stays with the same consumer for all periods, is $-\sum_{t=1}^{T} \delta^{t-1}\tau_t(s_t, m_t, \alpha_t)$.

Timing. At the beginning of period t, the individual has a health status s_t and, potentially, a contract in place. Each insurer observes the existing contract, and the status s_t, and unsigned insurers offer a contract $\tau_t^j(\cdot \mid s_t)$. After keeping or signing a new contract, the health shock m_t realizes, and the individual decides her treatment α_t, receiving the transfer $\tau_t(s_t, m_t, \alpha_t)$.

3 An Illustrative Example

To illustrate the main mechanisms at play, we study a model with $T = 2$ periods. In the first period, the individual has a known health status and, at $t = 2$, two potential health statuses $S = \{h, \ell\}$, representing a healthy and an unhealthy individual. The set of health shocks is given by $M = \{0, m\}$; that is, the individual may receive no health shock or a shock of size m. We also assume a linear cost for additional treatment, $k(\alpha, m) = \alpha$, and the probability of facing a negative shock, $G(m \mid s) = g(s)$, satisfies $g(h) < g(\ell)$, capturing that a healthier individual is less likely to receive a health shock.

Finally, the probability of being healthy next period depends on the health shock. When no health shock arrives, we assume that $F(h \mid \alpha, 0)$ is constant. Consequently, when no health shock arrives, treatment is not necessary as it does not provide health benefits. When the negative shock arrives, we denote $F(h \mid \alpha, m) = f(\alpha)$, and we assume this probability is increasing, concave, and differentiable in the additional treatment α. The contract characterization in this section follows the formal analysis in Section 4. We present the results delaying the formal analysis to that section.

3.1 Last Period

We solve the problem by backward induction, studying the optimal contract and healthcare choices in the last period and how those outcomes affect the contracting outcomes and healthcare choices in the first period. We study the optimal equilibrium outcomes in three different contracting environments: $\mathcal{C} = \{\mathtt{n}, \mathtt{c}, \mathtt{m}\}$, representing an individual with no insurance, facing a competitive market, or facing a monopolist insurer. For any contract $C \in \mathcal{C}$ and health status $s \in S$, let $V_t^{[C]}(s)$ be the period's t expected discounted utility that the individual obtains under contract C and health status s.

No insurance We start by characterizing the optimal healthcare choices without access to insurance. For ease in notation, we write $\hat{u}_t(\alpha_t; m_t) = u(y_t - m_t - \alpha_t)$ to represent the individual's utility at period t given the treatment choice α_t and shock m_t. For a given health status s, the value of not having insurance at $t = T$ is given by

$$V_2^{[\mathtt{n}]}(s) = \mathbb{E}_M \left[\max_{\alpha(m_2) \in A} \hat{u}_2(\alpha(m_2); m_2) \;\middle|\; s \right].$$

The choice of treatment occurs after the realization of the health shock. Because this is the last period, however, investing in treatment beyond the minimal treatment m_t increases the healthcare expenditure without receiving the benefit of better future outcomes. Consequently, the optimal treatment beyond the minimal treatment is zero; i.e., $\alpha_2^{[\mathtt{n}]}(m_2) = 0$. The expected utility in the last period is given by

$$V_2^{[\mathtt{n}]}(s) = \mathbb{E}_M[\hat{u}_t(0; m_t) \mid s].$$

Given our assumption that $g(\ell) > g(h)$, it can be readily verified that $V_2^{[\mathtt{n}]}(h) > V_2^{[\mathtt{n}]}(\ell)$; a healthy individual derives a larger expected payoff than an unhealthy one.

Competitive Insurance Market In a competitive market $J \geq 2$ firms compete by offering one-period contracts to the individual at periods $t = 1, 2$. In each period t, each firm observes the health status s_t and makes simultaneous offers. An offer constitutes a function $\tau(s, m, \alpha)$, which determines payment as a function of current period health status, health shock, and treatment.

We can simplify the description of an insurance contract significantly. First, when determining the contract to be offered to an agent with status s_t, we can omit the dependence of transfers on s_t since firms know this. Second, the assumption of contractable treatment implies that we can assume, without loss, that the insurance firm *directly chooses* treatment.[1] Hence we can refer to a one-period contract simply as $\mathcal{T} = (\alpha(m_t), \tau(m_t))$, containing shock-dependent treatment and a net transfer.

Insurers compete to attract consumers, leaving all the rents to the individual. This behavior

[1]Suppose a firm wants to implement treatment level α. This can be done by implementing large and negative transfers if the chosen treatment differs from the desired one.

implies that we can characterize the optimal competitive insurance by maximizing the individual's utility subject to a non-negative profit condition. Because treatment is contractible, insurers can pre-specify and induce the level of treatment of their choice.

For ease in notation, define $\tilde{u}_t(\alpha_t; m_t, \tau_t) = u(y_t - m_t - \alpha_t + \tau_t)$ to represent the individual's utility at period t given the treatment choice α_t, transfer from the insurance τ_t and realized shock m_t. For a given health status s, the expected utility for the optimal competitive insurance and treatment at $t = T$ is given by the solution to:

$$V_2^{[\mathsf{c}]}(s) = \mathbb{E}_M \left[\max_{(\alpha(m_2), \tau(m_2)) \in A \times \mathbb{R}} \tilde{u}_2(\alpha(m_2); m_2, \tau_2(m_2)) \;\middle|\; s \right],$$

subject to $-\mathbb{E}_M \left[\tau_2(m_2, \alpha(m_2)) \mid s\right] \geq 0$.

As before, the optimal choice of treatment beyond the minimum is zero regardless of the health shock, $\alpha_2^{[\mathsf{c}]}(m_2) = 0$. Following standard arguments, it is easy to show that the optimal contract involves full insurance at an actuarially fair rate, that is

$$c_2(s) = y_2 - g(s)m,$$

and $V_2^{[\mathsf{c}]}(s) = u\left(c_2(s)\right)$. As before, a healthy individual derives a larger expected payoff than an unhealthy individual, $V_2^{[\mathsf{c}]}(h) > V_2^{[\mathsf{c}]}(\ell)$.

Monopolized Insurance Market In a monopolized market, a single firm offers one-period contracts to the individual at periods $t = 1, 2$. As in competitive markets, the health status s_t is observed by the firm, and we can model treatment as directly chosen by the insurance firm. These observations allow us to write contracts as a shock-dependent treatment and a net transfer $\mathcal{T} = (\alpha(m_t), \tau(m_t))$.

A monopolist insurer extract rents from the individual. In particular, the monopolist extracts rents all the way up to the individual's outside option, i.e., their value without insurance, $V_2^{[\mathsf{n}]}(s)$. Given a health status s, the monopolist maximizes the net transfers obtained from the individual, subject to the individual's expected utility being higher than her outside option of not having insurance. In doing so, the insurer chooses the utility-maximizing treatment, extracting larger rents through the shock-dependent transfer. The monopolist's profit is given by the solution to

$$\pi(s) = \mathbb{E}_M \left[\max_{(\alpha(m_2), \tau(m_2)) \in A \times \mathbb{R}} -\tau_2(m_2) \;\middle|\; s \right]$$

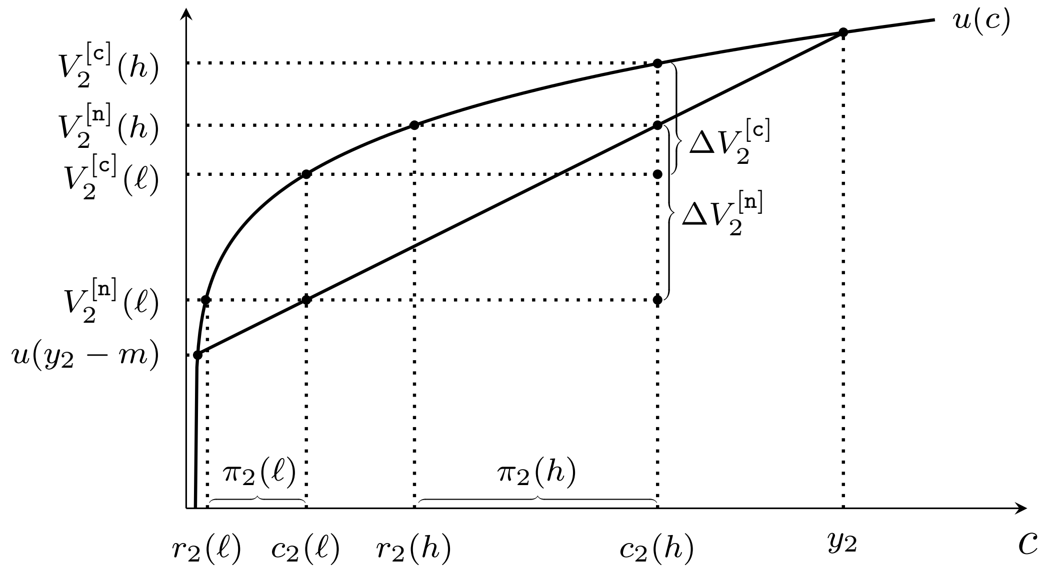

Figure 1: Last period ($T = 2$) outcomes. Depiction of market outcomes for an individual facing a monopolized market, a competitive market, or autarky. $V_2^{[C]}(s)$ represents the expected utility of the individual with health status $s \in \{h, \ell\}$ under a contract $C \in \{\mathtt{c}, \mathtt{n}, \mathtt{m}\}$. The consumption of an uninsured individual under a shock m_2 is represented by $y_2 - m_2$. Insured individuals are fully insured obtaining a consumption under health status s equal to $r_2(s)$ in a monopolized market and $c_2(s)$ in a competitive market.

subject to

$$V_2^{[\mathtt{m}]}(s) \equiv \mathbb{E}_M \left[\tilde{u}_2(\alpha(m_2); m_2, \tau_2(m_2)) \;\middle|\; s \right] \geq V_2^{[\mathtt{n}]}(s).$$

Using standard tools, we can show that three elements characterize the solution:

(i) For any realized health shock m_2, there is no treatment beyond the minimal treatment, $\alpha_2^{[\mathtt{m}]}(m_2) = 0$.

(ii) Full insurance, with a consumption level equal to the certain equivalent of the health shock:

$$r_2(s) \equiv u^{-1}(V_2^{[\mathtt{n}]}(s)),$$

(iii) Profits equal to the difference between the health shock's expected value and its certain equivalent

$$\pi(s) = c_2(s) - r_2(s)$$

Figure 1 depicts the equilibrium outcomes under each of the three contracts. In the absence of insurance and with a minimal treatment level, the health shocks lead to a consumption of y_2 when no shock arrives and $y_2 - m$ when it does. Different health statuses lead to different lotteries between these two consumption levels. In monopolized markets and markets in which individuals are uninsured, consumers obtain an expected payoff equal to $V_2^{[\mathtt{n}]}(s)$. Whereas under competition, the individual receives an expected payoff of $V_2^{[\mathtt{c}]}(s)$. When insured, the individual has constant consumption; i.e., consumption is invariant with respect to the health shock. Under a monopoly, the individual consumes $r_2(s)$ and, under competition, consumes an amount $c_2(s) > r_2(s)$. The monopolist's profits are represented by the gap $c_2(s) - r_2(s)$.

3.2 First Period

Now we turn to studying first-period decisions. Unlike the last period, when a health shock arrives, treatment beyond minimal will be optimal in equilibrium as a better treatment improves the prospects for future health outcomes. We show that an uninsured individual, a competitive firm, and a monopolist insurer will assess the attractiveness of future rents differently, leading to different treatment levels in equilibrium. In what follow, we denote by $\alpha_1^{[C]}(m_1)$ and $c_1^{[C]}(m_1)$ the optimal level of treatment and consumption under contract C and health shock m_1.

No insurance When deciding their optimal healthcare choices, an uninsured individual contemplates the impact that better treatment today will have on their future health. That is, the uninsured individual solves

$$V_1^{[\mathtt{n}]} = \mathbb{E}_M \left[\max_{\alpha(m_1)\in A} \left\{ \hat{u}_1(\alpha(m_1); m_1) + \delta \mathbb{E}_S \left[V_2^{[\mathtt{n}]}(s) \;\middle|\; m_1, \alpha(m_1) \right] \right\} \;\middle|\; s \right],$$

the individual considers the impact that better treatment has on their consumption today versus the prospects of better health outcomes tomorrow, which is captured by the discounted-expected continuation values.

As before, treatment choices occur after the realization of the negative shock. The optimal treatment depends on the realized m_1. When no health shock arrives, no further treatment is necessary; i.e., $\alpha_1^{[\mathtt{n}]}(0) = 0$. When a negative shock arrives, the optimal treatment is determined by the first-order condition of the problem, which is given by:

$$f'\left(\alpha_1^{[\mathtt{n}]}(m)\right) \delta \Delta V_2^{[\mathtt{n}]} = u_t'\left(c_1^{[\mathtt{n}]}(m)\right). \tag{1}$$

where

$$c_1^{[\mathtt{n}]}(m_1) = y_1 - m_1 - \alpha_1^{[\mathtt{n}]}(m_1)$$

is the equilibrium consumption under health shock m_1 and $\Delta V_2^{[\mathtt{n}]} = V_2^{[\mathtt{n}]}(h) - V_2^{[\mathtt{n}]}(\ell)$ is the incremental continuation value of having a better future health status.

The optimal treatment choice equates the marginal benefit of having better future outcomes with the marginal cost of higher healthcare expenditure today. The marginal benefit consists of the discounted incremental rent derived by a better health status $\delta\Delta V_2^{[\mathtt{n}]}$ weighted by the increase in the probability that this outcome occurs, $f'\left(\alpha_1^{[\mathtt{n}]}(m)\right)$. The marginal costs consist in the (marginal) utility lost today due to the increase in healthcare costs, $u_t'\left(c_1^{[\mathtt{n}]}(m)\right)$. Figure 1 illustrates the incremental rent derived by better health outcomes without insurance, $\Delta V_2^{[\mathtt{n}]}$.

Competitive Insurance Market As in the second period, competitive insurers leave all the rents to the individual. The competitive contract is characterized by the shock-dependent treatment and transfers maximizing the individual's expected-discounted utility subject to a non-negative profit condition. That is,

$$V_1^{[\mathtt{c}]} = \mathbb{E}_M \left[\max_{(\alpha(m_1),\tau(m_1))\in A\times\mathbb{R}} \left\{ \tilde{u}_1(\alpha(m_1); m_1, \tau_1(m_1)) + \delta \mathbb{E}_S \left[V_2^{[\mathtt{c}]}(s) \;\middle|\; m_1, \alpha(m_1) \right] \right\} \;\middle|\; s \right],$$

subject to $-\mathbb{E}_M\left[\tau_1(m_1, \alpha(m_1)) \mid s\right] \geq 0$.

The optimal contract involves full insurance at actuarially fair rates. The consumption is given by

$$c_1 = y_1 - g\left(\alpha_1^{[\mathtt{c}]}(m) + m\right),$$

where g is the probability of receiving a negative health shock in the first period. Full insurance occurs, as insurers have better risk tolerance than individuals. In their pressure to compete for consumers, they absorb the risks and give the rent to the individuals.

In the absence of a health shock, no treatment is necessary. When a health shock arrives, the optimal treatment is characterized by

$$f'\left(\alpha_1^{[\mathtt{c}]}(m)\right)\delta\Delta V_2^{[\mathtt{c}]} = u_t'\left(c_1^{[\mathtt{c}]}(m)\right), \tag{2}$$

where $\Delta V_2^{[\mathtt{c}]} = V_2^{[\mathtt{c}]}(h) - V_2^{[\mathtt{c}]}(\ell)$ is the incremental continuation value of having a better future health status insured by competitive firms. As with an uninsured individual, the optimal treatment choice equates the marginal benefits of having better future outcomes with the marginal cost of higher healthcare expenditure today. The main differences are that the marginal benefit considers continuation values under competition instead and that the marginal cost incorporates the transfer from the insurance company. Figure 1 visually illustrates the incremental rent derived by better health outcomes under competition, $\Delta V_2^{[\mathtt{c}]}$.

Comparing equations (1) and (2), we can identify two drivers of treatment decisions. On the costs side, different insurance contracts (or lack thereof) will lead to a different consumption profile when the negative shock arrives. Different consumption levels lead to different marginal costs of investing in better treatment. All else equal, the concavity of the utility function (risk aversion), leads lower-income individuals to invest less in healthcare, as they face a higher cost of giving up consumption today. When health risks are endogenous, health shocks act as poverty traps. Poorer people invest less in healthcare, obtaining worse health outcomes in the future and lower future wealth.

On the marginal benefit side, different contracts lead to different returns of investing in future health. The incremental rent of obtaining a better health status under contract C is given by $\Delta V_2^{[C]}$. The incremental rent may have different orderings depending on the distribution of health shocks in each status. Figure 1 illustrates a situation in which the incremental rent obtained by an uninsured person is larger than the one obtained under a competitive market, i.e., $\Delta V_2^{[\mathtt{n}]} > \Delta V_2^{[\mathtt{c}]}$. This ordering, of course, might reverse. We characterize the shocks and their induced ordering in the next section.

Monopolized Insurance Market As in the last period, monopolist insurers can extract rents up to the individual's outside option, which now considers continuation values with the possibility of being insured or uninsured in the future. Because the continuation values of being uninsured and facing a monopolist next period are the same, there is no need to distinguish their continuation values.

We can write the monopolist problem as maximizing the net transfers obtained from the individual, subject to the individual's expected utility being higher than her outside option of not having insurance. The monopolist's profits are given by the solution to:

$$\pi_1 = \mathbb{E}_M \left[\max_{(\alpha,\tau)\in A\times\mathbb{R}} -\tau_1(m_1) + \delta \mathbb{E}_S \left[\pi_2(s) \;\middle|\; \alpha(m_1), m_1 \right] \;\middle|\; s \right]$$

subject to

$$V_1^{[\mathtt{m}]} \equiv \mathbb{E}_M \left[\tilde{u}_1(\alpha(m_1); m_1, \tau_1(m_1)) + \delta \mathbb{E}_S \left[V_2^{[\mathtt{n}]}(s) \;\middle|\; \alpha(m_1), m_1 \right] \;\middle|\; s \right] \geq V_1^{[\mathtt{n}]}.$$

We can show that three elements characterize the solution:

(i) No treatment when no shock arrives, $\alpha_1^{[\mathtt{m}]}(0) = 0$. When the negative shock realizes, the optimal treatment is given by the solution to the following first-order condition:

$$f'\left(\alpha_1^{[\mathtt{m}]}(m)\right) \delta \left[\Delta V_2^{[\mathtt{n}]} + u_t'\left(c_1^{[\mathtt{m}]}\right) \Delta\pi_2 \right] = u_t'\left(c_1^{[\mathtt{m}]}\right), \tag{3}$$

where $\Delta\pi_2 = \pi_2(h) - \pi_2(\ell)$ is the incremental continuation rent the monopolist derives from facing a healthier individual in the future.

(ii) Full insurance, with a consumption level equal to:

$$c_1^{[\mathtt{m}]} \equiv u^{-1}\left(V_1^{[\mathtt{n}]} - \delta \mathbb{E}_M\left[\mathbb{E}_S\left[V_2^{[\mathtt{n}]}(s) \;\middle|\; \alpha_1^{[\mathtt{m}]}(m_1), m_1\right]\right]\right),$$

(iii) Profits equal to the difference between the health shock's expected value and the consumption level given to the individual

$$\pi_1 = y_1 - g\left(\alpha_1^{[\mathtt{m}]}(m) + m\right) - c_1^{[\mathtt{m}]}$$

Similar to competitive insurers, monopolist insurers provide full insurance. They do so because insurers are more tolerant to risk. The monopolist prefers to absorb the risk of the health shock and then extract, via lower transfer to the consumers, the generated rents.

Although the treatment choices induced by a monopolist (see equation 3) are similar in structure to those for uninsured individuals and individuals in competitive markets (equations 1 and 2), they follow a different set of incentives. Monopolists perceive rents in future insurance markets and are incentivized to steer individuals toward the monopolists' preferred health status. Better treatment increases the probability of a healthier individual by $f'\left(\alpha_1^{[\mathtt{m}]}(m)\right)$ leading to a change in future rents of $\delta\Delta\pi_2$. Their incentive to steer individuals is captured by $\Delta\pi_2$, which might be positive or negative depending on the nature of the health shock. We characterize this object in the next section. At the same time, the monopolist is constrained in its rent-extraction process by the individual's outside option. When choosing a treatment level, the monopolist needs to consider the impact treatment has on the consumer's utility and the slackness of the constraints. Better treatment has the costs of a marginal decrease in today's consumption, $u_t'\left(c_1^{[\mathtt{m}]}\right)$, and the benefit of improved health outcomes $\delta\Delta V_2^{[\mathtt{n}]}$ which arrive at a rate $f'\left(\alpha_1^{[\mathtt{m}]}(m)\right)$. While the constraint is expressed in utils, the monopolists' objective function is expressed in money. Equation (3) translates these two objects, using the Lagrange multiplier $\lambda^{-1} = u'(c_1^{[\mathtt{m}]})$.

Finally, the monopoly extracts rents by giving the individual a consumption level equal to the certain equivalent of period's transaction. That is, the individual receives consumption equal to the outside option, $V_1^{[\mathtt{n}]}$, net of change in future utility due to change in treatment decisions $\delta \mathbb{E}_M\left[\mathbb{E}_S\left[V_2^{[\mathtt{n}]}(s) \;\middle|\; \alpha_1^{[\mathtt{m}]}(m_1), m_1\right]\right]$. The monopoly profits are equal to the expected value of the lottery they induce with their treatment choices, $y_1 - g\left(\alpha_1^{[\mathtt{m}]}(m) + m\right)$, net of the consumption given to individuals, $c_1^{[\mathtt{m}]}$.

4 Competition and Insurance Markets

We now turn into a general characterization of insurance contracts in different competitive environments. To simplify the characterization of the contracts, we keep the assumption that there are two health statuses $S = \{h, \ell\}$, representing a healthy and an unhealthy individual. We denote by $F(h \mid s_t, \alpha_t, m_t) = f(s_t, \alpha_t, m_t)$ the probability of being healthy next period given the current health status, s_t, treatment choice, α_t, and health shock m_t. We denote the derivative of $f(\cdot)$ with respect to α by $f_\alpha(\cdot)$.

4.1 No Insurance

We now characterize the optimal healthcare choices of uninsured individuals. We do this recursively, accounting for the impact of current decisions on the evolution of future health outcomes. For a given health status s, the value of not having insurance at $t \leq T$ is given by

$$V_t^{[\mathtt{n}]}(s) \equiv \mathbb{E}_M \left[\max_{\alpha(m_t) \in A} \left\{ \hat{u}_t\left(\alpha(m_t); m_t\right) + \delta \mathbb{E}_S \left[V_{t+1}^{[\mathtt{n}]}\left(s_{t+1}\right) \mid s, m_t, \alpha(m_t) \right] \right\} \;\middle|\; s \right] \tag{4}$$

where $V_{T+1}^{[\mathtt{n}]} = 0$. As in the example, the individual decides her optimal healthcare level conditional on the health shock's realization. In doing so, she foresees her treatment choices' impact on her future health outcomes.

Lemma 1. *The optimal healthcare for an uninsured individual is given by the minimal treatment at the last period, $\alpha_T^{[\mathtt{n}]}(m_t) = 0$. At any period $t < T$, for a given health shock m_t, the optimal level of treatment $\alpha_t^{[\mathtt{n}]}(m_t)$ is given by the solution to:*

$$f_\alpha\left(s, m_t, \alpha_t^{[\mathtt{n}]}(m_t)\right) \delta \Delta V_{t+1}^{[\mathtt{n}]} = u_t'\left(c_t^{[\mathtt{n}]}(m_t)\right) k'\left(\alpha_t^{[\mathtt{n}]}(m_t); m_t\right). \tag{5}$$

where $\Delta V_{t+1}^{[\mathtt{n}]} = V_{t+1}^{[\mathtt{n}]}(h) - V_{t+1}^{[\mathtt{n}]}(\ell)$ is the incremental rent of being healthier next period, and

$$c_t^{[\mathtt{n}]}(m_t) = y_t - m_t - k(\alpha_t^{[\mathtt{n}]}(m_t), m_t) \tag{6}$$

is the consumption when shock m_t realizes under the optimal treatment $\alpha_t^{[\mathtt{n}]}(m_t)$.

Uninsured consumers choose their treatment based on the marginal costs of giving up consumption today versus the benefit of improved health outcomes in the future. The main difference with respect to Section's 3 example is that treatment costs are shock-dependent, affecting the marginal impact on consumption by a factor of $k'\left(\alpha_t^{[\mathtt{n}]}(m_t); m_t\right)$. Similarly, on the marginal benefit side, the increase in future health outcomes now takes into consideration that the transition probability $f_\alpha(\cdot)$ is affected by the current health status, s, the realized shock m_t, and the choice of treatment $\alpha_t^{[\mathtt{n}]}(m_t)$.

4.2 Competitive Market

In a competitive market, $J \geq 2$ firms compete by offering one-period contracts to the consumer at every period $t \leq T$. In each period t, every firm observes the current health status s_t, making simultaneous offers. A contract specifies shock-dependent treatment and a net transfer $\mathcal{T}_t = (\tau_t(m_t), \alpha_t(m_t))$. The individual observes the offered contracts and chooses its best option.

A symmetric equilibrium of this one-period game corresponds to a contract $\mathcal{T}^{[\mathtt{c}]}$ and an acceptance rule for the consumer such that: the consumer's acceptance rule is sequentially rational and offering contract $\mathcal{T}^{[\mathtt{c}]}$ is optimal given the consumer's acceptance rule and that other firms are offering $\mathcal{T}^{[\mathtt{c}]}$.

Given information symmetry, standard arguments imply that $\mathcal{T}^{[\mathtt{c}]}$ must maximize the consumer's utility subject to a zero profit condition, i.e., it is the solution to:

$$V_t^{[\mathtt{c}]}(s) \equiv \mathbb{E}_M \left[\max_{\substack{\alpha(m_t)\in A,\\ \tau(m_t)\in\mathbb{R}}} \left\{ \tilde{u}_t\left(\alpha(m_t); m_t, \tau(m_t)\right) + \delta \mathbb{E}_S \left[V_{t+1}^{[\mathtt{c}]}\left(s_{t+1}\right) \mid s, m_t, \alpha(m_t) \right] \right\} \;\middle|\; s \right], \quad (7)$$

subject to

$$-\mathbb{E}_M \left[\tau(m_t) \mid s\right] \geq 0$$

Lemma 2. *In a competitive health insurance market the optimal contract $\mathcal{T}^{[\mathtt{c}]}$ satisfies:*

(i) The contract offers full insurance at actuarially fair rates, i.e., there exists constant $c_t^{[\mathtt{c}]}$ such that, for any realization of the health shock m_t, the consumption satisfies:

$$\begin{aligned} c_t^{[\mathtt{c}]}(s) &= \mathbb{E}_M \left[y_t - m_t - k(\alpha_t^{[\mathtt{c}]}(m_t), m_t) \;\middle|\; s \right] \\ &= y_t - m_t - k(\alpha_t^{[\mathtt{c}]}(m_t), m_t) + \tau(\alpha_t^{[\mathtt{c}]}(m_t), m_t) \end{aligned} \quad (8)$$

(ii) The contract provides the minimal treatment at the last period, $\alpha_T^{[\mathtt{c}]}(m_t) = 0$ for every m_t and, at any period $t < T$, for a given health shock m_t, provides the level of treatment $\alpha_t^{[\mathtt{c}]}(m_t)$ which is given by the solution to:

$$f_\alpha(s, m_t, \alpha_t^{[\mathtt{c}]}(m_t)) \delta \Delta V_t^{[\mathtt{c}]} = u'(c_t^{[\mathtt{c}]}(s)) k'(\alpha_t^{[\mathtt{c}]}(m_t), m_t), \quad (9)$$

where $\Delta V_t^{[c]} = V_2^{[c]}(h) - V_2^{[c]}(\ell)$ *is the incremental rent of being healthier next period.*

4.3 Monopolized Markets

In a monopolized market, a single firm offers one-period contracts to the consumer at every period $t \leq T$. In each period t, the monopolist observes the current health status s, and makes a take-it-or-leave-it offer to the individual. A contract specifies shock-dependent treatment and a net transfer $\mathcal{T}_t = (\tau_t(m_t), \alpha_t(m_t))$ that are contingent on the health status. The individual observes the offered contract and chooses its best option. An equilibrium of this one-period game corresponds to a contract $\mathcal{T}^{[m]}$ and an acceptance rule for the consumer such that: the consumer's acceptance rule is sequentially rational and offering contract $\mathcal{T}^{[m]}$ is optimal given the consumer's acceptance rule.

The monopolist insurer can extract rents up to the individual's outside option. That is, given a health status s, the monopolist maximizes the net transfers obtained from the individual, subject to the individual's expected utility being higher than her outside option of not having insurance. In doing so, the insurer chooses the utility-maximizing treatment to be able to obtain more rent. The monopolist's profits are given by the solution to

$$\pi_t(s) = \mathbb{E}_M \left[\max_{\substack{\alpha(m_t) \in A, \\ \tau(m_t) \in \mathbb{R}}} \left\{ -\tau(m_t) + \delta \mathbb{E}_S \left[\pi_{t+1}(s_{t+1}) \;\middle|\; s, \alpha(m_t), m_t \right] \right\} \;\middle|\; s \right]$$

subject to

$$V_t^{[m]}(s) \equiv \mathbb{E}_M \left[\tilde{u}_t(\alpha(m_t); m_t, \tau(m_t)) + \delta \mathbb{E}_S \left[V_{t+1}^{[n]}(s_{t+1}) \;\middle|\; s, \alpha(m_t), m_t \right] \;\middle|\; s \right] \geq V_t^{[n]}(s).$$

Lemma 3. *In a monopolized health insurance market the optimal contract* $\mathcal{T}^{[m]}$ *satisfies:*

(i) The contract offers full insurance at the certain equivalent of the incremental insurance value offered by the monopolists, i.e., there exists constant $c_t^{[m]}$ *such that, for any realization of the health shock* m_t*, the consumption satisfies:*

$$c_t^{[m]}(s) = u^{-1} \left(V_t^{[n]}(s) - \delta \mathbb{E}_M \left[\mathbb{E}_S \left[V_{t+1}^{[n]}(s_{t+1}) \;\middle|\; s, \alpha_t^{[m]}(m_t), m_t \right] \;\middle|\; s \right] \right), \tag{10}$$

(ii) The contract provides the minimal treatment at the last period, $\alpha_T^{[m]}(m_t) = 0$ *for every* m_t *and, at any period* $t < T$*, for a given health shock* m_t*, provides the level of treatment*

$\alpha_t^{[\mathtt{m}]}(m_t)$ *which is given by the solution to:*

$$f_\alpha(s, m_t, \alpha_t^{[\mathtt{m}]}(m_t))\delta \left[u'(c_t^{[\mathtt{m}]}(s))\Delta\pi_t + \Delta V_t^{[\mathtt{n}]}\right] = u'(c_t^{[\mathtt{m}]}(s))k'(\alpha_t^{[\mathtt{m}]}(m_t), m_t), \tag{11}$$

where $\Delta\pi_t = \pi_t(h) - \pi_t(\ell)$ *is the incremental profit the monopolist derives from healthier individuals.*

(iii) Profits equal to the difference between the health shock's expected value and the consumption level given to the individual

$$\pi_t(s) = \mathbb{E}_M \left[y_1 - k(\alpha_t^{[\mathtt{m}]}(m_t), m_t) - m_t \;\middle|\; s\right] - c_t^{[\mathtt{m}]}(s) \tag{12}$$

5 Conclusions

In this article, we explored how competition shapes health-insurance contracts when the risks of future health shocks are endogenous to the treatment choices induced by the contract. We found that market competition is a key factor determining the type of insurance contracts observed in equilibrium. Insurers are incentivized to steer consumers to the most profitable health status from the insurers' perspective. This status might or might not be aligned with a healthier individual. In steering consumers, insurers have to consider the contract's impact on the individual's intertemporal utility and their incentive to quit the insurer in favor of their outside option. Investing in better treatments improves future health prospects at the cost of less consumption today. Different insurance contracts affect this tradeoff by changing the consumption given to individuals when a health shock arrives, affecting the cost of investing in future health, and by affecting the continuation payoffs of having improved health.

In future analysis, we plan to incorporate two extensions exploring different aspects of competition and regulation. The first extension examines the impact of consumer lock-down periods, i.e., when individuals are not allowed to change providers for a fixed period of time. Lockdown periods induce a hybrid between competition and monopoly. While insurers compete to attract customers, giving away their rents to consumers, they act as a monopoly during the lockdown period, generating future rents and incentives to steer the healthcare choices of individuals. The second extension further explores the role of competition and its relation to health outcomes. We will explore a model in which insurers are symmetric but differentiated from the individuals' perspective. We will explore how changing the degree of insurance differentiation affects the tradeoff between higher rents of competition and potentially worse health outcomes.

References

CAPPS, C., DRANOVE, D. and SATTERTHWAITE, M. (2003). Competition and market power in option demand markets. *RAND Journal of Economics*, pp. 737–763.

COOPER, Z., GIBBONS, S., JONES, S. and MCGUIRE, A. (2011). Does hospital competition save lives? evidence from the english nhs patient choice reforms. *The Economic Journal*, **121** (554), F228–F260.

EINAV, L., FINKELSTEIN, A., RYAN, S. P., SCHRIMPF, P. and CULLEN, M. R. (2013). Selection on moral hazard in health insurance. *American Economic Review*, **103** (1), 178–219.

GAYNOR, M., MORENO-SERRA, R. and PROPPER, C. (2013). Death by market power: reform, competition, and patient outcomes in the national health service. *American Economic Journal: Economic Policy*, **5** (4), 134–66.

GHILI, S., HANDEL, B., HENDEL, I. E. and WHINSTON, M. D. (forthcoming). Optimal Long-Term Health Insurance Contracts: Characterization, Computation, and Welfare Effects. *Review of Economic Studies*.

GOWRISANKARAN, G., NEVO, A. and TOWN, R. (2015). Mergers when prices are negotiated: Evidence from the hospital industry. *American Economic Review*, **105** (1), 172–203.

HANDEL, B. and HO, K. (2021). The industrial organization of health care markets. In K. Ho, A. Hortaçsu and A. Lizzeri (eds.), *Handbook of Industrial Organization*, *Handbook of Industrial Organization, Volume 5*, vol. 5, Elsevier, pp. 521–614.

HANDEL, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, **103** (7), 2643–82.

HO, K. and LEE, R. S. (2017). Insurer competition in health care markets. *Econometrica*, **85** (2), 379–417.

Appendix

Proof of Lemma 1 Because the individual does not have insurance, consumption is constrained to the disposable income available after the realization of the shock, which determines (6). The optimal treatment at $t < T$, equation (5), follows from the first-order condition of the problem (4), which is also sufficient given our concavity assumptions. □

Proof of Lemma 2 We prove the lemma using the Kuhn-Tucker theorem. Let λ be the Lagrange multiplier of the problem. If $\lambda = 0$, the restriction is not binding, and the first-order condition with respect to $\tau(m_t)$ becomes $\tilde{u}'(\alpha(m_t); m_t, \tau(m_t)) > 0$. This implies that the objective function is increasing in the transfer, and the transfer increases until the restriction binds. When $\lambda > 0$, the set of first-order conditions is given by

$$[\tau(m_t)] : \tilde{u}'\left(\alpha_t^{[\mathtt{c}]}(m_t); m_t, \tau_t^{[\mathtt{c}]}(m_t)\right) - \lambda = 0$$
$$[\alpha(m_t)] : -u'\left(\alpha_t^{[\mathtt{c}]}(m_t); m_t, \tau(m_t)\right) k'(\alpha_t^{[\mathtt{c}]}(m_t), m_t) + f_\alpha(s, m_t, \alpha_t^{[\mathtt{c}]}(m_t))\delta \Delta V_t^{[\mathtt{c}]} = 0$$

The first set of first-order conditions implies that the marginal utility of consumption should be equal in every state, i.e., full insurance. The second first-order condition translates into equilibrium condition (9) when the full insurance result is factored in. Finally, consumption is found by pining down the transfers that deliver full insurance, give the maximal rent to the individual, and satisfy the non-negative profits constraint. This process delivers equation (8). □

Proof of Lemma 3 We prove the lemma using the Kuhn-Tucker theorem. Let λ be the Lagrange multiplier of the problem. If $\lambda = 0$, the restriction is not binding, and the first-order condition with respect to $\tau(m_t)$ becomes $-g(m_t \mid s) < 0$, where $g(m_t \mid s)$ is the probability the shock m_t arrives given the current health status s. This implies that the objective function decreases in the transfer, and the transfer decreases until the utility of the agent binds with its outside option. When $\lambda > 0$, the set of first-order conditions is given by

$$[\tau(m_t)] : -g(m_t \mid s) + \lambda g(m_t \mid s)\tilde{u}'\left(\alpha_t^{[\mathtt{m}]}(m_t); m_t, \tau_t^{[\mathtt{m}]}(m_t)\right) = 0$$
$$[\alpha(m_t)] : f_\alpha(s, m_t, \alpha_t^{[\mathtt{m}]}(m_t))\delta \Delta \pi_{t+1}$$
$$+ \lambda\left[-u'\left(\alpha_t^{[\mathtt{m}]}(m_t); m_t, \tau(m_t)\right) k'(\alpha_t^{[\mathtt{m}]}(m_t), m_t) + f_\alpha(s, m_t, \alpha_t^{[\mathtt{m}]}(m_t))\delta \Delta V_t^{[\mathtt{n}]}\right] = 0$$

The first set of first-order conditions implies that the marginal utility of consumption should be equal in every state, delivering the full insurance result. Let $c_t^{[\mathtt{m}]}$ be the full insurance consumption. The Lagrange multiplier is equal to $\lambda^{-1} = u'\left(c_t^{[\mathtt{m}]}\right)$. Replacing the multiplier in the second first-order condition translates into equilibrium condition (11) when the full insurance result is factored in. Consumption is given by the minimal certain-consumption that gives the consumer a net present value of $V_t^{[\mathtt{m}]}(s)$, delivering (10). Finally, profits are determined by the difference between the health shock's expected value and the consumption guaranteed to consumers. □