WARWICK

ECONOMICS

The Right to Health and the Health Effects of Denials

Sonia Bhalotra & Manuel Fernandez

(This paper also appears as CAGE Discussion paper 586)

October 2021(Updated March 2023) No: 1376

Warwick Economics Research Papers

ISSN 2059-4283 (online)

ISSN 0083-7350 (print)

The Right to Health and the Health Effects of Denials*

Sonia Bhalotra†
University of Warwick

Manuel Fernández‡
Universidad de los Andes

10 January 2023

Abstract

We investigate supply-side barriers to medical care in Colombia, where citizens have a constitutional right to health, but insurance companies impose restrictions. We use administrative data on judicial claims for health as a proxy for unmet demand. We validate this using the health services utilization register, showing that judicial claims map into large, pervasive decreases in medical consultations, procedures, hospitalizations and emergency care. This manifests in population health outcomes. We identify increases in mortality pervasive across cause, age, sex, and income, with larger increases for cancer and individuals over the age of fifty.

JEL classifications: G22, I11, I13, I18, K38, K42
Keywords: Health care, health insurance, mortality, right-to-health, litigation, accountability, universal-health-coverage, Colombia

*We acknowledge funding from ESRC grant ES/M010263/1 for the Human Rights Big Data and Technology Project (hrbdt.ac.uk) at the Human Rights Centre, University of Essex. We are grateful to Damian Clarke, Marty Gaynor, Paul Hunt, Grant Miller, Sharun Mukand, Clara Sandoval-Villalba, Bob Town, Marcos Vera-Hernandez, Simon Weidenholzer, Carmel Williams and Nicolas Ziebarth for helpful comments and discussions.

†University of Warwick, United Kingdom, CEPR, IZA, IEA, CAGE. E-mail: sonia.bhalotra@warwick.ac.uk

‡Universidad de los Andes, Colombia, and IZA. E-mail: man-fern@uniandes.edu.co.

1 Introduction

Universal health coverage has recently garnered widespread support as a policy objective (Bloom et al., 2018). It is seen as marking the third great transition in public health, following the demographic and epidemiological transitions of the past two centuries (Rodin and de Ferranti, 2012). However, it is increasingly recognized that there are weak links between coverage and population health outcomes. Some studies have highlighted low uptake (Banerjee et al., 2010, 2021), while others have highlighted poor service quality (Mohanan et al., 2014; Powell-Jackson et al., 2015; Kruk et al., 2018). We instead investigate supply-side denials or restrictions on access to care.

The setting is Colombia, where a state commitment to the right to health for all (Article 48) was written into the new Constitution formulated in 1991, following which a new health care system that progressively reached nearly universal health coverage was put in place.[1] The Constitution provides judicial protection of the right to health and a series of other human rights (Article 86), allowing citizens who fear that their rights are jeopardized because of restrictions to file claims for redress with the Constitutional Court.[2] These judicial claims (*tutelas*) are costless, simple, can be filed with any judge within the local jurisdiction, and have preferential proceedings such that the judge is mandated to return a decision within ten days. This is a major 'experiment' in accountability.

The use of tutelas and complaints as an enforcement mechanism has become widespread in Colombia, leading to what is referred as judicialization of health care (Yamin and Parra-Vera, 2009, 2010; Vargas et al., 2010; Parra-Vera and Yamin, 2013; Andia and Lamprea, 2019). Health insurers have systematically restricted enrollees' access to required medical services, primarily through delays or denials for examinations, medicines, surgeries, or treatments that a health professional has already prescribed (Maya, 2008; Corte Constitucional, 2008; Yamin and Parra-Vera, 2009, 2010; Parra-Vera and Yamin, 2013; Arrieta-Gómez, 2018).[3] Citizens have responsed to these restrictions by pursuing litigation. Between 2010 and 2016, more than 675,000 tutelas (judicial claims) were filed against health insurers. On average, 80% of health-related claims were decided in favor of claimants. We provide what appears to be the first quantitative analysis of the content of the judicial claims

[1] At 98 percent, Colombia has one of the highest levels of comprehensive insurance coverage among developing economies.

[2] Colombia is one of 63 countries that has an independent constitutional court, a high court that deals primarily with constitutional law. Other countries, including the US, delegate constitutional judicial authority to their general court system, with the final decision-making power resting in the Supreme Court.

[3] See also the periodical reports of the Colombian Ombudsman office (Defensoría del Pueblo, 2004, 2007, 2009, 2010, 2011, 2012, 2013, 2015); and a report by the Ministry of Health (Ministerio de Salud y Proteccion Social, 2018).

instrument using comprehensive administrative data.

A key feature of our study is that we use data on judicial claims to measure unmet demand for medical care. We show how this is (empirically) identified, and then estimate the extent to which it impacts population health. We define unmet demand as the set of health services requested (demanded) but not delivered because of supply-side barriers, irrespective of whether they are part of the mandatory benefit plan or not, but conditional on being prescribed by an independent health professional.[4]

Organization of health care Supply-side denials reflect remediable features of the organization of health care. Colombia's commitment to health coverage has been delivered, since 1993, through a managed competition model of health insurance. Health insurers are responsible for organizing the provision of health services, but tight regulation constrains the extent to which insurers can vary the premium, the content of the benefits package, or risk-based enrollee selection, the strategic variables that are typically available to insurers to manipulate. Competition between insurers for enrollees is thus driven by the network of providers and service quality (Miller et al., 2013). In practice, insurers have to be authorized to operate in a municipality,[5] so competition occurs at a sub-national level, but markets tend to be highly concentrated. Insurers get a risk-adjusted fixed fee per enrollee that is common across insurers, while the costs they incur depend on a variable demand for medical services. These factors generate an incentive to impose barriers to restrict access to medical care as a cost-reduction strategy. This is reinforced by the punishments for doing so being small. A weak punishment regime reflects the weak administrative capacity that characterizes many low and middle-income countries (Banerjee et al., 2021).

Theoretical framework To guide the empirical strategy, we sketch a conceptual framework that captures the decision of insurers to deny health claims, and the decision of citizens to litigate when faced with a denial. Denials and the share of denials that result in litigation are equilibrium objects that jointly determine the tutela rate (i.e., the number of tutelas against health insurers per 1,000 enrollees). We show that the tutela rate (which is observed) is a scalar multiple of unmet demand (which is unobserved). The model clarifies the determinants of movements in the tutela rate and the role of demand vs. supply forces; see Section 4 and Appendix Section A. It thus also helps us to set out clearly the identifying assumptions, and to devise tests of these assumptions.

[4]The majority of the medical services restricted, more than 65%, are part of the mandatory benefits package, and close to one-third are outside the benefit plan but are deemed necessary by an independent health professional.

[5]Municipalities are decentralized subdivisions of which there are 1,120 in the country.

Data We merge numerous administrative data sets, each of which contains the entire population. Data on all judicial claims (tutelas) and complaints were obtained by filing a petition under the Right to Information Act. Mortality, a widely used marker of population health, was obtained from vital statistics data. We also obtained administrative data on enrollment by insurers, and a register containing the universe of medical services provided, classified by type of diagnoses (e.g., neoplasms vs. mental disorders) and type of service (consultations, procedures, hospitalization, emergencies).

Mechanism We demonstrate the presumed mechanism by showing that the utilization of medical services is decreasing in unmet demand. This provides a profile of the size and scope of the impact of restrictions, and it validates our approach of proxying unmet demand with judicial claims. Importantly, it also makes it unlikely that our results for mortality are driven by demand side variation. This is because demand shocks will tend to move health service utilization and mortality in the same direction. In contrast, we expect supply-side denials to reduce service utilization and increase mortality.

Identification We nevertheless adopt an empirical strategy that increases our confidence that we isolate supply-side restrictions from demand-side factors. By virtue of using insurer-municipality-year data, we can control not only for cross-sectional heterogeneity at the insurer-municipality level but also for shocks at the insurer level and at the municipality level.[6] To address the concern that there remain relevant omitted variables that vary at the insurer-municipality-year level, we additionally introduce a placebo test, and an instrumental variables approach that isolates supply-side variation, see Section 5.

The residual threat to our strategy is dynamic selection. In particular, if individuals change insurers in response to restrictions, this could generate selectivity (or endogenous heterogeneity) in the risk profile of enrollees across insurers.[7] We therefore model enrollee mobility across insurers using individual data. Only 14 percent of enrollees change insurers in any year (even though it is relatively inexpensive and easy to do so). Importantly, we are able to demonstrate that switching insurers is unrelated to the number of judicial claims (or complaints, discussed below) made in the year before the change.

To provide an external check on the estimates obtained using tutela (judicial claims) data, we use an alternative measure of supply-side restrictions: the number of complaints

[6] These controls capture relevant baseline differences and shocks to health, including infections, demographics, information, the efficiency of provision, system-level changes, cultures of claiming, or the leniency of local courts.

[7] Recall that our baseline strategy controls for insurer-municipality fixed effects which account for any cross-sectional heterogeneity that is constant over time, and for insurer x year and municipality x year shocks but dynamic selection operates at the insurer-municipality-year level.

against health insurers made to the National Health Superintendent (Supersalud), the agency in charge of oversight of the system. Between 2013 and 2017, the years for which administrative data on complaints are available, there were 1.5 million individual complaints against health insurers, the majority of which were related to restrictions imposed by insurers on access to requested medical services. Importantly, different agencies collect and manage the claims and complaints. An additional advantage of the complaints data is that we can identify the age and sex of the plaintiff in these data, which allows us to analyze heterogeneity in the effects of complaints on health outcomes along these dimensions.

Main Findings We find large effects of denials (unmet demand) on population-level mortality rates. The marginal effects are similar irrespective of whether we use judicial claims (tutelas) or complaints to measure unmet demand. The results are also broadly similar whether or not we use an instrument, increasing our confidence that selection is not a major issue. Using the IV specification, we estimate that doubling the tutela rate (i.e., a 100% increase in judicial claims) is associated with an increase of between 0.13 and 0.19 standard deviations of the mortality rate, which corresponds to an increase of between 15.4 and 28.9 percent over the average mortality rate.[8] Mortality rates are used as a measure of population health because they are consistently available over a long period from vital statistics data. However, mortality is an extreme measure that represents the tip of the iceberg of underlying morbidities. We might therefore regard our estimates for mortality as a lower bound of the impact of denials on population health.

The impact of denials on mortality is pervasive, being evident across cause of death, age, sex, and broad income groups. The strongest impacts are on mortality rates due to certain cancers, which is consistent with the fact that the diagnoses more commonly invoked in tutela claims are neoplasms. Impacts are also stronger among people over 50.

We proceeded to estimate the extent to which denials hamper health service utilization. This is a direct measure of denials, and the driver (or mechanism) for increased mortality. We estimate that doubling the tutela rate reduces the per capita number of *hospitalizations* by 0.71 standard deviations, of *emergencies* by 0.65 standard deviations, of *consultations with health professionals* by 0.91 standard deviations, and for *medical procedures* by 0.13

[8] Although not directly comparable in concept or context, to benchmark these effect sizes, we note that studies analyzing Medicaid expansion in the US have found impacts of health insurance coverage on mortality in a wide range, from no effect (Finkelstein and McKnight, 2008; Weathers and Stegman, 2012) to a reduction in all-cause mortality of between 6.1 percent (Sommers et al., 2012) and 9.4 percent (Miller et al., 2021). Goodman-Bacon (2018) estimates that, after the introduction of Medicaid, mortality among nonwhite children on Medicaid fell by 20 percent, leading to a reduction in aggregate nonwhite child mortality rates of 11 percent. More recently, Goldin et al. (2020) exploits experimental variation from a randomized outreach study finding that the average per-month effect of getting health insurance on two-year mortality was approximately -0.18 percentage points among middle-aged adults.

standard deviations. Thus denials reduce the utilization of the full range of medical services. We also find that they reduce utilization across the range of common diagnoses. As indicated earlier, these results not only delineate a mechanism but also validate the use of tutelas (or complaints) as a measure of unmet demand.

We investigated whether unmet demand is greater in municipalities with weaker insurer-level competition and found no significant association. We similarly find no evidence that increased mortality emerges from municipalities where insurers face lower competition. This is consistent with a recently evolving literature on the industrial organization of healthcare markets which shows, for the American market, that the impact of insurer competition on welfare, negotiated provider prices, and premia is theoretically ambiguous (Ho and Lee, 2017). In our setting too, it is unclear that enforcing more competition will induce better provision, for instance because greater market power may enhance the bargaining power of insurers when negotiating contracts with providers, generating cost reductions that lower the incentive to impose restrictions on demand.

Relevance and Contributions Our study is relevant in light of the current wave of expansion in health care provision in many countries, enhanced by a global agenda for universal health coverage. Despite considerable expansions in coverage, the WHO estimates that at least half of all people in the world do not receive the health care they need. With public budgets often stretched, effective delivery and accountability issues are increasingly important.

Our results highlight the costs to population health and to the public purse associated with the incentives of providers being unaligned with constitutional commitments. There is too much regulation insofar as insurers are not allowed to set prices or select enrollees and, at the same time, too little regulation insofar as insurers are not sufficiently penalized when complaints against them are received and upheld.

We discuss broadly related research in more detail in Section 8. Here we provide a brief discussion in order to delineate where the contributions of this paper lie. Our work is related to a literature studying the health effects of expanding health insurance coverage which highlights the importance of studying outcomes alongside access (see, among others, Baicker et al. (2013); Miller et al. (2013); Gruber et al. (2014); Goodman-Bacon (2018); Goldin et al. (2020); Miller et al. (2021)). We contribute to this literature by identifying the extent to which slippage between coverage and access can hamper the realization of the potential benefits of coverage.

Although the incentive structure that allows insurers to restrict access to health care is particular to the Colombian system, similar problems arise whenever health insurance cov-

erage does not guarantee effective access to medical care, i.e. whenever de-jure and de-facto access are not equivalent. For instance, studies of Medicaid have noted supply shortages and uncompetitive physician fees (Currie et al., 1994; Government Accountability Office, 2011; Rosenbaum, 2014; Polsky et al., 2015; Oostrom et al., 2017), as well as administrative burdens that distort physicians' behavior and reduce the availability of providers, with potentially harmful effects on patients (Cutler and Ly, 2011; Dillender, 2018; Dunn et al., 2021). We contribute by identifying the size and spread of the unmet demand for medical services and the associated population-level health costs of denying medical care.

Organizational structure in the US health insurance market has been the subject of a number of recent studies (Gaynor et al., 2004; Einav and Levin, 2015; Duggan et al., 2023), all of which demonstrate that private insurers (or providers) have financial incentives to limit costs, which result in restrictions on patient access. Managed care is the dominant form of health insurance in the US and many other OECD countries, and there is a growing shift towards regulated managed competition marketplaces. Incentives for cost control are built into these systems (Gaynor et al., 2004; Einav and Levin, 2015). Einav and Levin (2015) observe that private insurers tend to limit marginally valuable care. We show that when regulatory institutions are weak, insurers also limit valuable care with rampant effects on population health. Duggan et al. (2023) show that hospital privatization results in an aggregate market-level decline in utilization, which they interpret as a decline in access to care. We similarly observe changes in utilization, but we additionally have data on judicial claims and complaints which provide objective measures of access to care (i.e., supply restrictions, distinct from any demand changes).

Our work is also related to a literature studying how the use of litigation by patients can influence the provision of health care (Kessler and McClellan, 1996; Danzon, 2000; Kessler and McClellan, 2002; Currie and MacLeod, 2008; Frakes, 2013; Frakes and Jena, 2016). Economists have focused on the effects of liabilities for malpractice on the quality and cost of medical care. We depart from this thread by focusing upon a context in which litigation is used not to deter medical malpractice but to enforce access to medical care, supported by the constitutionalization of health as a human right. The right to health is explicitly recognized, to a greater or lesser extent, in the written constitutions of more than half of all UN member countries (Kinney and Clark, 2004; Backman et al., 2008; Ruger et al., 2015). However, it is only in some countries- such as Colombia- that constitutional provisions accompany this notional commitment to guarantee enforcement, including the possibility of legal recourse.[9]

[9] Although they are less widespread in these countries, similar provisions exist in Argentina (Bergallo, 2011), Brazil (Biehl et al., 2009; Martins and Allen, 2020), Costa Rica (Monge, 2019) and South Africa (Wilson, 2011).

We do not (cannot) analyze the impacts of judicial accountability for the right to health. Instead, we exploit that judicial claims carry information on denials of prescribed medical treatments that are generally undocumented to analyze the impacts of denials, conditional on judge leniency.[10] Our setting draws attention to constitutional design and accountability issues, highlighting a lacuna in modeling the constitutionalization of citizen rights. An explicit commitment to fulfilling rights such as the right to health or education or food increases salience and, when supported by the powers of a constitutional court, provides a compelling framework for enacting human rights law. We are unaware of research that evaluates this framework's incentive structure or welfare consequences. In broadly related work, Fox and Stephenson (2011) model judicial review, analyzing its impacts on the incentives of elected leaders to posture by enacting bold but possibly ill-advised policies and on voter welfare. Maskin and Tirole (2004) consider the implications of constitutional design for public choices, analyzing the strengths and weaknesses of electoral accountability and the role of non-accountable officials, including judges.

The rest of this paper unfolds as follows. Section 2 outlines the relevant features of the Colombian health care system and the process of judicialization of access to health care. Section 3 describes the data sources. Section 4 discusses a theoretical framework. The econometric model and identification strategy are presented in Section 5. In Sections 6 and 8 we present and discuss the results and robustness checks. Section 9 concludes.

2 Institutional Background and Policy Context

Structure. The health care system in Colombia is organized as a tightly regulated managed competition model of health insurance.[11] Individuals enter the system by enrolling with a health insurer operating in their municipality of residence. The insurance company is responsible for guaranteeing enrollees access to required medical services through their independently contracted network of providers (e.g., hospitals, clinics, laboratories). Citizens can freely choose which insurer to enroll with, and which providers within the insurer network to use. There are two regimes that have a common structure but differ in their target populations: the contributory regime (CR) covers formal sector employees and those in self-employment that are able to pay, while the subsidized regime (SR) covers the low-income population. Affiliation to the CR is mandatory for all formal employees, while eligibility for

[10] As indicated earlier, insurer-municipality fixed effects and municipality x year fixed effects purge variation in judge leniency from our estimates.

[11] For a characterization of the managed competition model of health insurance, see Enthoven (1978); Londoño and Frenk (1997). We provide an overview of characteristics relevant to our analysis; an in-depth description is available in Gaviria et al. (2006); Maya (2008); Glassman et al. (2009). Analysis of the SR is provided in Miller et al. (2013).

the SR is means-tested.[12]

Insurers are constrained by regulation of three key strategic variables. First, insurers cannot set premia but receive a standardized risk-adjusted capitation payment per enrollee defined periodically by the government.[13] Second, they must offer a standardized compulsory benefit package (known as POS) that includes preventive care services and essential procedures and medications.[14] Finally, insurers cannot deny enrollment on the basis of demographics or pre-existing conditions, so they cannot directly engage in risk selection.[15]

The high coverage of the system is a marker of success. The average share of the population enrolled with an insurer in either the CR or the SR between 2010 and 2018 was 89 percent (Appendix Figure B.1), with half of the covered population in each regime. Between 2010 and 2018, there were 80 active insurers on average per year in the country, 48 in the CR and 32 in the SR. The average number of insurers per municipality was 6.5 in the CR, and 4.2 in the SR (Panel (a) of Appendix Figure B.2). However, approximately 10 percent of municipalities have a single insurer in the SR (Panel (b) Appendix Figure B.2), while between 10 and 20 percent of all municipalities have at most two insurers in the CR (see Panel (c) of Appendix Figure B.2).

Restrictions. There is extensive evidence that insurers have been systematically imposing barriers to limit the utilization of medical services (Maya, 2008; Abadia et al., 2009; Yamin and Parra-Vera, 2009; Rodríguez, 2012; Arrieta-Gómez, 2018). These barriers take the form of delays or denials of authorizations for exams, medicines, surgeries, or treatments prescribed by an independent health professional.[16] This is a well-known problem inside the country,

[12]The contributory regime regime is financed by payroll contributions and corporate income taxes, while the subsidized regime is financed by 'solidarity' contributions from the CR and transfers from both subnational governments and the central government.

[13]There is a cost-sharing scheme through co-payments for outpatient care services that applies only to individuals in the contributory regime who earn more than twice the minimum wage (Buitrago et al., 2021). However, only a small fraction of the population is affected by these co-payments, partly because half of the enrollees in the system are in the subsidized regime and partly because the threshold value of twice the minimum wage is above the seventh decile of the income distribution. Importantly, co-payments do not vary across insurers.

[14]If a required service or medication is not included in the benefits package, users have to pay for it unless it is deemed essential and they lack the resources to cover the cost. In such cases, the insurer has to guarantee the timely provision of the service, and the government reimburses the insurer for the cost incurred.

[15]The capitation can vary by age, gender, municipality of residence and regime, but is common across insurers conditional on these variables. The composition of the POS is also common across insurers, being defined by the Ministry of Health and changed periodically. It is designed to be comprehensive enough to cover most population health care needs.

[16]For example, an insurer may impose an inappropriate administrative cost, pecuniary or other, that, if not met, leads to a delay or rejection of an authorization. They might exploit 'gray areas' in the benefits package definition, arguing that a part of treatment, examination, or medication was not explicitly included in the benefits package to delay or deny the entire treatment.

and press articles documenting these practices abound (see Appendix Figure B.3). That insurers have an incentive to impose these barriers is largely explained by the pay-structure they face: they are mostly financed by capitation payments, which are set at a constant rate for the year, but the costs they incur are variable and directly determined by the utilization of care services. It is then in the financial interest of insurers to limit the utilization of services and medications to reduce variable costs. Although some limitations on health care use can be efficient (counterbalancing demand-side incentives leading to ex-post moral hazard), others may be inefficient and harmful (Miller et al., 2013). The financial incentive is compounded by the inability of the agencies in charge of oversight of the health care system to enforce the law, which prohibits these practices (Maya, 2008; Corte Constitucional, 2008; Rodríguez, 2012).

Tutela Writs. In response to restrictions, citizens have increasingly used judicial claims available under Colombian law, the most salient being the tutela writ. The Colombian Political Constitution, enacted in 1991, explicitly recognized a broad set of fundamental rights for all its residents, among which is the right to health.[17] Crucially, the Constitution recognized the rights and created a legal enforcement mechanism called the tutela writ. This is a legal claim designed to get express access to the judicial system when rights are violated. The claim is i) costless; ii) can be filed with any judge within the local jurisdiction; iii) need only contain the basic facts so that a judge can address the case; and iv) has preferential proceeding so that the judge has to return a decision in the first instance within ten days.

The Constitutional Court of Colombia has stated explicitly that when insurers impose barriers to access health services, even those not included in the mandatory benefit package, it violates the right to health.[18] An individual can file a tutela against an insurer when a service is delayed or denied and, if the judge decides in favor of the claimant, the insurer is forced to provide it. No additional penalty is imposed upon the insurer beyond the mandate to authorize the services requested. Since some denials do not lead to tutela claims and some claims are not granted, this incentivizes the insurer to continue restricting access. Due to

[17]The right to health is broadly defined as the right to access health services in a timely and efficient manner to allow for the preservation, improvement, and promotion of health. See Articles 44, 48, and 49 of the Colombian Political Constitution.

[18]In section 3.2.1.3 of sentence T-760 of 2008, the Court states: "*The Constitutional jurisprudence has indicated that access to a required health service, contemplated in the mandatory plans, is an autonomous fundamental right. To this extent, the denial of health services contemplated in the POS violates the fundamental right to health; therefore, it is an enforceable and justiciable provision through a Tutela writ.*". Later on, discussing services not included in the mandatory benefit package, the Court states that "*when a person requires a health service that is not included in the Mandatory Benefit Plan, and lacks resources to cover its cost, the entities in charge of ensuring the provision of the service must adhere to their responsibility and, consequently, ensure access to it. However, it is the Government that has to bear the cost of the service since it must guarantee the effective enjoyment of the right.*" Corte Constitucional (2008).

its simplicity and fast turnaround, the tutela has become a very popular instrument in the country. The number of tutelas invoking *any* fundamental right increased from 0.3 to 12.7 per 1,000 inhabitants between 1992 and 2016 (See Panel (a) of Figure 1),[19] with tutela writs invoking the right to health accounting for 20 to 40 percent of all claims.

The growth over time has been substantial: in 1999, the total number of tutelas presented to courts invoking the right to health was 21,301 (0.53 per 1,000 inhabitants), but by 2016 the number reached 163,977 (3.36 per 1,000 inhabitants), a more than sevenfold increase in a span of 17 years.[20]

Four features of the data on tutelas invoking the right to health are noteworthy. First, between 65.9 and 82.4 percent are made against insurers (Panel (a) of Figure B.4).[21] Second, more than 65% of the services requested in tutela writs during the period of analysis were part of the compulsory benefit package (Panel (b) of Figure B.4). This shows that a large share of litigation is to access services that people should be able to get without going to court.[22] Third, between 72.2 and 86 percent of tutelas are decided in the first instance in favor of the claimant (Panel (c) of Figure B.4). This implies that, in the eyes of Colombian jurisprudence, citizen claims are more often than not recognized as valid. Fourth, although the share of enrollees in each regime is approximately equal, only between 12.7 and 36.7 percent of tutelas against insurers are made by individuals in the subsidized regime (Panel (d) of Figure B.4). Thus the low-income population is less likely to use this judicial instrument to protect their rights.

Complaints to the Supersalud. Tutela writs are a powerful legal instrument in Colombia, but they are not the only mechanism that people have to enforce access to restricted medical services. Individuals can also make complaints (and requests) to the National Health Superintendent (Supersalud), the agency in charge of overseeing the health care system. The

[19]Unless stated otherwise, the statistics reported in this section are taken from reports produced by the Colombian Ombudsman office (Defensoría del Pueblo, 2004, 2007, 2009, 2010, 2011, 2012, 2013, 2015) and the Ministry of Health and Social Protection (Ministerio de Salud y Proteccion Social, 2018).

[20]The extent to which individuals were using tutela writs invoking the right to health during the early 2000s led the Colombian Constitutional Court to make a major pronouncement: Ruling T-760 of 2008. In the ruling, the Court provided a thorough analysis of the problems that led to the judicialization of the health care system and ordered a set of changes, including i) full clarity about the procedures and medications that were included in the benefit packages; ii) periodical audits of the insurers in order to inform users of their performance; iii) strengthening the regulatory agency overlooking the system; and iv) a revision of the capitation payments that had been unchanged for several years. The number of tutelas invoking the right to health declined temporarily between 2008 and 2009, but by 2015 they were back to the pre-ruling levels. We can only study the period after the 2008 pronouncement because the tutela data are only available after then.

[21]Less than 4 percent are made against providers, and between 15 and 30 percent are made against local authorities, governmental agencies or other actors in the system.

[22]The remaining 35% of claims are nevertheless for medical services prescribed by a medical professional.

Supersalud can apply sanctions against insurers, ranging from economic penalties to the removal of the license to operate, and can mediate disputes between insurers and enrollees regarding access to requested services. The complaints process is simple: citizens can file complaints to the Supersalud in person, on the telephone, in writing, through an online platform, or using the social network accounts (Facebook and Twitter) of the agency. The Supersalud notifies the defendant, who is asked to respond to the complaint. If mediation is necessary, the Supersalud decides on the case within ten days.

Between 2013 and 2017, the years for which the data on complaints is available, people made 1,261,247 complaints against insurers, an average of 7.83 complaints per 1,000 enrollees in the system per year. This number is 2.5 times larger than the tutela rate during the same period[23] (Panel (b) of Figure 1). Nevertheless, Tutelas and complaints are different instruments that can be used to enforce access to restricted care services. Panel (a) of Figure 2 provides a scatter plot of tutela against complaint rates for every insurer × municipality × year cell.[24] The correlation between the two instruments is 0.35.

3 Data

We use five sources of administrative data for the analysis. These data are available either at the individual level for the entire population (vital statistics), as insurer × municipality × year averages (enrollees, tutelas, complaints), or as municipal × year averages (health services register). There are 1,120 municipalities, 80 insurers on average per year,[25] and, in the most comprehensive dataset, we have information for 2010–2017.

Vital Statistics. As a marker of the population-level health costs of denials, we use overall, cause-specific, age-specific, and sex-specific mortality rates at the level of insurer × municipality × year between 2010 and 2017, which we calculate from individual data on mortality. We use mortality because it is one of the few objective, well-measured health outcomes available over time (Finkelstein and McKnight, 2008), and because the vital statistics registers include information on the municipality of residence of the deceased and the insurer with which he or she was enrolled, if enrolled.

Enrollment. We use individual-level administrative data on the insurer in which each

[23]The higher rate of complaints relative to tutelas might reflect two differences between them: first, a complaint is easier to make. The Supersalud provides many different channels for users to make complaints, while to file a tutela, you must write and deliver a document to a judge. Second, for a judge to rule a tutela in favor of the claimant the individual not only has to show that access to a health service was restricted, but that this restriction violates their right to health.

[24]Between 2013 and 2016, the years for which the data on tutelas and complaints overlap, we have 45,403 observations at insurer × municipality × year level.

[25]If an insurer is operating in both the subsidized or contributory regime we treat it as two different entities.

person is enrolled at the beginning of each year from 2010–2017. When necessary, we aggregate this data into the relevant unit of analysis. The data comes from the Ministry of Health and Social Protection registries.

Tutelas. We use the number of tutelas filed against health insurers and invoking the right to health. We obtained these data from the ombudsman's office after we filed a formal petition supported by right-to-information legislation. The data covers the universe of health-related tutelas put forward to Colombian courts between 2010 and 2016. The information is available at the insurer × municipality × year level. At this disaggregated level, the data are restricted to the number of claims made, so we cannot differentiate by the type of medical care requested or whether the claims were successful.

Complaints. We use the universe of complaints to the National Health Superintendent from 2013-2017. We can identify which insurance company is involved in the complaint, the main (broad) reason for the complaint, and the age and sex of the claimant. We only use complaints against health insurers. We further exclude complaints about administrative procedures unrelated to the provision of health services and any general information requests. We also obtained these data following a right-to-information request.

Utilization of Health Services. We use data that contain all medical services effectively supplied in each municipality × year during 2010–2017, obtained from the Ministry of Health and Social Protection of Colombia and publicly available through an online platform (SISPRO). This information can be disaggregated by type of service: consultation, hospitalization, procedure, or emergency, and by the medical diagnosis associated with it, classified in ICD-10 codes. This information cannot be disaggregated by insurer.

4 Theoretical Framework

This section discusses the theoretical framework that underlies our empirical strategy. A *formalization* is presented in Appendix Section A; here we provide a simplified schematic description focusing on the salient features of the problem. The aim is to make explicit the main factors that determine the decision of insurers to restrict access to a service and the decision of enrollees to file tutelas or complaints. This allows us to characterize what is meant by unmet demand, and how unmet demand relates to the prevalence of judicial claims.

The basic structure of the problem is summarized in the decision tree shown in Figure 3. To simplify, suppose there are only two periods, all agents are enrolled with a health insurer, and tutelas are the instrument available to citizens to litigate if their insurer denies them prescribed medical care. In the initial period, individuals observe their health status and, if

required, get prescribed a treatment, procedure, or medication. In the Colombian health care system, these prescriptions can be authorized or denied/delayed (henceforth restricted) by the health insurer (even if they were prescribed or suggested by a medical professional). Any medical service prescribed but not readily authorized is enforceable through litigation, but not all individuals use legal instruments. This is captured by the separation in the branches leading to end node B on the one hand and end nodes C and D on the other. Since only a fraction of tutelas are decided in favor of the plaintiff, i) some restricted services will not be delivered (i.e., end nodes B or C), and ii) some restricted services will be delivered as a result of litigation (i.e., end node D). In the final period, new health outcomes are realized.

Let α_i be the fraction of enrollees in insurer i that file tutelas conditional on having their requested services restricted. In terms of the decision tree, α_i corresponds to the ratio of enrollees that end up in nodes C or D, relative to the total number of enrollees in nodes B, C, or D. It follows that

$$Tutela_i = \alpha_i Restriction_i, \tag{4.1}$$

where $Restriction_i$ is the fraction of enrollees that face restrictions and $Tutela_i$ is the fraction of enrollees that file tutelas. Equation 4.1 describes the direct link between the tutela rate, which we observe, and the restriction rate, which we do not observe.

Both α_i and $Restriction_i$ are equilibrium objects that jointly determine $Tutela_i$. We show in Appendix Section A that α_i is i) increasing in the (average) improvements in health agents get from receiving the restricted care services; and ii) increasing in the (average) probability that litigation is decided in favor of the plaintiff. On the other hand, $Restriction_i$ is i) increasing in the (average) cost of providing the services, ii) decreasing in the (average) probability that litigation is decided in favor of the plaintiff, and iii) decreasing in α_i.

These results clarify, for instance, that it is not the case that restrictions are typically on expensive treatments (and, indeed, the data reveal that tutelas are claimed for all sorts of treatments, expensive and not, see Appendix Figure B.5). This is because some expensive treatments are for life-threatening conditions that citizens are more likely to litigate for and that judges are more likely to deem valid.

The probability that litigation is decided in favor of the plaintiff may vary across municipalities with judge leniency. Judge leniency would be an instrument for *claims granted*, while we want an instrument for the *volume of claims made*. This is because we ask a different question, which is what the health cost of denials & delays is. In our analysis, judge leniency is absorbed by municipality or municipality-year fixed effects.

As long as receiving medical care cannot worsen health, with the demand for medical care held constant, a larger (supply-side driven) restriction rate will translate into worse average health outcomes among the population. It follows from Equation 4.1 that an increase in the restriction rate that is not demand-driven will be reflected in a higher tutela rate and associated with i) lower utilization of health care services and ii) a higher mortality rate. These are the two critical empirical predictions of the model that we test.

The predicted associations will be attenuated by the fact that a fraction of services are delivered after favorable judicial rulings (node D).[26] But the fact that some services are nevertheless restricted implies there is an unmet demand for health care, defined as the share of services that are requested but not delivered. More formally,

$$\begin{aligned} unmetDemand_i &= Restriction_i - ClaimsUpheld_i \\ &= \left(\alpha_i^{-1} - \beta_i\right) Tutela_i, \end{aligned} \tag{4.2}$$

where β_i is the fraction of enrollees that receive medical care as a result of litigation, and $\left(\alpha_i^{-1} - \beta_i\right) \geq 0$.

Panel (b) of Figure 2 shows the raw correlation between the (log) tutela and (log) mortality rates using municipality $\times$ insurer $\times$ year data. Panel (c) replicates the exercise using the complaint rate. Consistent with the model predictions, there is a clear positive association. We argue that this reflects the impact of restricting access to health services on health outcomes, as described in Equation 4.1, which results in an unrealized demand for health care, as described in Equation 4.2. We now proceed to discuss the empirical strategy used to test this hypothesis.

5 Empirical Strategy

For expositional purposes, we focus here on tutela rates, but each model is also estimated using complaint rates, an alternative and independently generated measure of unmet demand. Since each observation corresponds to an average (i.e., a rate) over the enrollees in the insurer-municipality-year triplet, we apply analytical weights in the estimation, but all results are qualitatively unchanged if we estimate an unweighted regression.

5.1 Supply Restrictions and Mortality: Three-Way Fixed Effects and IV

We first estimate a model of the form:

[26] If some judges are more lenient than others, this is captured by municipality x year fixed effects.

$$Mort_{i,m,t} = \tau Tutela_{i,m,t} + \delta X_{i,m,t} + \phi_{m,t} + \theta_{i,t} + \varphi_{i,m} + \epsilon_{i,m,t}, \tag{5.1}$$

where $i \in \{1, \ldots, I\}$ indexes health insurers, $m \in \{1, \ldots, M\}$ indexes municipalities, and $t \in \{1, \ldots, T\}$ indexes years, $Mort_{i,m,t}$ is the mortality rate, $Tutela_{i,m,t}$ is the rate of tutela claims, $\phi_{m,t}$ are municipality specific flexible time trends; $\theta_{i,t}$ are health insurer specific flexible time trends; and $\varphi_{i,m}$ are municipality $\times$ health insurer fixed effects. $X_{i,m,t}$ are time-varying covariates. The main identification challenge is to isolate the variation in the tutela rate induced by supply-side restrictions from demand driven changes. Once that is done, the parameter τ can be interpreted as the effect of supply-side restrictions to medical care access on mortality.

In the theoretical framework, we derived that $Tutela_{i,m,t} = \alpha_{i,m,t} Restriction_{i,m,t}$, where $\alpha_{i,m,t} \in [0, 1]$ is the fraction of enrollees that file tutelas (i.e., that litigate) conditional on having a medical service restricted. The equation makes clear that we need to consider factors influencing the restriction rate and the prevalence of litigation (see Section 4 and Appendix Section A). The likely drivers of demand for both are absorbed by the fixed effects. In particular we are able to control not only for cross-sectional heterogeneity at the insurer-municipality level $\varphi_{i,m}$ but also for shocks at the insurer level $\theta_{i,t}$ and at the municipality level $\phi_{m,t}$. These controls capture relevant baseline differences and/or shocks to health including infections, demographics, information, efficiency of provision, system-level changes, cultures of claiming or the leniency of local courts.

We address the concern that there remain relevant omitted variables that vary at the insurer-municipality-time level by additionally introducing an instrumental variables approach and, later, we also investigate dynamic selection. To construct an instrumental variable, we leverage the fact that insurers operate across multiple municipalities. The key idea is that there are insurer-specific (supply-side) factors that determine the restriction rate and that are orthogonal to *local* (municipality-level) demand for health care. Natural examples of insurer-specific supply-side factors are cost structures determined by different service provider contracts and managerial practices. (In terms of the model parameters, the cost structure would be captured by the distribution of c_i^x, and efficiency would be reflected in the value of $\tilde{c}_i$.).

The instrument is the leave-one-out insurer-specific tutela rate, calculated using information from municipalities in which the insurer operates, but excluding the municipality of interest m:

$$Z_{i,m,t} = \sum_{k \neq m} \nu_{i,k,t} Tutela_{i,k,t}, \tag{5.2}$$

where $\nu_{i,k,t} \equiv \frac{N_{i,k,t}}{\sum_{k \neq m} N_{i,k,t}}$, and $N_{i,k,t}$ is the total number of enrollees in health insurer i, municipality k, and year t. In practice, we calculate $Z_{i,m,t}$ using only municipalities outside the department in which m is located to allow for correlated demand shocks between municipalities that are close to each other. The exclusion restriction holds if the restriction rate in municipalities outside the department where m is located, conditional on the structure of fixed effects described above, does not directly affect the mortality rate in municipality m.

The time-varying covariates $X_{i,m,t}$ include the one-period lagged mortality rate, which serves as a measure of average health outcomes in the previous year and accounts for the possibility of mean-reversion. Although it is not common practice because most municipalities have only a few insurers (Appendix Figure B.2), we also want to account for the possibility that people switch to other insurers due to restrictions, changing the risk profile of enrollees. To do this, we include the market share of the health insurer in the municipality-year,[27] and the change in that market share from the previous year. Importantly, in Section 6.2 we directly provide evidence that individuals do not respond to increased restrictions by changing insurers.

5.2 Supply Restrictions and Utilization of Health Services: Shift-Share - IV

The mechanism through which restrictions affect health outcomes is by limiting the utilization of health care services. We test this prediction using a similar strategy. As the health services data are only available at the municipal $\times$ year level, we modify the empirical strategy for the analysis of mechanisms using a shift-share IV. The premise remains that insurer-specific restriction rates are correlated across municipalities because of supply-side determinants that are orthogonal to local demand for health care. The estimated model takes the form:

$$y_{m,t} = \pi Tutela_{m,t} + \gamma W_{m,t} + \phi_m + \epsilon_{m,t}, \tag{5.3}$$

where $y_{m,t}$ is a measure of health services used, including patient consultations, hospitalizations, procedures, and emergencies, all defined on a per enrollee basis. $W_{m,t}$ is a vector of controls, and ϕ_m is a municipality fixed effect. To construct the instrument, note that the

[27] Market share is defined as total share of enrollees the health insurer has within the municipality.

municipal tutela rate is a weighted average of the tutela rates of each insurer operating in the municipality, where the weights correspond to market shares:

$$Tutela_{m,t} = \sum_i s_{i,m,t} Tutela_{i,m,t}, \tag{5.4}$$

where $s_{i,m,t}$ is the municipal share of enrollees of insurer i at time t. We can decompose $Tutela_{i,m,t}$ into an aggregate, insurer level rate, and an idiosyncratic, municipal level rate:

$$Tutela_{i,m,t} = Tutela_{i,t}^{\neg m} + \mu_{i,m,t}, \tag{5.5}$$

where $Tutela_{i,t}^{\neg m}$ is the tutela rate of insurer i at time t at the national level, calculated leaving out municipality m. As before, we calculate $Tutela_{i,t}^{-m}$ using only municipalities outside the department in which m is located. The instrument is then constructed to utilize variation from the aggregate rates ($Tutela_{i,t}^{\neg m}$), and not from the local idiosyncratic rates ($\mu_{i,m,t}$):

$$B_{m,t} = \sum_i s_{i,m,t_0} Tutela_{i,t}^{-m}, \tag{5.6}$$

where s_{i,m,t_0} is the municipal share of enrollees of insurer i in a baseline year $t_0 = 2010$. Defining the market share in the baseline year ameliorates concerns that people might switch insurers because of restrictions, which would endogenously change the risk-pool distribution. This said, below we show that there is no evidence of endogenous mobility of enrollees across insurers.

Like all shift-share IVs, the instrument has two components: First, how exposed a municipality is to restrictions by an insurer, given by the respective market share s_{i,m,t_0}. Second, how many restrictions were imposed by each insurer on average outside the municipality m, proxied by the tutela rate. The recent literature on shift-share IV shows that the validity of the instrument in this set-up can be argued in terms of the exposure variables (s_{i,m,t_0}) (Goldsmith-Pinkham et al., 2020), or the aggregate shocks ($Tutela_{i,t}^{-m}$) (Borusyak et al., 2022). In our setting, both alternatives are plausible. However, following the discussion in the previous section, we argue that the restriction rate in municipalities *outside* the department where m is located is unrelated to local demand for health care, which implies exogeneity of the aggregate shocks. To further account for local demand changes and allow for mean reversion, we include in the vector of controls $W_{m,t}$ one-period lags of the mortality

rate and each of the four utilization rates.

6 Results: Impact of Restrictions on Mortality

6.1 Main Results

The estimates are presented in Table 1. We show results with and without the instrument, and with and without the time-varying controls. We consistently provide estimates for both tutela and complaint rates as these provide alternative sources of variation in the same underlying variable (restrictions); see Section 2. The F-statistic on the excluded instrument in the first stage is shown and, in all cases, is above 19, which indicates the instrument has sufficient power. There is no evidence of mean reversion.

An increase in supply-side restrictions on access to health services leads to an increase in the mortality rate that is robust to the alternative specifications and both measures. It is both statistically significant and meaningful. The point estimates range between 0.107 (se 0.060) and 0.327 (se 0.080). To get a sense of magnitude, this implies that doubling the average tutela rate –an increase of 2.16 tutelas per 1,000 enrollees– is associated with an increase of between 0.06 and 0.22 SD of the mortality rate. The standardized estimates using complaints data are similar- doubling the average complaint rate –an increase of 4.84 complaints per 1,000 enrollees– is associated with an increase of between 0.11 and 0.22 SD of the mortality rate. These results corroborate the hypothesis that restrictions have important effects on average health outcomes, even on a health outcome as extreme as the mortality rate. The actual incidence of restrictions on population health will be larger as it will tend to include unmeasured impacts on morbidities that do not translate into mortality.

6.2 Robustness Checks and Extensions

Alternative Instrument To improve comparison with the results (below) on the utilization of medical services, Table 2 shows the results of estimating Equation 5.3 using the municipal-year (rather than the municipal-insurer-year) mortality rate as the dependent variable. For comparability, we report the same specifications as in Table 1. The F-statistic on the excluded instrument of the first stage is above 12. The results are similar to those from the three-way fixed effects approach and the corresponding IV.

For the instrumental variable specification that includes the full set of controls, the point estimates show that doubling the average tutela rate is associated with an increase of 0.09 SD in the mortality rate. A similar movement of the complaint rate distribution is associated with an increase of 0.12 SD of the mortality rate. The corresponding numbers for

the same specification in Table 1 were 0.15 and 0.12 standard deviations.[28]

Dynamic Effects Table 3 shows estimates of the baseline three-way fixed effects specification shown in Equation 5.1, but using past (one-year lagged) tutela or complaint rates. These would affect health outcomes if the impact of restrictions takes time to kick in, so results from this exercise are informative of the timing of effects. We find no significant impact of lagged rates, which shows that barriers to access have immediate (within the year) impacts on mortality.

Placebo Table 3 also shows estimates with future supply-side restrictions, captured by the one-year forward rates, on the right hand side. These should not affect current health outcomes, so we expect a null result in this regression. This is what we find. If forward rates had turned out to be significant, this would have flagged that we have failed to measure a relevant unobservable trend at the insurer x municipality x year level.

Dynamic Selection If individuals change insurers as a response to restrictions, this could change the risk profile of enrollees over time. This poses a potential threat to the identification strategy. If individuals who switch insurers when they perceive an increase in supply-side barriers are the ones in worse health (which seems likely), then the remaining enrollee pool will have better health, and our estimates, showing a positive correlation between mortality rates and restrictions will tend to be conservative. However, if individuals who switch insurers when they perceive an increase in supply-side barriers are the ones in better health, then the remaining enrollee pool will have worse health, and the estimated positive correlation between mortality rates and restrictions may be spurious. In fact, this concern is undermined by our finding (reported below) that restrictions are associated with lower service utilization. Nevertheless, we now directly investigate it.

Recall that, in the baseline specification, we include insurer $\times$ municipality fixed effects, which account for any cross-sectional heterogeneity in risk profiles across insurers in local markets. We now allow that there is time-varying insurer $\times$ municipality variation driven by citizens moving across insurers within a municipality. To do this, we acquired individual-level administrative data with information about the insurer with which each individual is enrolled at the start of the year.

[28]Finding that the estimates using municipal-year variation are similar to those using insurer-municipal-year variation indicates that unobservables at the insurer-year and the insurer-municipal level are uncorrelated with the tutela (or complaints) rate. In this way, it mitigates concerns that unobservable trends drive our results.

We first describe the share of all enrollees that move insurer in a year and find that, averaged across the years, this is 14%. We next investigate if the switching of insurer is endogenous, that is, if it changes with the restriction rate (i.e., with the share of tutelas or complaints). The estimates are in Table 4. The dependent variable is an indicator equal to one if an individual changed insurer between year $t-1$ and t. The independent variable of interest is the tutela or complaint rate in year $t-1$, defined at the insurer $\times$ municipality $\times$ year level.[29] We condition upon individual fixed effects and initial insurer $\times$ year and initial municipality $\times$ year flexible trends, where "initial" refers to the insurer or municipality of the individual when she first appeared in the data. Table 5 repeats this exercise, dividing the sample by sex and age groups.

There is no evidence of individuals changing insurers as a response to increased restrictions. This is true regardless of the age and sex of the individual or the subsystem (contributory or subsidized) in which the insurer operates. These results make it unlikely that dynamic selection is the mechanism driving our results.

6.3 Heterogeneity in Impact

By Regime. In Table 6 we repeat the analysis, dividing the sample between the contributory (CR) and subsidized (SR) regimes, which have different target populations, the SR covering the most economically vulnerable segments of the population. Earlier we noted (Panel (b) of Figure 1) that tutela and complaint rates are significantly larger in the CR than the SR, which suggests that the CR population either faces more restrictions or has a greater tendency to use judicial instruments. It is hard to identify which, but both are plausible. We display only the specifications with the full set of controls.

In both regimes, we find evidence that supply-side restrictions lead to an increase in mortality rates. The IV results indicate that, in the CR, doubling the average rate of tutelas and complaints respectively leads to increases of 0.19 and 0.14 SD of the mortality rate. In the SR, a similar change of each of the two measures leads to an increase of 0.14 and 0.27 SD deviations in the mortality rate. The complaint rate estimates are consistent with these results but imprecise.

By Cause of Death. A notable feature of judicial claims is that they pertain to all forms of medical care, supported by a wide range of medical diagnoses. Panel (a) of Figure B.5 shows that of all tutelas between 2000 and 2015, 9.3% requested appointments with

[29] These data are not available at the individual level but, even if they were, the filing of a tutela or complaint is potentially endogenous, being a function of the individual's (unobservable) underlying health and their preferences.

health specialists, 12.8% requested exams, 15.5% surgeries, 17.1% medications, and 18.0% treatments. Moreover, a recent report by the Ministry of Health and Social Protection (Ministerio de Salud y Proteccion Social, 2018) shows that 11.6% of tutelas are supported by medical diagnoses of neoplasia, 9.9% refer to diseases of the nervous system, and 8.7% to diseases of the circulatory system (Panel (b) of Figure B.5).

Although denials appear pervasive, some health conditions require more urgent or comprehensive care, making it plausible that the impact of restrictions varies by cause. We therefore investigate cause-specific mortality rates; see Figure 4. The data on tutelas and complaints by cause are only available aggregated at the national level, so we use the municipality-insurer-year level rate as before. To facilitate exposition, we focus on the 15 most prevalent causes of death according to the classification of diseases defined by the Colombian National Statistical Agency, which has 71 categories, and we only report results for the IV specification using tutela rates. The Figure displays both the coefficient estimates and the standardized effects.

There are two main takeaways from this exercise. First, all of the point estimates are positive, and most are statistically significant at standard levels, indicating that restrictions increase mortality rates pretty much across the spectrum. This is consistent with the wide range of diagnoses supporting the tutela claims. Second, some conditions exhibit a stronger response to restrictions. The standardized effect on the mortality rate from certain forms of cancers and sequelae of accidents and assaults is between 0.21 and 0.22 standard deviations, and from cerebrovascular, respiratory and cancer of breast is between 0.15 and 0.18 standard deviations. In contrast, the impacts on mortality from digestive and pulmonary disease are relatively small, lying between 0.05 and 0.07 standard deviations.

By Age and Sex The information on complaints is available at the individual level, allowing us to identify the age group and the sex of the person affected in the complaint (who is not necessarily the person who files the claim), so we can create group-specific rates. The information on vital statistics is also available at the individual level so that we can construct group-specific mortality rates. We estimate the three-way fixed effects model for different subsamples by combining the two sources.[30]

As a starting point, Panels (a) and (b) of Figure B.6 show the average complaint and mortality rates for males and females and four age groups: 0-5, 6-24, 25-49, and 50 or above. Two patterns are evident in the data: first, complaint rates have a U-shaped pattern across age groups. The lowest rate is at ages 6-24, while the largest is at age 50-plus. Second, the

[30]We cannot identify age and sex in the tutela data, but the impacts of tutela and complaint rates on aggregated mortality rates were similar.

complaint rate tends to be larger for females. In both cases, differences could be explained by a higher propensity to make complaints or by a higher prevalence of restrictions.

The estimated effects are presented in Figure 5. To facilitate exposition, we report results only for the instrumental variable specification with the full set of controls. The standardized effects show larger impacts of restrictions on older people. The impact on mortality rates for individuals above 50 is close to 0.31 standard deviations. This is consistent with older people being more sensitive to access to medical care. The standardized effect is slightly larger for males than for females (0.20 vs. 0.17 SD), although we cannot reject that they are equal at standard significance levels.

By Competition in the Insurance Market A mechanism that could explain the prevalence of supply-side restrictions is limited competition between insurers within municipalities. As shown in Panel (a) of Appendix Figure B.2, the average number of insurers per municipality was 6.5 in the CR, and 4.2 in the SR, but these numbers conceal important heterogeneities: approximately 10 percent of municipalities have a single insurer in the SR (Panel (b) Appendix Figure B.2), while between 10 and 20 percent of all municipalities have at most two insurers in the CR (see Panel (c) of Appendix Figure B.2). Moreover, the average cross-municipality Herfindahl-Hirschman Index (HHI) of market concentration, defined in terms of the share of enrollees, is 5,251 in the contributory regime and 5,463 in the subsidized regime. This is more than twice the lower threshold used by the Justice Department of the United States to classify an industry as highly concentrated.

To investigate the role of competition, we estimated Equation 5.3 interacting the tutela or complaint rates with the HHI index, or with alternative binary classifications of municipalities based on the HHI index. We find no clear evidence that impacts of restrictions on mortality emerge predominantly from low-competition markets- the interaction term is, in general, statistically insignificant at standard levels across specifications.[31] This is not entirely surprising. For example, higher market concentration can generate economies of scale or increase insurers' power to negotiate lower prices with providers, reducing the incentive to use denials to control costs. These trade-offs have been studied in the literature on the industrial organization of health care markets (Gaynor et al., 2015; Ho and Lee, 2017), and they suggest that enforcing more competition may not be enough to induce a better provision of services.

Overall, the results described in this section provide compelling evidence that restrictions on access to health care impose significant health costs on the population. The identified

[31] We also find that the descriptive association of tutelas with the HHI is not statistically significant. Results are available on request.

increases in mortality rates are robust to varying the controls (two and three-way fixed effects models, time-varying controls), the estimator (with and without IV, using the two-way or three-way variation), and the measure of restrictions (tutela vs. complaint rates). This stability of the estimates mitigates concerns that selection drives our results, and we find no evidence of endogenous mobility across insurers. In the next section, we investigate the impacts of restrictions on the likely mechanism- the utilization of health care services. This allows us not only to verify the mechanism in principle but also to identify the size of these effects, and their distribution by service type.

7 Mechanisms: Do Restrictions Cramp Utilization of Medical Services?

Estimates of Equation 5.3 are presented in Tables 7 and 8. The first table considers consultations and procedures, and the second considers hospitalizations and emergencies, all on a per enrollee basis. We show results using the fixed-effects OLS specification and the shift-share IV and, as before, for tutela and complaint rates. The F-statistic on the excluded instrument of the first stage is above 16. The mean and standard deviation of each variable (reported in the tables) shows that, on average, in a year, there are 2.02 (SD 1.16) consultations with health professionals, 1.55 (SD 1.12) medical procedures, 0.036 (SD 0.024) hospitalizations, and 0.112 (SD 0.097) emergencies.

The evidence is clear that supply-side restrictions on access to medical care reduce *all types* of service utilization.[32] To get a sense of magnitude, doubling the average tutela rate reduces the rate of hospitalizations by 0.71 SD, of emergencies by 0.65 SD, for consultations with health professionals by 0.91 SD, and for procedures by 0.15 SD. Doubling the average complaint rate reduces the rate of hospitalizations by 0.41 SD, emergencies by 0.33 SD, consultations by 0.34 SD, and procedures by 0.19 SD.

The results are robust and the effect sizes, relative to the reported means and SD are large. These results indicate that restrictions on access are not temporary, or not only around delay, but that there is a lot of outright denial. Moreover, the evidence lines up with the results for mortality, delineating the mechanism and showing that it is active across the different margins analyzed.

[32] If, for example, restrictions were predominantly on consultations, then one might imagine that as (preventive) consultations fall, there are knock-on effects that lead to an increase in hospitalizations and emergencies. Thus, the direct effects seem to dominate any knock-on effects of denials. It remains possible that there is some substitution within each of the broad categories of consultations, procedures, hospitalizations, and emergency care.

8 Discussion

Our work is broadly related to a literature studying the health effects of expanding health insurance coverage. Available research suggests that expanding health insurance increases medical service utilization (Feldstein, 1977; Manning et al., 1987; Currie and Gruber, 1996a; Gaviria et al., 2006; Finkelstein, 2007; Card et al., 2008; Finkelstein et al., 2012; Weathers and Stegman, 2012; Baicker et al., 2013; Miller et al., 2013; Gruber et al., 2014; Goodman-Bacon, 2018), but with mixed impacts on health outcomes. This highlights the importance of studying outcomes alongside access. Health outcomes that have been shown to improve with expansion of health insurance coverage in the United States include child mortality, birth weight (Currie and Gruber, 1996a,b; Camacho and Conover, 2013; Chou et al., 2014; Gruber et al., 2014; Goodman-Bacon, 2018), self-reported mental and physical health (Gaviria et al., 2006; Finkelstein et al., 2012; Weathers and Stegman, 2012; Baicker et al., 2013), and all-cause mortality, especially among older adults and minorities (Sommers et al., 2012; Goldin et al., 2020; Miller et al., 2021). However, some studies find no effect on measured physical health outcomes (Baicker et al., 2013) including mortality (Finkelstein and McKnight, 2008; Weathers and Stegman, 2012). We contribute to this literature by identifying the extent to which slippage between coverage and access can hamper the realization of the potential benefits of coverage.

Although the incentive structure that allows insurers to restrict access to health care is particular to the Colombian system, analogous problems arise whenever health insurance coverage does not guarantee effective access to medical care. For example, there is evidence that Medicaid recipients in the United States face barriers to access ambulatory care because of low rates of physician participation and shortages of primary care providers (MAG, 1994; Currie et al., 1994; Government Accountability Office, 2011; Rosenbaum, 2014; Sonchak, 2015). Similarly, physicians treating Medicaid patients experience greater difficulty referring them to speciality care relative to the privately insured (Government Accountability Office, 2011; Felland et al., 2013). Recent evidence also shows that administrative burdens can reduce the availability of providers, with potentially harmful effects on patients (Cutler and Ly, 2011; Dillender, 2018; Dunn et al., 2021). These are examples of supply-side constraints that can lead to an under-provision of medical services even when they are part of the benefit plan. However, neither for the US, nor for any other country, are we aware of estimates of the population-level health costs of denying medical care, and we make a contribution on this front.

We also contribute to a literature studying how the use of litigation by patients can affect the provision of health care. Economists have focused upon the effects of liabilities for

malpractice on the quality and cost of medical care (Kessler and McClellan, 1996; Danzon, 2000; Kessler and McClellan, 2002; Currie and MacLeod, 2008; Frakes, 2013; Frakes and Jena, 2016). Results from this literature indicate that the fear of tort liabilities encourages physicians to engage in 'defensive medicine', administering treatments that potentially have little medical impact but that vastly increase the cost of providing health care. Some studies show that reforms that reduce tort liabilities reduce both the number of lawsuits filed and medical expenditures (Kessler and McClellan, 1996; CBO, 2004), with no effect on mortality or the quality of care (Kessler and McClellan, 1996; Frakes and Jena, 2016), although there is also evidence that, in specific situations like childbirth, they can increase medical complications (Currie and MacLeod, 2008).

We depart from this literature by focusing upon a context in which litigation is used not to deter medical malpractice, but to enforce access to medical care, supported by constitutionalization of health as a human right. The idea of institutionalizing rights to health has legal foundations in the Universal Declaration of Human Rights of 1948,[33] and, more explicitly, in the International Covenant on Economic, Social and Cultural Rights (ICESCR) of 1966, a legally binding treaty signed and later ratified by 169 countries.[34] The right to health is explicitly recognized –to a greater or lesser extent– in the written constitutions of more than half of all UN member countries (Kinney and Clark, 2004; Backman et al., 2008; Ruger et al., 2015).

In some cases, this notional commitment is accompanied by constitutional provisions to guarantee its enforcement, including the possibility of legal recourse. In Colombia, as discussed, use of an enforcement mechanism has become widespread, leading to what is referred as 'judicialization of health care' (Yamin and Parra-Vera, 2009, 2010; Parra-Vera and Yamin, 2013). Although it seems less widespread in these countries, a similar provision has been made available in Argentina (Bergallo, 2011), Brazil (Gauri, 2004; Biehl et al., 2009, 2012), Costa Rica (Wilson, 2011; Norheim et al., 2014), and South Africa (Cooper, 2011).[35]

[33]Article 25 of the Declaration states that: "Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family, including food, clothing, housing and medical care..."

[34]The ICESCR declares in its Article 12 that "the States parties... recognize the right of everyone to the enjoyment of the highest attainable standard of physical and mental health." It is not, according to the standard interpretation, a right to be healthy, but a set of entitlements that include "the right to a system of health protection which provides equality of opportunity for people to enjoy the highest attainable level of health,... taking into account both the individual's biological and socio-economic preconditions and a State's available resources." (Alston, 2010, general comment 14, N. 8-9).

[35]The United States has not ratified the ICESCR, and its constitution makes no reference to health rights, but it has nevertheless experienced right to health litigation in specific segments of the population. In the 1976 landmark decision *Estelle vs. Gamble*, the Supreme Court of the United States established the right to medical care for prisoners after a Texas inmate, J. W. Gamble, sued the State Department of Corrections for lack of adequate medical treatment. The failure of correctional officials to honor these rights has resulted in protracted litigation involving hundreds of cases (Rold, 2008; Ruger et al., 2015).

Most of the literature studying judicial enforcement of the right to health is qualitative, often focusing on individual cases or a small group of cases. We provide what appears to be the first quantitative analysis, using unique data on judicial claims made to enforce access to medical care. Our study is particularly relevant in the context of sharp expansions in health care provision across countries, enhanced by a global push for universal health coverage. Where the incentives of providers are not necessarily aligned with constitutional commitments, questions of effective delivery and accountability become increasingly important.

9 Conclusion

Our results show that supply-side restrictions on access to health care services manifest as substantial reductions in medical consultations, hospitalizations, emergency care, and medical procedures. In line with this, we identify increases in mortality rates. The estimates are robust to varying the estimator and the measure of restrictions (tutela vs. complaint rates). There is no evidence that citizens migrate across insurers in response to restrictions, mitigating concerns over dynamic selection. The estimates are robust to allowing for mean reversion, and for other relevant time-varying covariates. They are pervasive across social groups defined by economic and demographic characteristics, and across causes of death. They do not vary by insurer-concentration indices.

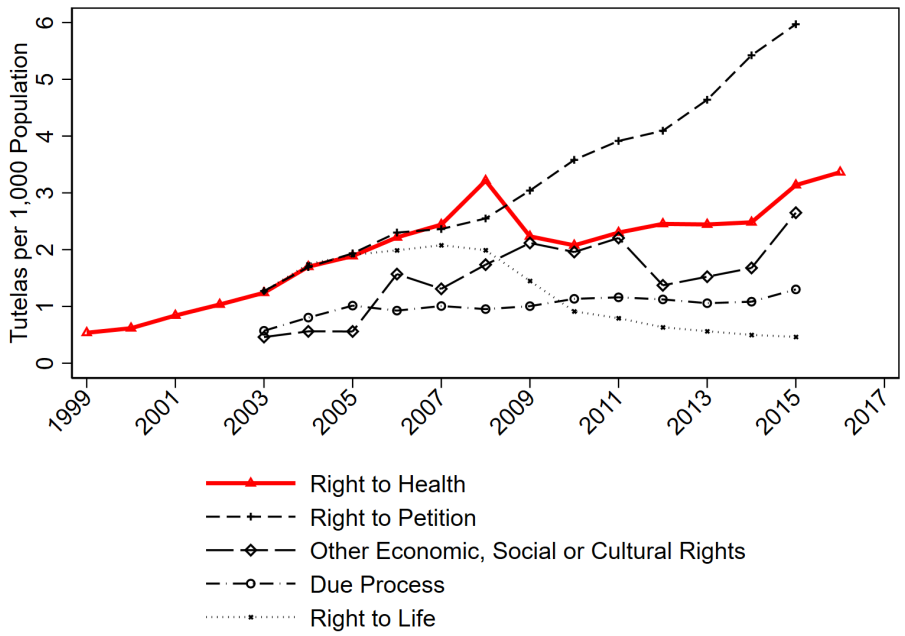
Our results are relevant to emerging health care insurance systems in many developing countries. Identifying an effective design for these new regimes is important because public budgets are tight, and health and economic status are more intricately tied among the poor. Our results are also more widely relevant as richer countries increasingly have to grapple with rising demand that current provisions are, in general, unable to satisfy.

Colombia has made a constitutional commitment to the right to health and acted to implement universal health care coverage ahead of many other countries. It has in place a well-functioning, widely accessed institution designed to protect citizen rights. However, the organization of the delivery of health care services, which operates through insurance companies (mostly but not entirely privately owned), has design flaws. There is too much regulation insofar as insurers are not allowed to set prices or select enrollees and, at the same time, too little regulation insofar as insurers are not penalized when complaints against them are received and upheld. Our estimates provide a lower bound on the population-level health gains that can be achieved through re-designing the system.

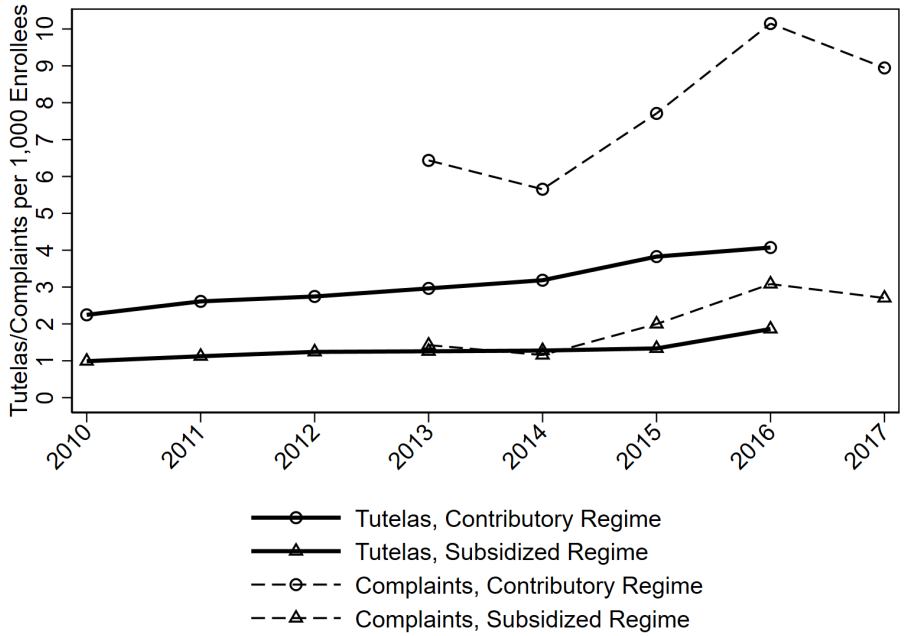
Tables and Figures

Figure 1: Trends in Tutelas and Complaints

(a) Tutelas per 1,000 Inhabitants by Fundamental Right Invoked



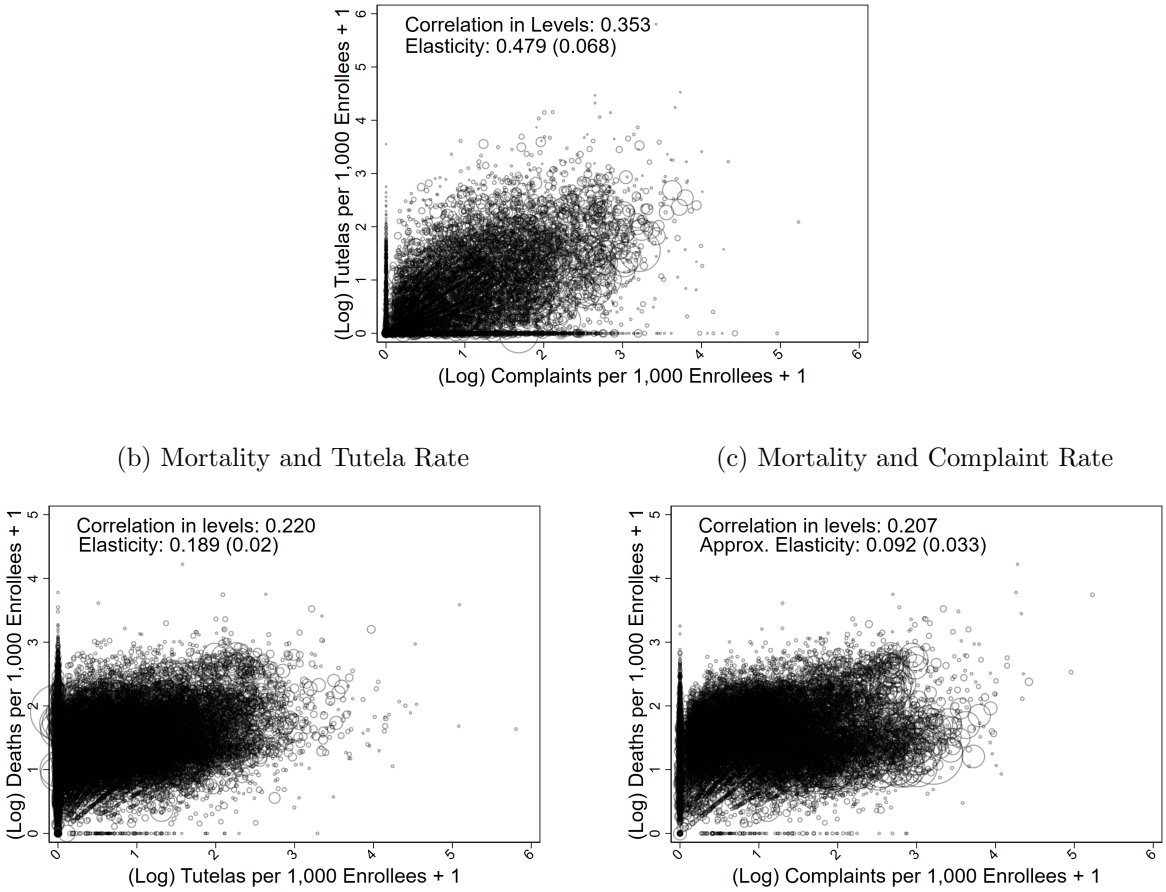
(b) Complaints and Tutelas Against Health Insurers per 1,000 Enrollees by Regime



Notes: The numbers reported in the right-to-health series of Panel (a) include complaints made against other actors in the system different from insurers (i.e., IPSs, Secretary of Health and Social Protection, local authorities). See footnote 20 for a discussion on the peak of health related tutelas in 2008. The numbers in Panel (b) are based on Defensoría del Pueblo (2004, 2007, 2009, 2010, 2011, 2012, 2013, 2015), and include only complaints and tutelas made against insurers.

Figure 2: Correlations between Mortality, Tutela and Complaint Rates

(a) Tutela and Complaint Rate

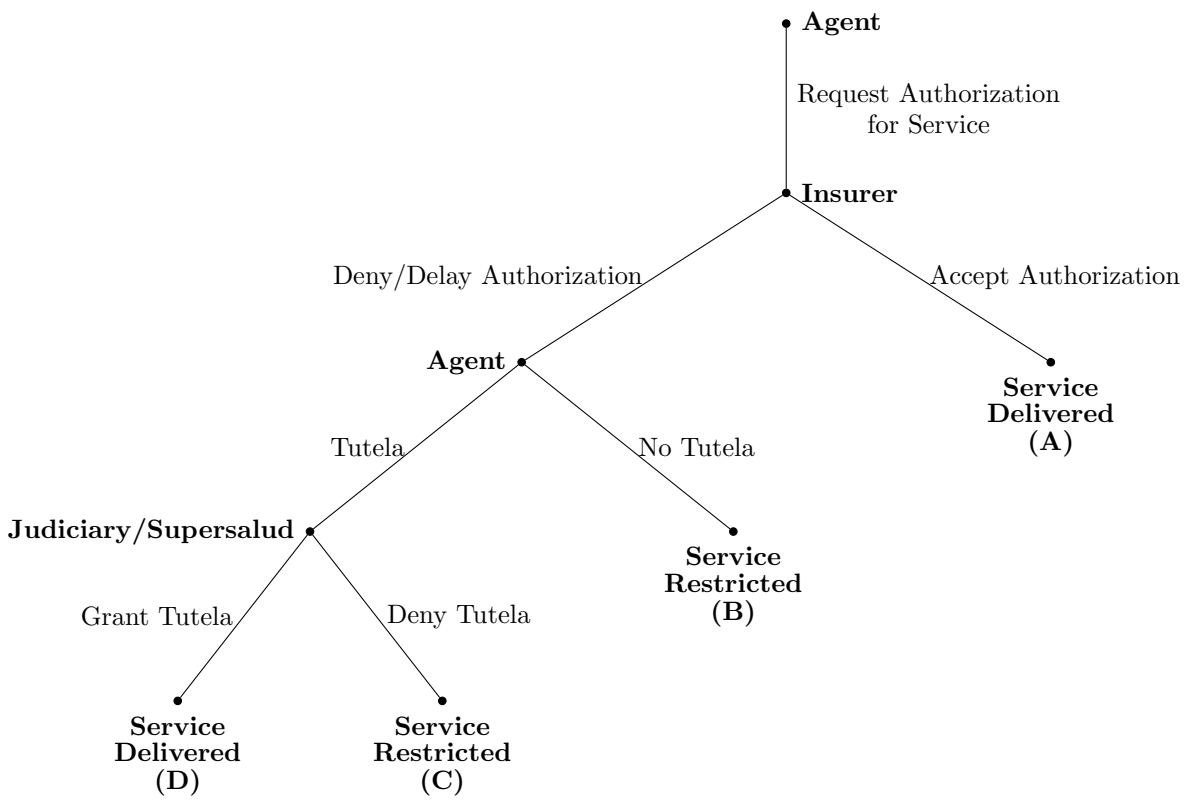


(b) Mortality and Tutela Rate

(c) Mortality and Complaint Rate

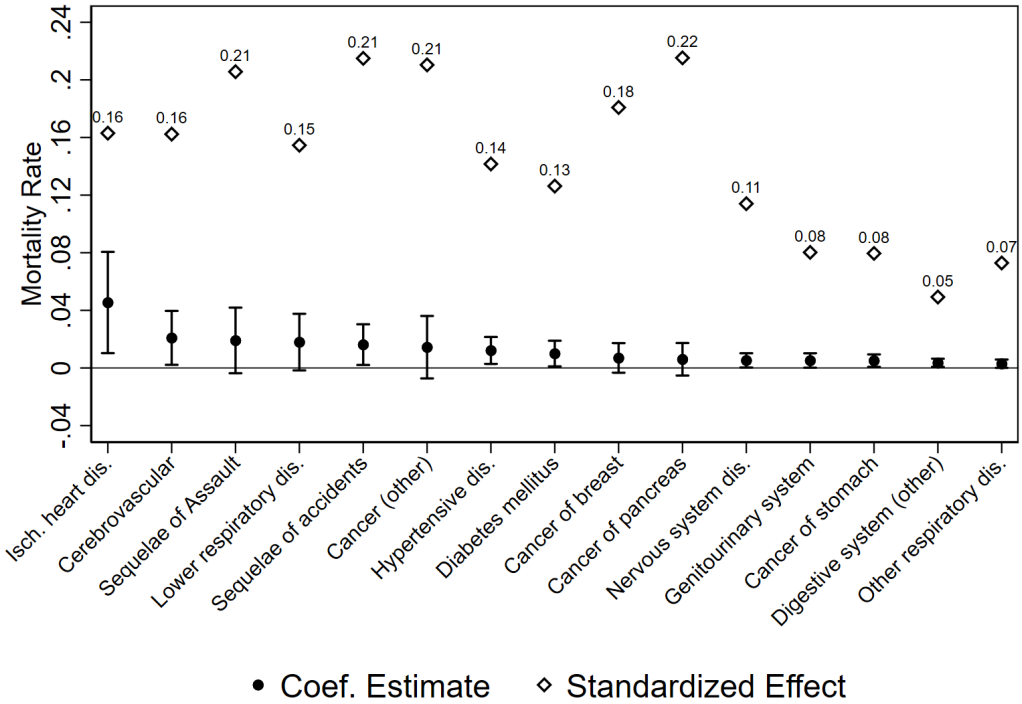
Notes: The scatter plot in Panel (a) shows the relation between the (log) of the tutela and (log) complaint rates (plus one) defined at the municipality × insurer × year level. The scatter plot in Panel (b) shows the relation between the (log) of the tutela and (log) mortality rates (plus one) defined at the same level. Panel (c) repeats the exercise but using the complaint rate. The size of each circle is proportional to the total share of enrollees. The approximate elasticity is measured by the estimated slope of a regression of the respective variables. The sources of data to construct the figure are described in Section 3.

Figure 3: Decision Tree



The figure depicts the relevant features of the decision of the insurer to deny authorized medical services, the decision of the citizen to litigate against the denial, and the decision of the judiciary to uphold the denial or not. The decision-making is discussed in Section 4 and formalized in Appendix Section A. An empirical model designed to identify supply-side restrictions or unmet demand from the observables in the data is set out in Section 5.

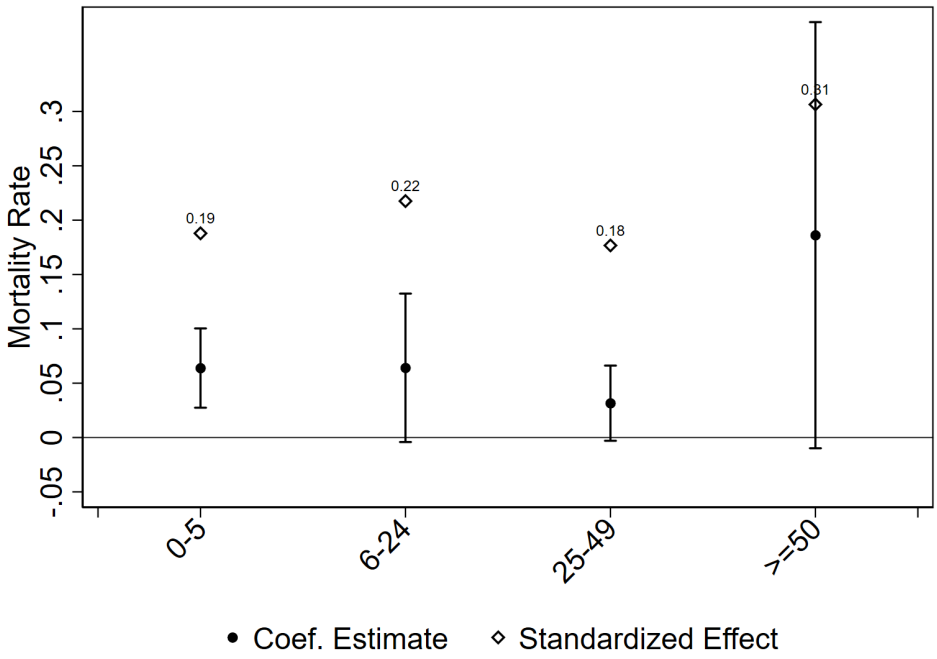
Figure 4: Effect of Restrictions on Access to Medical Care on Mortality Rates **by Cause of Death**



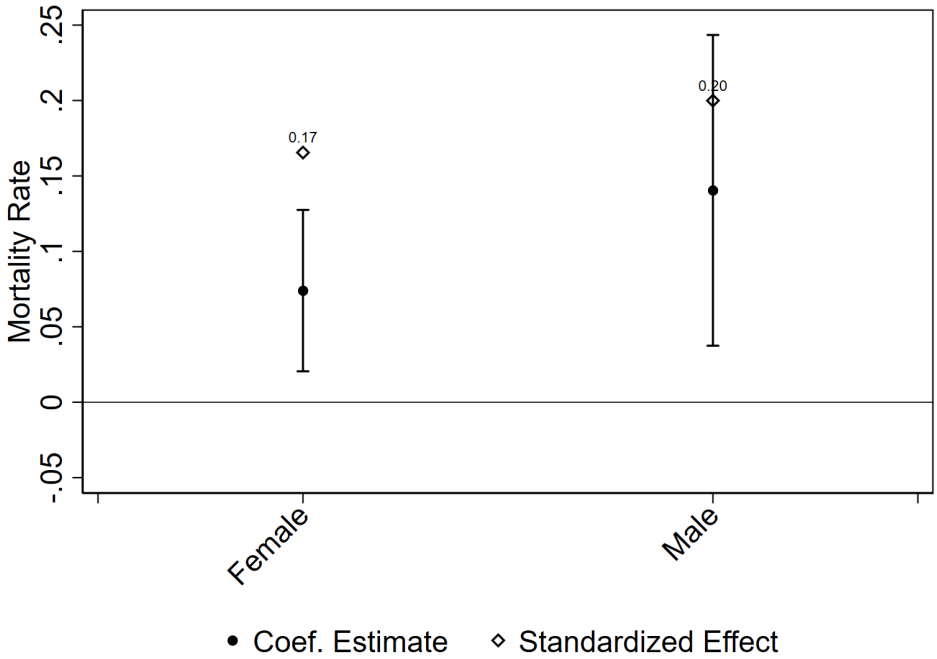
Notes: The Figure shows the estimated effect of supply-side restrictions to access health care on cause-specific mortality rates. Restrictions are measured using the municipality × insurer × year tutela rate. The estimates correspond to the instrumental variable specification discussed in Section 5.1. The standardized effect is calculated as the effect of doubling the average tutela rate on the cause-specific mortality rate, divided by its standard deviation.

Figure 5: Effect of Restrictions on Access to Medical Care on Mortality Rates **by Age and Sex**

(a) By Age



(b) By Sex



Notes: The Figure shows the estimated effect of supply-side restrictions to access to health care on age-specific (Panel (a)) and sex-specific (Panel (b)) mortality rates. The model specification uses complaint rates since the tutela data can not be disaggregated by age or sex. The standardized effect is calculated as the effect of doubling the average complaint rate on the age or sex-specific mortality rate, divided by its standard deviation.

Table 1: Effect of Restrictions on Access to Medical Care on Mortality Rates: Three-Way Fixed Effects IV Model

	Dep. Var.: Mortality Rate Subsidized and Contributory Regime							
	OLS (a)	IV (b)	OLS (c)	IV (d)	OLS (e)	IV (f)	OLS (g)	IV (h)
Tutela Rate	0.164***	0.283***	0.126***	0.327***				
	(0.031)	(0.068)	(0.029)	(0.080)				
Complaint Rate					0.210***	0.107*	0.137***	0.113*
					(0.028)	(0.060)	(0.014)	(0.063)
Controls								
Mortality Rate (t-1)			-0.005	-0.006			-0.000	0.000
			(0.004)	(0.004)			(0.000)	(0.000)
Insurer Market Share in Municipality			-0.981***	-0.867**			-0.284***	-0.267
			(0.022)	(0.339)			(0.033)	(0.273)
Δ Insurer Market Share in Municipality			0.594**	0.309			-0.041	0.048
			(0.176)	(0.311)			(0.146)	(0.137)
Observations	74,171	70,732	52,306	49,243	51,939	48,909	41,319	38,557
Fixed Effecs								
Municipality × Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Insurer × Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality × Insurer	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First Stage								
F stat. First Stage		28.43		19.77		25.32		22.36
Standardized Effect	0.08	0.13	0.06	0.15	0.22	0.11	0.14	0.12
Summary Stats.								
Av. Mortality Rate	3.87							
	(4.61)							
Av. Tutela Rate	2.16							
	(6.29)							
Av. Complaint Rate	4.84							
	(7.18)							

*** 1 percent ** 5 percent * 10 percent.

Notes: The Table shows the estimated effect of supply-side restrictions to access health care on mortality rates. The IV estimates correspond to the three-way fixed effects IV model discussed in Section 5.1. All rates are defined per 1,000 enrollees, and each regression includes analytical weights proportional to the number of enrollees in the insurer × municipality × year. We report cluster-robust standard errors, clustered by municipality and insurer. The standardized effect is calculated as the effect of doubling the average tutela or complaint rate on the mortality rate, divided by its standard deviation.

Table 2: Effect of Restrictions on Access to Medical Care on Mortality Rates: **Shift-Share Approach**

	Dep. Var.: Mortality Rate Subsidized and Contributory Regime							
	OLS (a)	IV (b)	OLS (c)	IV (d)	OLS (e)	IV (f)	OLS (g)	IV (h)
Tutela Rate	0.237***	0.486***	0.101***	0.196***				
	(0.023)	(0.074)	(0.013)	(0.046)				
Complaint Rate					0.053***	0.104***	0.053***	0.118***
					(0.011)	(0.017)	(0.013)	(0.014)
Observations	7,795	7,795	6,587	6,587	5,544	5,544	5,508	5,508
Fixed Effecs								
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First Stage								
F stat. First Stage		12.13		16.22		84.47		66.68
Standardized Effect	0.11	0.23	0.05	0.09	0.06	0.11	0.06	0.12
Controls: lagged utilization and lagged mortality	No	No	Yes	Yes	No	No	Yes	Yes

*** 1 percent ** 5 percent * 10 percent.

Notes: The Table shows the estimated effect of supply-side restrictions to access health care on mortality rates. The estimates in this table use municipality-year data that aggregates over insurers. The IV estimates correspond to the shift-share IV discussed in Section 5.2. All rates are defined per 1,000 enrollees, and each regression includes analytical weights proportional to the number of enrollees in the municipality $\times$ year. We report cluster-robust standard errors clustered by municipality. The standardized effect is calculated as the effect of doubling the average tutela or complaint rate on the mortality rate, divided by its standard deviation.

Table 3: Effect of Restrictions on Access to Medical Care on Mortality Rates: Three-Way Fixed Effects Placebo and Dynamics Model

	Dep. Var.: Mortality Rate Subsidized and Contributory Regime							
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Tutela Rate (t+1)	-0.0004	-0.0097						
	(0.0003)	(0.0129)						
Tutela Rate (t-1)			0.0013	0.0015				
			(0.0013)	(0.0013)				
Complaint Rate (t+1)					0.0003	0.0002		
					(0.0004)	(0.0004)		
Complaint Rate (t-1)							-0.0001	-0.0003
							(0.0002)	(0.0005)
Observations	52,361	40,012	63,153	63,109	41,352	34,992	32,554	32,533
Fixed Effecs								
Municipality × Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Insurer × Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality × Insurer	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Standardized Effect	-0.00	-0.00	0.00	0.00	0.00	0.00	-0.00	-0.00
Controls: lagged mortality, market share, and change in market share	No	Yes	No	Yes	No	Yes	No	Yes

*** 1 percent ** 5 percent * 10 percent.

Notes: The Table shows the estimated effect of supply-side restrictions to access health care on mortality rates, using either lagged or future restrictions instead of contemporary ones. The estimates correspond to the three-way fixed effects model discussed in Section 5.1. All rates are defined per 1,000 enrollees, and each regression includes analytical weights proportional to the number of enrollees in the insurer × municipality × year. We report cluster-robust standard errors, clustered by municipality and insurer. The standardized effect is calculated as the effect of doubling the average tutela or complaint rate on the mortality rate, divided by its standard deviation.

Table 4: Dynamic Selection: Restrictions on Access to Medical Care and Probability of Changing Insurers

	Dep. Var.: =1 if insurer changed between *t* and *t-1*, =0 otherwise					
	Subsidized and Contributory	Subsidized and Contributory	Contributory	Contributory	Subsidized	Subsidized
Tutela Rate (t-1)	0.00039		0.00003		0.00092	
	(0.00022)		(0.00022)		(0.00075)	
Complaint Rate (t-1)		0.00036		0.00010		0.00088
		(0.00019)		(0.00023)		(0.00068)
Observations	163525077	131963216	89,968,529	72,008,004	70,177,496	56,632,724
Summary Stats.						
Share Insurer Changed	0.14	0.14	0.14	0.14	0.14	0.14
	(0.35)	(0.35)	(0.34)	(0.34)	(0.35)	(0.35)
Fixed Effecs						
Enrollee	Yes	Yes	Yes	Yes	Yes	Yes
Init. Insurer × Year	Yes	Yes	Yes	Yes	Yes	Yes
Init. Municipality × Init. Insurer	Yes	Yes	Yes	Yes	Yes	Yes

*** 1 percent ** 5 percent * 10 percent.

Notes: The Table shows the estimated effect of supply-side restrictions on the probability of changing insurers. The dependent variable is an indicator variable equal to one if the individual changed insurer between year t and $t-1$. The regressor of interest is the tutela or complaint rate at year $t-1$. Each regression includes individual fixed effects and initial insurer × year and initial municipality × year flexible trends, whereby "initial" we mean the insurer or municipality of the individual when she first appeared in the data. We report cluster-robust standard errors clustered by municipality and insurer.

Table 5: Dynamic Selection: Restrictions on Access to Medical Care and Probability of Switching Insurers, **by Sex and Age**

	Dep. Var.: =1 if insurer changed between *t* and *t-1*, =0 otherwise					
	Female ages 0-4	Female ages 5-18	Female ages 19-44	Female ages 45-59	Female ages 60-69	Female ages $\geq$ 70
Tutela Rate (t-1)	0.00060	0.00038	0.00044	0.00042	0.00034	0.00022
	(0.00035)	(0.00027)	(0.00024)	(0.00023)	(0.00020)	(0.00017)
Observations	3,646,451	16,479,870	37,954,156	17,316,501	7,067,446	7,340,276
Summary Stats.						
Share Insurer Changed	0.15	0.13	0.17	0.12	0.09	0.07
	(0.36)	(0.34)	(0.38)	(0.32)	(0.28)	(0.25)
Fixed Effecs						
Enrollee	Yes	Yes	Yes	Yes	Yes	Yes
Init. Insurer × Year	Yes	Yes	Yes	Yes	Yes	Yes
Init. Municipality × Init. Insurer	Yes	Yes	Yes	Yes	Yes	Yes

	Dep. Var.: =1 if insurer changed between *t* and *t-1*, =0 otherwise					
	Male ages 0-4	Male ages 5-18	Male ages 19-44	Male ages 45-59	Male ages 60-69	Male ages $\geq$ 70
Tutela Rate (t-1)	0.00062	0.00035	0.00037	0.00043	0.00027	0.00020
	(0.00028)	(0.00045)	(0.00039)	(0.00026)	(0.00018)	(0.00015)
Observations	3,711,203	15,993,053	27,538,410	12,498,595	5,036,638	4,853,376
Summary Stats.						
Share Insurer Changed	0.15	0.13	0.18	0.14	0.11	0.08
	(0.36)	(0.33)	(0.38)	(0.35)	(0.31)	(0.26)
Fixed Effecs						
Enrollee	Yes	Yes	Yes	Yes	Yes	Yes
Init. Insurer × Year	Yes	Yes	Yes	Yes	Yes	Yes
Init. Municipality × Init. Insurer	Yes	Yes	Yes	Yes	Yes	Yes

*** 1 percent ** 5 percent * 10 percent.

Notes: The Table shows the estimated effect of supply-side restrictions on the probability of changing insurers, dividing the sample by sex and age groups. The dependent variable is an indicator variable equal to one if the individual changed insurer between year t and $t-1$. The regressor of interest is the tutela or complaint rate at year $t-1$. Each regression includes individual fixed effects and initial insurer × year and initial municipality × year flexible trends, whereby "initial" we mean the insurer or municipality of the individual when she first appeared in the data. We report cluster-robust standard errors clustered by municipality and insurer.

Table 6: Effect of Restrictions on Access to Medical Care on Mortality Rates **by Regime**: Three-Way Fixed Effects IV Model

	Contributory Regime				Subsidized Regime			
	OLS (a)	IV (b)	OLS (c)	IV (d)	OLS (e)	IV (f)	OLS (g)	IV (h)
Tutela Rate	0.114***	0.334**			0.238***	0.386***		
	(0.024)	(0.124)			(0.061)	(0.032)		
Complaint Rate			0.149***	0.099			0.074**	0.484
			(0.014)	(0.063)			(0.026)	(0.294)
Observations	32,559	32,559	22,781	22,456	18,463	14,995	17,579	14,714
Fixed Effecs								
Municipality × Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Insurer × Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality × Insurer	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First Stage								
F stat. First Stage		7.38		19.6		159.76		11.31
Standardized Effect	0.07	0.19	0.22	0.14	0.08	0.14	0.04	0.27
Summary Stats.								
Av. Mortality Rate	3.55				4.17			
	(5.41)				(3.71)			
Av. Tutela Rate	3.12				1.31			
	(7.67)				(4.55)			
Av. Complaint Rate	7.83				2.08			
	(8.65)				(3.76)			

*** 1 percent ** 5 percent * 10 percent.

Notes: The Table shows the estimated effect of supply-side restrictions to access health care on mortality rates, conditional on the regime. The IV estimates correspond to the three-way fixed effects IV model discussed in Section 5.1. All rates are defined per 1,000 enrollees, and each regression includes analytical weights proportional to the number of enrollees in the insurer × municipality × year. We report cluster-robust standard errors, clustered by municipality and insurer. The standardized effect is calculated as the effect of doubling the average tutela or complaint rate on the mortality rate, divided by its standard deviation.

Table 7: Effect of Restrictions on Access to Medical Care on Utilization of Medical Services I: Shift-Share Approach

	Consultations per Enrollee				Procedures per Enrollee			
	OLS (a)	IV (b)	OLS (c)	IV (d)	OLS (e)	IV (f)	OLS (g)	IV (h)
Tutela Rate	-0.22***	-1.06**			-0.15**	-0.17**		
	(0.04)	(0.43)			(0.05)	(0.08)		
Complaint Rate			-0.22***	-0.56***			-0.06**	-0.29***
			(0.02)	(0.08)			(0.02)	(0.07)
Observations	6,587	6,587	5,508	5,508	6,583	6,583	5,499	5,499
Av. Dependent Variable	2.019				1.549			
	(1.160)				(1.123)			
Fixed Effecs								
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First Stage								
F stat. First Stage		16.22		66.68		16.22		66.68
Standardized Effect	-0.19	-0.91	-0.13	-0.34	-0.13	-0.15	-0.04	-0.19
Controls: lagged utilization and lagged mortality	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*** 1 percent ** 5 percent * 10 percent.

Notes: The Table shows the estimated effect of supply-side restrictions to access health care on the utilization of care services. The estimates in this table use municipality-year data that aggregates over insurers. The IV estimates correspond to the shift-share IV discussed in Section 5.2. Each regression includes analytical weights proportional to the number of enrollees in the municipality $\times$ year. We report cluster-robust standard errors clustered by municipality. The standardized effect is calculated as the effect of doubling the average tutela or complaint rate on the utilization rate, divided by its standard deviation.

Table 8: Effect of Restrictions on Access to Medical Care on Utilization of Medical Services II: Shift-Share Approach

	Hospitalizations per Enrollee				Emergencies per Enrollee			
	OLS (a)	IV (b)	OLS (c)	IV (d)	OLS (e)	IV (f)	OLS (g)	IV (h)
Tutela Rate	-0.005***	-0.017**			-0.013***	-0.063**		
	(0.001)	(0.005)			(0.003)	(0.025)		
Complaint Rate			-0.004***	-0.012***			-0.016***	-0.037***
			(0.001)	(0.002)			(0.003)	(0.006)
Observations	6,581	6,581	5,464	5,464	6,583	6,583	5,497	5,497
Av. Dependent Variable	0.036				0.112			
	(0.024)				(0.097)			
Fixed Effecs								
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First Stage								
F stat. First Stage		16.22		66.84		16.21		67.07
Standardized Effect	-0.20	-0.71	-0.14	-0.41	-0.13	-0.65	-0.15	-0.33
Controls: lagged utilization and lagged mortality	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*** 1 percent ** 5 percent * 10 percent.

Notes: The Table shows the estimated effect of supply-side restrictions to access health care on the utilization of care services. The estimates in this table use municipality-year data that aggregates over insurers. The IV estimates correspond to the shift-share IV discussed in Section 5.2. Each regression includes analytical weights proportional to the number of enrollees in the municipality $\times$ year. We report cluster-robust standard errors clustered by municipality. The standardized effect is calculated as the effect of doubling the average tutela or complaint rate on the utilization rate, divided by its standard deviation.

References

Abadia, C. E., Oviedo, D. G., Abadía, E., and Oviedo, D. G. (2009). Bureaucratic Itineraries in Colombia. A theoretical and methodological tool to assess managed-care health care systems. *Social Science and Medicine*, 68(6):1153–1160.

Alston, P. (2010). The General Comments of the UN Committee on Economic, Social and Cultural Rights. Doc. E/C.12/2000/21.

Andia, T. S. and Lamprea, E. (2019). Is the judicialization of health care bad for equity? A scoping review. *International Journal for Equity in Health*, 18(1):61.

Arrieta-Gómez, A. I. (2018). Realizing the fundamental right to health through litigation: The Colombian case. *Health and Human Rights*, 20(1):133–145.

Backman, G., Hunt, P., Khosla, R., Jaramillo-Strouss, C., Fikre, B. M., Rumble, C., Pevalin, D., Páez, D. A., Pineda, M. A., Frisancho, A., Tarco, D., Motlagh, M., Farcasanu, D., and Vladescu, C. (2008). Health systems and the right to health: an assessment of 194 countries. *The Lancet*, 372(9655):2047–2085.

Baicker, K., Taubman, S. L., Allen, H. L., Bernstein, M., Gruber, J. H., Newhouse, J. P., Schneider, E. C., Wright, B. J., Zaslavsky, A. M., and Finkelstein, A. N. (2013). The Oregon Experiment — Effects of Medicaid on Clinical Outcomes. *New England Journal of Medicine*, 368(18):1713–1722.

Banerjee, A., Finkelstein, A., Hanna, R., Olken, B. A., Ornaghi, A., and Sumarto, S. (2021). The Challenges of Universal Health Insurance in Developing Countries: Experimental Evidence from Indonesia's National Health Insurance. *American Economic Review*, 111(9):3035–3063.

Banerjee, R., Ziegenfuss, J. Y., and Shah, N. D. (2010). Impact of discontinuity in health insurance on resource utilization. *BMC Health Services Research*, 10(1):195.

Bergallo, P. (2011). Argentina. In Yamin, A. E. and Gloppen, S., editors, *Litigating Health Rights*, chapter 3, pages 43–75. Harvard University Press.

Biehl, J., Amon, J. J., Socal, M. P., and Petryna, A. (2012). Between the Court and the Clinic: Lawsuits for Medicines and the Right to Health in Brazil. *Health Human Rights*, 14(1):E36–52.

Biehl, J., Petryna, A., Gertner, A., Amon, J. J., and Picon, P. D. (2009). Judicialisation of the right to health in Brazil. *The Lancet*, 373(9682):2182–2184.

Bloom, D. E., Khoury, A., and Subbaraman, R. (2018). The promise and peril of universal health care. *Science*, 361(6404):eaat9644.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-Experimental Shift-Share Research Designs. *The Review of Economic Studies*, 89(1):181–213.

Buitrago, G., Miller, G., and Vera-Hernández, M. (2021). Cost-Sharing in Medical Care Can Increase Adult Mortality Risk in Lower-Income Countries. *medRxiv Working Paper*.

Camacho, A. and Conover, E. (2013). Effects of subsidized health insurance on newborn

health in a developing country. *Economic Development and Cultural Change*, 61(3):633–658.

Card, D., Dobkin, C., and Maestas, N. (2008). The impact of nearly universal insurance coverage on health care utilization: Evidence from medicare. *American Economic Review*, 98(5):2242–2258.

CBO (2004). The Effects of Tort Reform: Evidence from the States. Technical report, Congressional Budget Office.

Chou, S.-Y. Y., Grossman, M., and Liu, J.-T. T. (2014). The impact of National Health Insurance on birth outcomes: A natural experiment in Taiwan. *Journal of Development Economics*, 111:75–91.

Cooper, C. (2011). Health Rights Litigation: Cautious Constitutionalism. In Yamin, A. E. and Gloppen, S., editors, *Litigating Health Rights: Can Courts Bring more Justice to Health?*, chapter 8, pages 190–229. Harvard University Press.

Corte Constitucional (2008). SENTENCIA N° T-760 de 2008. July 31, 2008. Magistrado Ponente: Manuel José Cepeda.

Currie, J. and Gruber, J. (1996a). Health Insurance Eligibility, Utilization of Medical Care, and Child Health. *The Quarterly Journal of Economics*, 111(2):431–466.

Currie, J. and Gruber, J. (1996b). Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women. *Journal of Political Economy*, 104(6):1263–1296.

Currie, J., Gruber, J., and Fischer, M. (1994). Physician Payments and Infant Mortality: Evidence from Medicaid Fee Policy. Technical Report 2, National Bureau of Economic Research, Cambridge, MA.

Currie, J. and MacLeod, W. B. (2008). First Do No Harm? Tort Reform and Birth Outcomes. *The Quarterly Journal of Economics*, 123(2):795–830.

Cutler, D. M. and Ly, D. P. (2011). The (Paper)Work of Medicine: Understanding International Medical Costs. *Journal of Economic Perspectives*, 25(2):3–25.

Danzon, P. M. (2000). Liability for medical malpractice. volume 1 of *Handbook of Health Economics*, pages Chapter 26: pp 1339 – 1404. Elsevier.

Defensoría del Pueblo (2004). La Tutela y el Derecho a la Salud: 1999-2003. Reporte Institucional.

Defensoría del Pueblo (2007). La tutela y el derecho a la salud periodo 2003–2005. Reporte Institucional.

Defensoría del Pueblo (2009). La Tutela y el Derecho a la Salud: 2006-2008. Reporte Institucional.

Defensoría del Pueblo (2010). La Tutela y el Derecho a la Salud 2009. Reporte Institucional.

Defensoría del Pueblo (2011). La Tutela y el Derecho a la Salud 2010. Reporte Institucional.

Defensoría del Pueblo (2012). La Tutela y el Derecho a la Salud 2011. Reporte Institucional.

Defensoría del Pueblo (2013). La Tutela y el Derecho a la Salud 2012. Reporte Institucional.

Defensoría del Pueblo (2015). La Tutela y los Derechos a la Salud y a la Seguridad Social 2014. Reporte Institucional.

Dillender, M. (2018). What happens when the insurer can say no? Assessing prior authorization as a tool to prevent high-risk prescriptions and to lower costs. *Journal of Public Economics*, 165:170–200.

Duggan, M., Gupta, A., Jackson, E., and Templeton, Z. S. (2023). The Impact of Privatization: Evidence from the Hospital Sector. *NBER*.

Dunn, A., Gottlieb, J., Shapiro, A., Sonnenstuhl, D., and Tebaldi, P. (2021). A Denial a Day Keeps the Doctor Away. Technical report, National Bureau of Economic Research, Cambridge, MA.

Einav, L. and Levin, J. (2015). Managed Competition in Health Insurance. *Journal of the European Economic Association*, 13(6):998–1021.

Enthoven, A. C. (1978). Consumer-Choice Health Plan. *New England Journal of Medicine*, 298(12):650–658.

Feldstein, M. (1977). Quality Change and the Demand for Hospital Care. *Econometrica*, 45(7):1681–1702.

Felland, L. E., Lechner, A., and Sommers, A. (2013). Access to Specialty Care for Medicaid Patients: Policy Issues and Options. Technical report, The Commonwealth Fund.

Finkelstein, A. (2007). The aggregate effects of health insurance: Evidence from the introduction of medicare. *Quarterly Journal of Economics*, 122(1):1–37.

Finkelstein, A. and McKnight, R. (2008). What did Medicare do? The initial impact of Medicare on mortality and out of pocket medical spending. *Journal of Public Economics*, 92(7):1644–1668.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., and Baicker, K. (2012). The oregon health insurance experiment: Evidence from the first year. *Quarterly Journal of Economics*, 127(3):1057–1106.

Fox, J. and Stephenson, M. C. (2011). Judicial Review as a Response to Political Posturing. *American Political Science Review*, 105(2):397–414.

Frakes, M. (2013). The impact of medical liability standards on regional variations in physician behavior: Evidence from the adoption of national-standard rules. *American Economic Review*, 103(1):257–276.

Frakes, M. and Jena, A. B. (2016). Does medical malpractice law improve health care quality? *Journal of Public Economics*, 143:142–158.

Gauri, V. (2004). Social Rights and Economics: Claims to Health Care and Education in Developing Countries. *World Development*, 32(3):465–477.

Gaviria, A., Medina, C., Mejía, C., McKenzie, D., and Soares, R. R. (2006). Assessing Health Reform in Colombia: From Theory to Practice. *Economía*, 7(1):29–72.

Gaynor, M., Ho, K., and Town, R. J. (2015). The Industrial Organization of Health-Care Markets. *Journal of Economic Literature*, 53(2):235–284.

Gaynor, M., Rebitzer, J. B., and Taylor, L. J. (2004). Physician Incentives in Health Maintenance Organizations. *Journal of Political Economy*, 112(4):915–931.

Glassman, A. L., Escobar, M.-L., Giuffrida, A., Giedion, U., and Giedion, A. (2009). From few to many : ten years of health insurance expansion in Colombia. Technical report, Inter-American Development Bank and The Brookings Institution.

Goldin, J., Lurie, I. Z., and McCubbin, J. (2020). Health Insurance and Mortality: Experimental Evidence from Taxpayer Outreach. *The Quarterly Journal of Economics*, 136(1):1–49.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*.

Goodman-Bacon, A. (2018). Public Insurance and Mortality: Evidence from Medicaid Implementation. *Journal of Political Economy*, 126(1):216–262.

Government Accountability Office (2011). Medicaid and CHIP: Most Physicians Serve Covered Children but Have Difficult Referring Them for Specialty Care. Technical Report June, Government Accountability Office.

Gruber, J., Hendren, N., and Townsend, R. M. (2014). The great equalizer: Health care access and infant mortality in Thailand. *American Economic Journal: Applied Economics*, 6(1 A):91–107.

Ho, K. and Lee, R. S. (2017). Insurer Competition in Health Care Markets. *Econometrica*, 85(2):379–417.

Kessler, D. and McClellan, M. (1996). Do doctors practice defensive medicine? *The Quarterly Journal of Economics*, 111(2):353–390.

Kessler, D. and McClellan, M. (2002). Malpractice law and health care reform: Optimal liability policy in an era of managed care. *Journal of Public Economics*, 84(2):175–197.

Kinney, E. D. and Clark, B. A. (2004). Provisions for health and health care in the constitutions of the countries of the world. *Cornell International Law Journal*, 37(2):285–355.

Kruk, M. E., Gage, A. D., Joseph, N. T., Danaei, G., García-Saisó, S., and Salomon, J. A. (2018). Mortality due to low-quality health systems in the universal health coverage era: a systematic analysis of amenable deaths in 137 countries. *The Lancet*, 392(10160):2203–2212.

Londoño, J.-L. and Frenk, J. (1997). Structured pluralism: towards an innovative model for health system reform in Latin America. *Health Policy*, 41(1):1–36.

MAG (1994). Access of Medicaid Recipients to Outpatient Care. *New England Journal of Medicine*, 330(20):1426–1430.

Manning, W. G., Newhouse, J. P., Duan, N., Keeler, E. B., Leibowitz, A., and Marquis, M. S. (1987). Health insurance and the demand for medical care: evidence from a randomized experiment. *American Economic Review*, 77(3):251–277.

Martins, A. and Allen, S. (2020). Litigation to Access Health Services: Ally or Enemy of Global Public Health? *Annals of Global Health*, 86(1).

Maskin, E. and Tirole, J. (2004). The Politician and the Judge: Accountability in Government. *American Economic Review*, 94(4):1034–1054.

Maya, E. (2008). El derecho a la salud en la perspectiva de los derechos humanos y del sistema de inspección, vigilancia y control de quejas en materia de salud *. Technical Report 4, Procuraduría General de la Nación.

Miller, G., Pinto, D., and Vera-Hernández, M. (2013). Risk Protection, Service Use, and Health Outcomes under Colombia's Health Insurance Program for the Poor. *American Economic Journal: Applied Economics*, 5(4):61–91.

Miller, S., Johnson, N., and Wherry, L. R. (2021). Medicaid and Mortality: New Evidence From Linked Survey and Administrative Data. *The Quarterly Journal of Economics*, 136(3):1783–1829.

Ministerio de Salud y Proteccion Social (2018). Informe de Cumplimiento de la Vigencia 2017 Orden 30 de la Sentencia T-760 DE 2008. Reporte Institucional Ministerio de Salud y Protección Social.

Mohanan, M., Bauhoff, S., La Forgia, G., Babiarz, K. S., Singh, K., and Miller, G. (2014). Effect of Chiranjeevi Yojana on institutional deliveries and neonatal and maternal outcomes in Gujarat, India: a difference-in-differences analysis. *Bulletin of the World Health Organization*, 92(3):187–194.

Monge, A. (2019). Litigation and access to health care: evidence from 3,124 cases. Technical report, Cornell University.

Norheim, O. F., Wilson, B. M., and Wilson, B. M. (2014). Health rights litigation and access to medicines: Priority classification of successful cases from Costa Rica's constitutional chamber of the supreme court. *Health and Human Rights*, 16(2):47–61.

Oostrom, T., Einav, L., and Finkelstein, A. (2017). Outpatient Office Wait Times And Quality Of Care For Medicaid Patients. *Health Affairs*, 36(5):826–832.

Parra-Vera, O. and Yamin, A. (2013). La sentencia T-760 de 2008, su implementación e impacto: retos y oportunidades para la justicia dialógica. *Clérico, L., Aldao, M., Ronconi, L(coords.), Tratado de derecho a la salud.*, pages 2591–2649.

Polsky, D., Richards, M., Basseyn, S., Wissoker, D., Kenney, G. M., Zuckerman, S., and Rhodes, K. V. (2015). Appointment Availability after Increases in Medicaid Payments for Primary Care. *New England Journal of Medicine*, 372(6):537–545.

Powell-Jackson, T., Mazumdar, S., and Mills, A. (2015). Financial incentives in health: New evidence from India's Janani Suraksha Yojana. *Journal of Health Economics*, 43:154–169.

Rodin, J. and de Ferranti, D. (2012). Universal health coverage: the third global health transition? *The Lancet*, 380(9845):861–862.

Rodríguez, C. (2012). La Judicialización de la Salud: Síntomas, Diagnóstico y Prescripciones. In Bernal, O. and Gutierrez, C., editors, *La salud en Colombia: logros retos y recomendaciones*, chapter 12, pages 507–560. Ediciones Uniandes-Universidad de los Andes.

Rold, W. J. (2008). Thirty years after Estelle v. Gamble: A legal retrospective. *Journal of Correctional Health Care*, 14(1):11–20.

Rosenbaum, S. (2014). Medicaid Payments and Access to Care. *New England Journal of Medicine*, 371(25):2345–2347.

Ruger, J. P., Ruger, T. W., and Annas, G. J. (2015). The Elusive Right to Health Care under U.S. Law. *New England Journal of Medicine*, 372(26):2558–2563.

Sommers, B. D., Baicker, K., and Epstein, A. M. (2012). Mortality and Access to Care among Adults after State Medicaid Expansions. *New England Journal of Medicine*, 367(11):1025–1034.

Sonchak, L. (2015). Medicaid Reimbursement, Prenatal Care and Infant Health. *Journal of Health Economics*, 44:10–24.

Vargas, I., Vázquez, M. L., Mogollón-Pérez, A. S., and Unger, J.-P. (2010). Barriers of access to care in a managed competition model: lessons from Colombia. *BMC Health Services Research*, 10(1):297.

Weathers, R. R. and Stegman, M. (2012). The effect of expanding access to health insurance on the health and mortality of Social Security Disability Insurance beneficiaries. *Journal of Health Economics*, 31(6):863–875.

Wilson, B. M. (2011). Costa Rica Health Rights Litigation: Causes and Consequences. In Yamin, A. E. and Gloppen, S., editors, *Litigating Health Rights: Can Courts Bring More Justice to Health?*, chapter 6, pages 132–154. Harvard University Press.

Yamin, A. E. and Parra-Vera, O. (2009). How Do Courts Set Health Policy? The Case of the Colombian Constitutional Court. *PLoS Medicine*, 6(2):e1000032.

Yamin, A. E. and Parra-Vera, O. (2010). Judicial Protection of the Right to Health in Colombia: From Social Demands to individual Claims to Public Debate. *Hastings International and Comparative Law Review*, 33(2):431–460.

For Online Publication

A Restrictions on Access to Health Care Services and Litigation: A Simple Model

In this section we develop a simple decision model that captures the salient features of the health care system in Colombia. The aim of the model is to make explicit the main factors that determine the decision of insurers to restrict access to a service, and the decision of enrollees to file tutelas or complaints. There are two periods. In the initial period agents observe their health status, and, if required, get prescribed a treatment, procedure or medication. These prescriptions can be authorized or denied/delayed (henceforth restricted) by the health insurer. Any medical service that is prescribed but not readily authorized is enforceable through litigation, but not all individuals choose to use legal instruments.

Agents. There is a continuum of agents, each enrolled with a health insurer. Suppose the initial endowment of health of an agent is given by the value of the variable $h_0 \in H = \{h_0^1 \dots h_0^H\}$, where larger values indicate better health. For each health endowment there is a unique corresponding medical service, $x \in X = \{x^1 \dots x^H\}$ that a health professional prescribes, equal across agents.[36] We assume h_0, or alternatively x,[37] has a probability distribution f that is common across agents. If a service x is prescribed and delivered, the health level in the final period is $h^x \geq h^{\neg x}$, where $h^{\neg x}$ is the health level of an agent that was prescribed x but did not receive it.[38] The assumption is that receiving medical care is at least as good as not getting it.

If an insurer denies an authorization, an agent can use litigation to enforce its access. To simplify, we assume tutelas are the only legal instrument available. An agent j that files a tutela incurs an idiosyncratic cost $\eta_j \in \mathbb{R}_+$, expressed in health units, which is independent of the service requested and the value of which is unobserved to the insurer. Filing a tutela has no pecuniary cost, so we think of η_j as capturing effort (the effort required to learn about the process) and psychic costs (the stress of filing the claim). We assume η_j has a known cumulative distribution function G across the population. A tutela requesting a service x is decided in favor of the plaintiff with an exogenous and known probability $\beta^x \in [0, 1]$. This probability can vary depending on the characteristics of the service requested, given that some restrictions more clearly contravene an individual's right to health. Finally, suppose that agents only care about their health and wish to maximize a utility function U, where $U' > 0$ and $U'' < 0$.

If an agent j faces a restriction on a service x she will find it optimal to file a tutela if

$$\beta^x U\left(h^x - \eta_j\right) + (1 - \beta^x) U\left(h^{\neg x} - \eta_j\right) \geq U\left(h^{\neg x}\right), \tag{A.1}$$

[36] Agents with the same health endowment get prescribed the same medical service.

[37] There is a one-to-one correspondence between H and X.

[38] We assume health evolves in a deterministic way, although adding a stochastic component delivers qualitatively similar results.

or, alternatively, if

$$\beta^x \geq \frac{U\left(h^{\neg x}\right) - U\left(h^{\neg x} - \eta_j\right)}{U\left(h^x - \eta_j\right) - U\left(h^{\neg x} - \eta_j\right)}. \tag{A.2}$$

The inequality states that as long as $\Delta h^x \equiv h^x - h^{\neg x} > \eta_j$, that is, if the health gain from receiving the service is larger than the idiosyncratic health cost, there will be a value of β^x for which the agent finds it optimal to file a tutela.

The ratio on the right hand side of inequality A.2 is an increasing function of η_j, which takes a minimum value of zero at $\eta_j = 0$. This implies that for a given $\beta^x > 0$, there is a lower threshold of the idiosyncratic cost, $\underline{\eta^x}$, such that all agents that get x restricted and have $\eta_j \leq \underline{\eta^x}$ will file a tutela. Hence, from the point of view of the insurer, the probability that an agent will file a tutela if x is restricted is $\alpha^x \equiv G(\underline{\eta^x})$. Note $\underline{\eta^x}$ is determined by two exogenous factors in the model: i). how much improvement in health an agent gets from receiving the service (Δh^x); and ii). the probability that litigation for x is decided in favor of the plaintiff (β^x).

Health Insurers. Insurers, indexed by $i \in \{1, ..., I\}$, receive a capitation payment y for each enrollee, which is exogenously determined. If an insurer authorizes a requested service x, it pays the full cost c_i^x. The costs of service provision can vary across insurers reflecting the fact that they can contract with their own independent network of providers. If they restrict a service, and the affected agent files a tutela that is upheld, the insurer pays the full cost of the service plus an additional idiosyncratic cost $\tilde{c}_i \in \mathbb{R}_+$. We think of this idiosyncratic cost as capturing factors like the administrative costs of dealing with the claims and general managerial efficiency. There are no additional penalties reflecting the fact that each case is treated individually and generates no precedent.

An insurer receiving a request to authorize a service x finds it optimal to restrict access if

$$\alpha^x \left[\beta^x \left(y - c_i^x - \tilde{c}_i\right) + \left(1 - \beta^x\right) y\right] + \left(1 - \alpha^x\right) y \geq y - c_i^x \tag{A.3}$$

or, alternatively, if

$$c_i^x \geq \tilde{c}_i \frac{\alpha^x \beta^x}{\left(1 - \alpha^x \beta^x\right)}. \tag{A.4}$$

This implies that every authorization for services for which this inequality holds will be restricted by insurer i, regardless of who makes the request. That is, some treatments, procedures or medications will always face supply side barriers to access. Let's define the subset of restricted services by insurer i as $X_i^r \subseteq X$. Which services belong to X_i^r depends on three factors: i). the cost of providing the service (c_i^x); ii). the probability that, if x is restricted, an agent will file a claim that is upheld ($\alpha^x \beta^x$); and iii). the idiosyncratic cost $\tilde{c}_i$.

Restrictions and Tutela Rates. We can now characterize some of the aggregate quantities used in the empirical strategy. First, since every service $x \in X_i^r$ is restricted by insurer i, the fraction of its enrollees that face restrictions is equal to the fraction of enrollees that were prescribed those services:

$$Restriction_i = \sum_{x \in X_i^r} f(x). \tag{A.5}$$

Equation A.5 shows that, beyond the three factors that determine X_i^r, the restriction rate also depends on the demand for medical care, as defined by the probability distribution of services requested f.

Second, the fraction of enrollees of insurer i that will file a tutela is given by

$$Tutela_i = \sum_{x \in X_i^r} \alpha^x f(x). \tag{A.6}$$

Let $\alpha_i \equiv \frac{\sum_{x \in X_i^r} \alpha^x f(x)}{\sum_{x \in X_i^r} f(x)} \in [0, 1]$ be the fraction of enrollees in insurer i that file tutelas conditional on having their requested services restricted. It follows that

$$Tutela_i = \alpha_i Restriction_i. \tag{A.7}$$

Equation A.7 allows us to use the tutela rate to proxy the restriction rate which is not observed. In terms of the decision tree of Figure 3, α_i corresponds to the ratio of enrollees that end up in nodes C or D, relative to the total number of enrollees in nodes B, C, or D.

Holding the demand for medical care constant, a larger restriction rate will translate into worse average health outcomes in the final period. We assumed receiving medical care was at least as good as not getting it ($\Delta h^x \geq 0$), so restrictions can only lead to a deterioration of population health. It follows from Equation A.7 that an increase in the restriction rate that is not demand driven will be reflected in a higher tutela rate and associated with i). lower utilization of health care services and ii). a higher mortality rate. These are the two key empirical predictions of the model. The predicted associations will be attenuated by the fact that a fraction

$$ClaimsUpheld_i = \sum_{x \in X_i^r} \beta^x \alpha^x f(x) \tag{A.8}$$

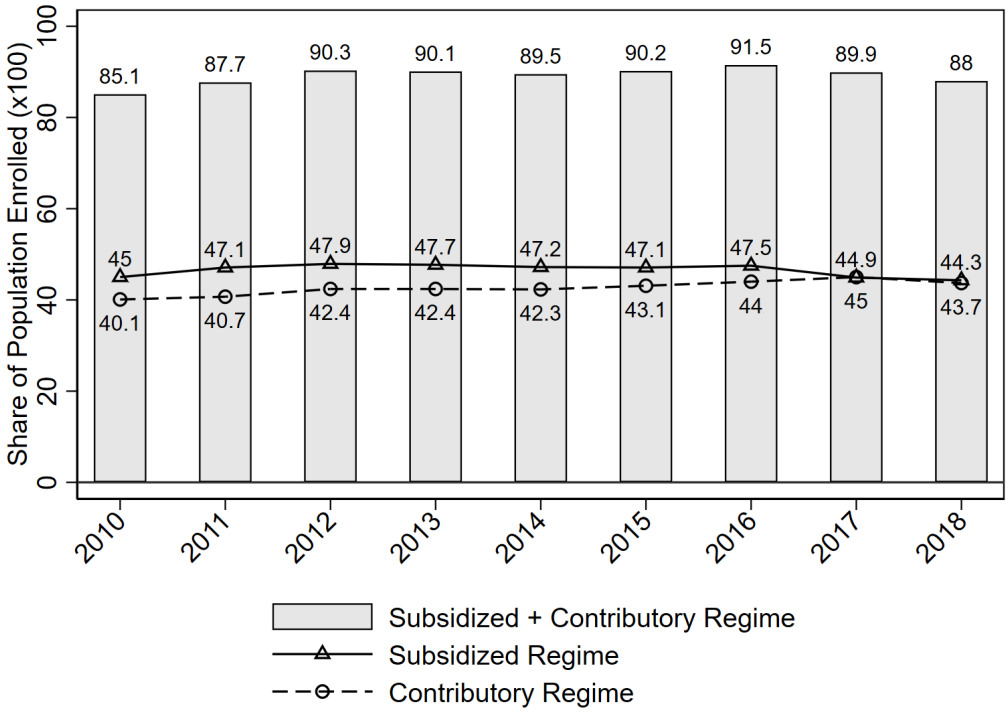
of services are delivered after favorable judicial rulings. But the fact that some services are nevertheless restricted implies there is an unmet demand for health care, defined as the share of services that are requested but not delivered. More formally,

$$
\begin{aligned}
unmetDemand_i &= Restriction_i - ClaimsUpheld_i \\
&= \left(\alpha_i^{-1} - \beta_i\right) Tutela_i,
\end{aligned}
\tag{A.9}
$$

where $\beta_i \equiv \frac{\sum_{x \in X_i^r} \beta^x \alpha^x f(x)}{\sum_{x \in X_i^r} f(x)} \in [0,1]$ is the fraction of enrollees that receive medical care as a result of litigation, and $\left(\alpha_i^{-1} - \beta_i\right) \geq 0$.

B Online Appendix

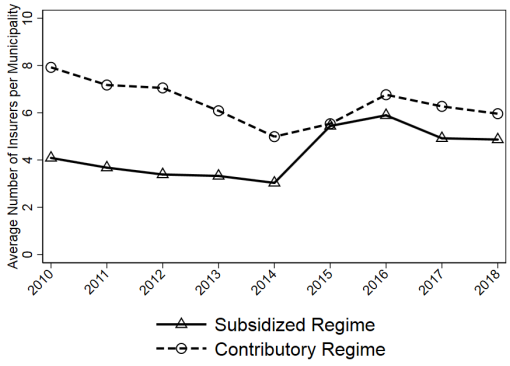
Figure B.1: Coverage of the Health System: Overall and by Regime



Notes: The bars show the share of the the population that is either in the contributory or subsidized regimes by year. The two lines report the share of population that is in the subsidized regime (solid) and contributory regime (dashed) by year.

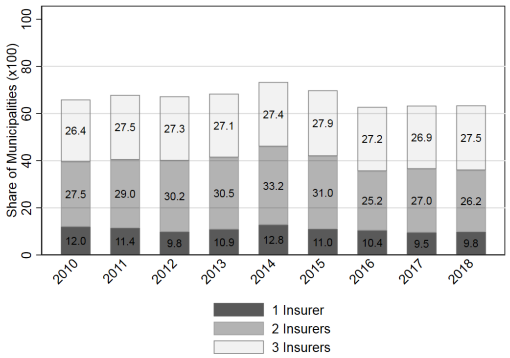
Figure B.2: Competitive Structure of the Health Care System

(a) Insurers per Municipality

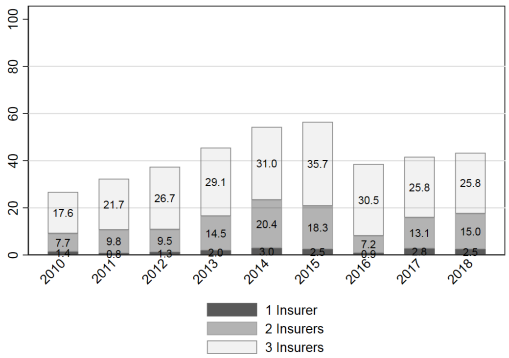


Share of Municipalities with at Most One, Two or Three Insurers

(b) Subsidized Regime



(c) contributory Regime



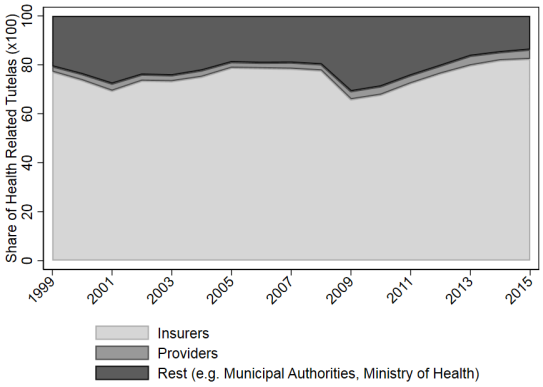
Notes: Panel (a) shows the average number of insurers per municipality and by regime. Panels (b) and (c) show the share ($\times$ 100) of municipalities that have at most 1, 2, or 3 insurers by regime.

Figure B.3: Examples of Press Coverage Documenting Insurers Limiting Access to Care Services

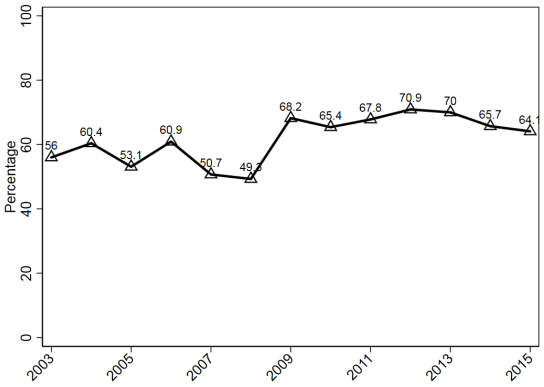
News Headline	Translation
EL UNIVERSAL Muere paciente esperando autorización para un medicamento en su EPS en Cali	**El Universal (30/01/2015).** A patient dies waiting for the health insurer to authorize a medication in Cali.
Personería denuncia que EPS no están entregando medicamentos EL TIEMPO A pacientes con cáncer les demoran hasta 10 días para la entrega de los medicamentos.	**El Tiempo (28/07/2020).** The ombudsman's office denounces heal insurers are not authorizing the delivery of medications. Cancer patients experience delays of 10 days for needed medications.
Semana **El 'paseo de la muerte' cobra la vida de un niño** Luego de dos órdenes judiciales, Sánitas habría demorado la remisión de un niño de 22 meses, lo que culminó en su muerte.	**Semana (30/09/2014).** The `walk of death' takes the life a child. Even after a judicial intervention, the health insurer delayed attention to a child 22 months of old, causing his death.
CARACOL RADIO **Joven muere después de pedir en vídeo autorización a su EPS en Cali** La mujer padecía de un cáncer de estómago y necesitaba atención especial	**Caracol (04/07/2016).** Young woman dies after asking in a video for an authorization to its health insurer. The women suffered from stomach cancer and needed special attention.
W RADIO **Personería de Tunja preocupada por incumplimiento sistemático de EPS** Los pacientes han tenido que recurrir a tutelas. La EPS está remitiendo a los usuarios a otras regiones del país, según la Personería, donde quedan desamparados.	**W Radio (26/02/2020).** ombudsman's office in Tunja worried because health insurers are systematically failing to comply. Patients are having to file tutelas.
Vanguardia **Niño con cáncer podría perder su pierna por demora de una cirugía** La vida de un menor de 14 años, quien padece de cáncer de huesos en la rodilla izquierda, está en riesgo por demoras en una cirugía que debe aprobar SaludVida EPS. Según la madre del menor, la entidad se estaría negando a autorizar la intervención quirúrgica de carácter prioritario, luego de que el médico tratante ordenara la suspensión de las quimioterapias.	**Vanguardia (14/03/2017).** A child with cancer might lose his leg because of delays in surgery. The life of 14 years old minor, who suffers from bone cancer in his left knee, is at risk for delays in a surgery that the health insurer must approve. According to the kid's mother, the entity is denying authorizing the priority surgical intervention.
EPS informa entrega de medicamento a paciente con cáncer cuando ya había muerto LA FM	**La FM (14/07/2019)**: Health insurers delivered medications for a patient with cancer after his death.

Figure B.4: Evolution of Tutelas Invoking the Right-to-Health by Defendant Body, Share of Services Requested Included in the Benefit Package, First Instance Decision by the Judge, and Regime

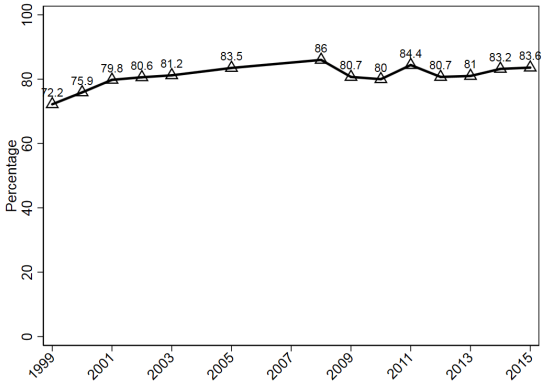
(a) Share of Health Related Tutelas by Defendant Body



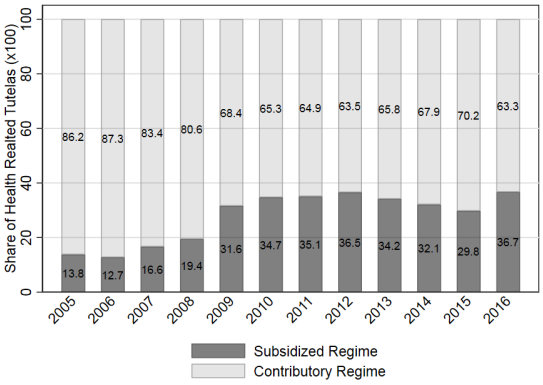
(b) Share of Services Demanded in Tutelas that are Included in the Benefit Package



(c) Share of Health Related Tutelas Decided in Favor of the Claimant



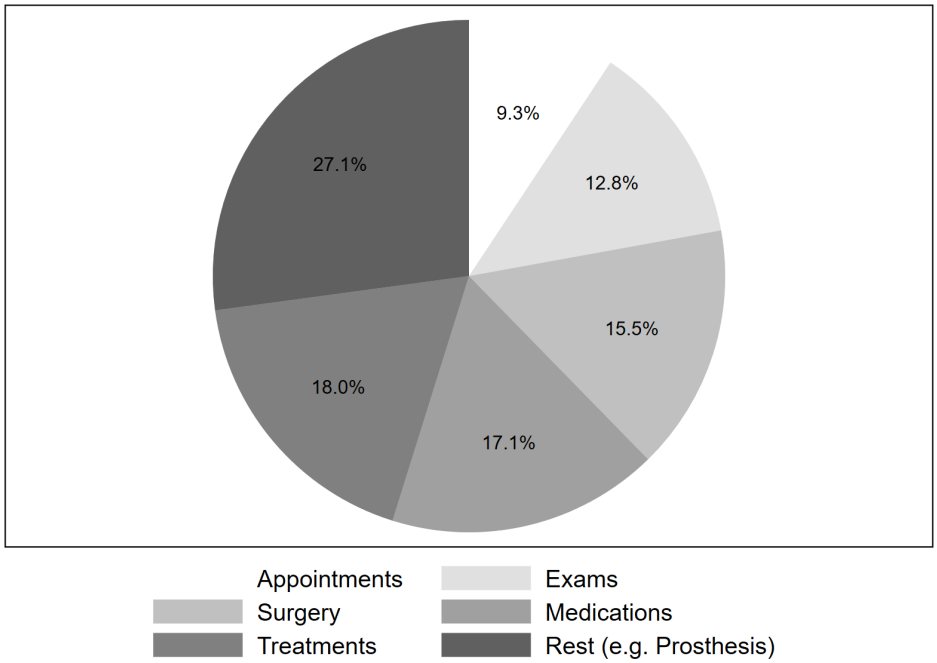
(d) Share of Health Related Tutelas by Regime



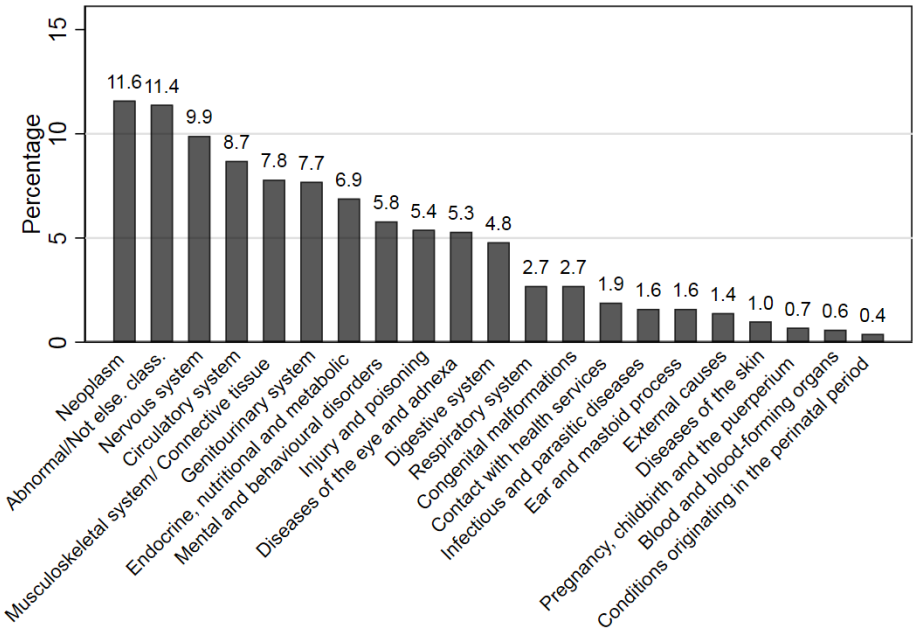
Source: Authors calculation based on Defensoría del Pueblo (2004, 2007, 2009, 2010, 2011, 2012, 2013, 2015). The numbers reported in Panels (c) and (d) include tutelas made against other actors in the system different from insurers (i.e. IPSs, Secretary of Health and Social Protection, local authorities).

Figure B.5: Type of Treatment Requested and Diagnoses Associated with the Tutelas Invoking the Right-to-Health

(a) Type of Treatment Requested in Tutelas Invoking the Right-to-Health



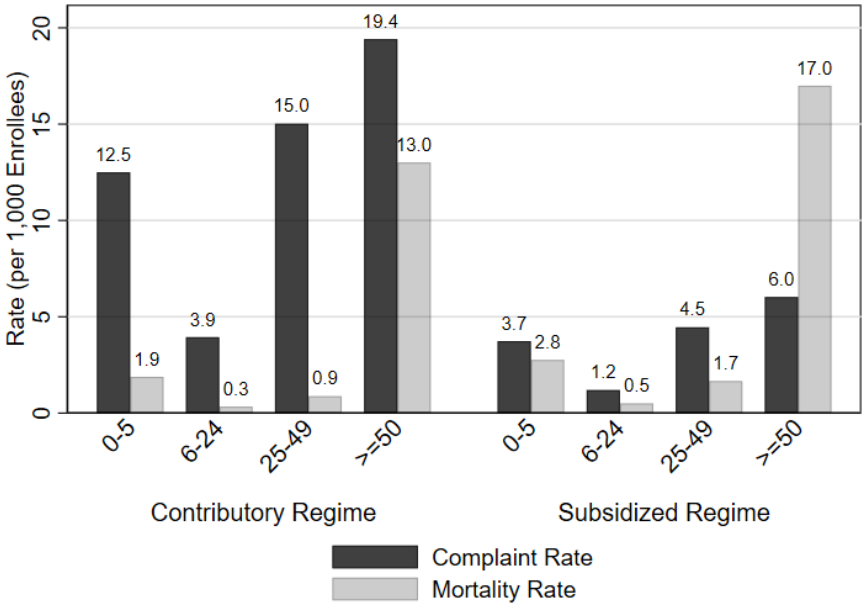
(b) Diagnoses Associated with the Tutelas in 2015



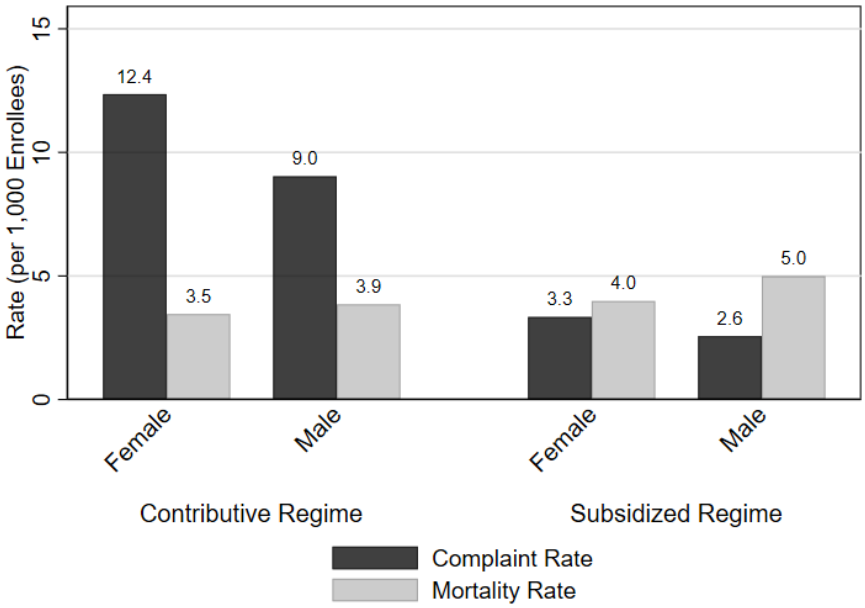
Source for Panel (a): authors calculation based on Defensoría del Pueblo (2004, 2007, 2009, 2010, 2011, 2012, 2013, 2015). Source for Panel (b): Ministerio de Salud y Proteccion Social (2018). The numbers reported in both panels include tutelas made against other actors in the system different from insurers (i.e. IPSs, Secretary of Health and Social Protection, local authorities).

Figure B.6: Complaint and Mortality Rates by Age and Sex Conditional on Regime

(a) By Age



(b) By Sex



Notes: Panel (a) reports the average complaint and mortality rates by age. Panel (b) reports the average complaint and mortality rates by sex. The divisions correspond to the people affected by the complaint, not those that file it

Table B.1: Complaints to the National Health Superintendence Between 2013 and 2017

	Overall		Contributory Regime		Subsidized Regime	
	Number	Share	Number	Share	Number	Share
Restriction of Access to Care Services*	737,785.0	44.1	565,273.0	47.7	119,172.0	32.1
Delays of Authorizations**	305,508.0	18.3	185,351.0	15.6	101,967.0	27.5
Problems with Affiliation	144,362.0	8.6	97,255.0	8.2	41,462.0	11.2
Restriction to Change/Choose Insurer/Provider	84,311.0	5.0	46,828.0	4.0	35,716.0	9.6
Denials of Authorizations**	75,309.0	4.5	51,778.0	4.4	18,445.0	5.0

Notes: The numbers in the overall column include complaints made against other actors in the system different from insurers (i.e. IPSs, Secretary of Health and Social Protection, local authorities). The numbers by regime include only complaints made against insurers. *Includes lack of opportunity to get an appointment with a care specialist, lack of opportunity to program a surgery, lack of opportunity to program an exam, and lack of opportunity to get a medication amongst others. **Includes authorizations for appointments with care specialists, surgeries, exams, and medications.