The Economics of Orbit Use: Theory, Policy, and Measurement

by

Akhil Rao

B.S., University of California, 2012

M.A., University of Colorado, 2016

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Economics

2019

This thesis entitled:
The Economics of Orbit Use: Theory, Policy, and Measurement
written by Akhil Rao
has been approved for the Department of Economics

Prof. Daniel Kaffine

Prof. Jonathan Hughes

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Rao, Akhil (Ph.D., Economics)

The Economics of Orbit Use: Theory, Policy, and Measurement

Thesis directed by Prof. Daniel Kaffine

Earth's orbits are a congestible resource with novel dynamic externalities. In this dissertation my coauthors and I examine the nature of orbit use externalities, study the policy choice space to classify existing policies and identify a class of optimal policies, consider the extent to which technological advancements can mitigate these externalities, and calculate the magnitude and time path of both an optimal satellite tax and the welfare gains from implementing it. Three key results emerge. First, open access to Earth's orbits drive the problems of excess collision risk and debris production. Left to their own devices, profit-maximizing firms may collapse the resource for generations by triggering a cascade of hazardous-fragment-producing collisions. Second, though the majority of extant policy discussions have focused on instruments targeting satellite launches, optimal policies will target satellites in orbit rather than the act of launching satellites. Despite physical uncertainty over collisions, price or quantity policy implementations are equivalent and either can maximize social welfare. Debris removal technologies cannot obviate the need for policy; they can only reduce equilibrium collision risk to the extent that satellite-owning firms pay for removal. Third, an optimal satellite tax (or orbit rental fee) for low-Earth orbit beginning in 2020 would start at approximately $40,000 USD per satellite per year, and grow at approximately 5.2% per year to preserve resource rents. The tax would increase the net present value of the satellite industry by around $1.75 trillion USD in 2020, and by over $4 trillion by 2040. Delaying action may be very costly: relative to a baseline of having begun optimal management in 2015, beginning optimal management in 2035 forgoes on the order of $4.6 trillion USD of permanent orbit use value in 2040.

Acknowledgements

This dissertation would not have been possible without involvement and support from more people than I can list.

I am profoundly grateful to my mom, dad, sister, and partner for their patience and support throughout this process. Thank you for listening to my rambling, calming my fears, reading my drafts, and reminding me that I need to eat, sleep, and take breaks.

Many people have given me constructive comments and feedback. I am grateful to Dan Kaffine, Jon Hughes, Martin Boileau, Miles Kimball, Matt Burgess, Giacomo Rondina, Alessandro Peri, Matt Butner, Kyoung-Gon Kim, Sean Ericson, Scott Holladay, and Anna Rubinchik for their detailed comments on various drafts and ideas. I am particularly grateful to Dan for his enthusiasm to advise a student with an unusual research idea, and for his guidance in pursuing it. Everything is cooler in space.

My knowledge of orbital mechanics has grown considerably through this project, and I am grateful to Kas McLean, Trevor Bennett, John Marino, Nick Campbell, and Sven Ericson for guiding my intuition and analytical thinking in this respect.

I am grateful to Francesca Letizia of the European Space Agency, Teri Grimwood of the Union of Concerned Scientists, and Debra Shoots of the NASA Orbital Debris Program Office for sharing data on collision risk, active satellites, and debris growth.

Funding for this research was generously provided by Center for Advancement of Research and Teaching in Social Science, the Reuben A Zeubrow Fellowship in Economics, and the CU Boulder Summer Dissertation Fellowship.

All errors are my own.

Contents

List of Tables

Table

List of Figures

Figure

Chapter 1

Introduction

Gravity wells of planetary bodies are congestible resources. Orbiting satellites face the risk of colliding with other satellites, and objects left in orbit can stay there for hundreds or thousands of years. The current legal environment on Earth precludes efficient allocation of orbital paths through property rights and markets, making Earth's orbits into an open access commons. The orbital congestion problem is made worse by the possibility of Kessler Syndrome - a cascade of collisions between objects in orbit, each collision producing many fragments and increasing the likelihood of another collision.

Relatively little is known about how to manage and measure orbital congestion under open access. Economic theory suggests that open access to a commons will drive the expected profits of commons use to zero, and that congestion can be managed by a system of property rights over the congested resource or a Pigouvian tax levied on the source of the externality. But what do these principles mean for orbital resources, given their novel dynamics and constrained legal environment? Will expected profit maximization under open access prevent or exacerbate positive feedbacks in debris growth? What are the relevant margins for optimal policy design, and to what extent can technological advances limit inefficient orbit use? How should the marginal external cost of orbit use be measured, and what does it imply for the efficient rental rate of orbit use? How much better off would society be by switching to optimal orbital management? Each chapter of my dissertation addresses one of these questions and related issues in orbital management.

In Chapter 2, Giacomo Rondina and I study the economic dynamics of open access and socially optimal orbit use. We focus on answering, "How will orbits be managed under open access, and how should they be managed?" We develop formal results and economic intuition for the nature of equilibrium congestion in orbit, and consider the effects of different time-varying physical and economic processes on open access orbit use. We show that the equilibrium collision risk under open access will be determined by the excess return on a satellite and that the economic feedback from expected congestion cost to profits can short-circuit explosive debris growth. However, firms under open access will not internalize longer-term debris accumulation. Consequently they may cause Kessler Syndrome despite limiting launches in response to collision risk. Increasing the decay rate of debris (e.g. through deorbit guidelines) or decreasing the amount of launch debris generated (e.g. through technological advances) may exacerbate equilibrium congestion. The interaction between debris accumulation, collision risk, and profit maximization may also cause endogenous fluctuations in orbital stocks around the steady state. This chapter connects to the literatures on open access and common resource problems.

In Chapter 3, I derive economic principles governing the choice of space traffic control policies. I focus on answering, "How should regulators choose amongst management policies, and to what extent can debris removal technology reduce the scope of the externality?" I show that policies which target satellite ownership, such as satellite taxes or permits, achieve greater expected social welfare than policies which target satellite launches, such as launch taxes or permits. Price or quantity policies can achieve equal expected social welfare due to the symmetry of uncertainty between regulators and firms. I also show that active debris removal can reduce the risk of runaway debris growth no matter how it is financed, but can only reduce the risk of satellite-destroying collisions if satellite owners pay for it. Technical solutions to space traffic control tend to emphasize launch restrictions or public funding of debris removal technology development and use, but often ignore that current and prospective orbit users dissipate rents under open access. While satellite-focused policies can achieve first-best orbit use, attempts to control orbital debris growth and collision risk through launch fees or debris removal subsidies under open access may be ineffective or backfire. This chapter connects to the literatures on congestion control, commons management, and time consistent policy design.

In Chapter 4, Dan Kaffine, Matt Burgess, and I calibrate the model developed in the preceding chapters and use it to estimate the time paths of an optimal satellite tax or rental fee, and the improvement in satellite safety and long-run global welfare from moving to a globally-harmonized optimal orbital management regime. We estimate physical and economic parameters from historical data on satellite and debris counts and collision probabilities over 1957—2017 and aggregate satellite industry revenues and costs over 2006—2015. The estimation and simulation methods can be used to calculate optimal policy values and gains from optimal policy under alternate physical and economic modeling assumptions and given different starting points. To achieve the results of a counterfactual optimal satellite management regime beginning in 2006, an optimal satellite tax would be on the order of $37,500 per satellite in 2020. The optimal management regime would approximately halve the risk of satellite-destroying collisions in 2020 and more than double the net present value of global social welfare in 2020 from orbit use. This chapter connects to the literatures on congestion measurement, commons management, and integrated assessment modeling.

Answers to questions of orbit use have implications for other dynamically-evolving global commons problems, such as climate change. Kessler Syndrome is similar to runaway global warming, while active debris removal is analogous to carbon capture and sequestration technologies. Results from the analysis of orbit use suggest a case for pessimism regarding both runaway warming and carbon capture technologies: even if polluting firms believe that they will face negative effects from climate change, they may not stop polluting in time to prevent physical dynamics from reaching the tipping point. Unless the costs of their deployment are somehow charged to polluters, the development of technologies to mitigate and reverse climate change may create additional incentives to pollute. While the space domain is intrinsically interesting and, as commercial space use intensifies, of increasing importance, further research on orbit management can offer lessons about other dynamic commons problems as well.

Chapter 2

Cost in Space: Debris and Collision Risk in the Orbital Commons

Written with Giacomo Rondina

As society launches more satellites, the risk of collisions between orbiting objects increases. Such collisions can destroy satellites and produce orbital debris, further increasing the risk of future collisions. Collision risk and debris accumulation threaten active satellites and the future of human activity in space. How will open access affect orbital debris accumulation, satellite collision risk, and occurrence of Kessler Syndrome? These questions have been explored very little in economics, and have not yet been addressed in the physics and engineering or law and policy literatures.[1] In this paper, we examine the consequences of open access to orbit in the first long-run dynamic economic model of satellite launch, show that the equilibrium collision rate is determined by the excess return on a satellite, and consider the economics of debris accumulation.

Satellites produce debris over their lifecycle. Launching satellites produces orbital debris (spent rocket stages, separation bolts), satellites can produce some debris while in orbit (paint chips, lost tools, etc.), and satellites which are not deorbited or shifted to disposal orbits at the end of their life become debris.[2] Objects in orbit move at velocities higher than 5 km/s, so debris as small as 10 cm in diameter can be hazardous to active satellites.

[1] Wienzierl (2018) highlights a number of issues in the development of a space economy, from space debris to coordination problems and market design, which economists have the tools to address.

[2] Reusable rockets can significantly lower the cost of launching and the amount of launch debris generated. SpaceX and Blue Origin are currently the only launch providers who offer such vehicles.

Though debris does deorbit naturally, this process can be very slow, especially at higher altitudes. Debris as low as 900 km above the Earth's surface can take centuries to deorbit, while debris at 36000 km can take many millennia (Weeden (2010)). Compounding the problem is the fact that collisions between debris objects can generate even more debris, also moving at high velocities and capable of damaging operational satellites.[3] It is possible for debris accumulation to cause a cascading series of collisions between orbital objects, resulting in an expanding field of debris which can render an orbital region unusable and impassable for thousands of years. Engineers and physicists call this phenomenon "collisional cascading" or "Kessler Syndrome" (Kessler and Cour-Palais (1978)). Kessler Syndrome can cause large economic losses, directly from damage to active satellites and indirectly from limiting access to space (Bradley and Wein (2009), Schaub et al. (2015)). Proposed commercial uses of space, like low-Earth orbit (LEO) broadband internet constellations with global coverage, asteroid mining, and space-based solar power, could become infeasible if Kessler Syndrome occurs. Existing estimates of debris growth indicate that the risk of Kessler Syndrome is highest in LEO, placing imaging and future telecommunications satellites at risk and potentially reducing access to higher orbits (Kessler et al. (2010)). Currently, there are more than 1,000 operating satellites in orbit, up to 600,000 pieces of debris large enough to cause satellite loss, and millions of smaller particles that can degrade satellite performance (Ailor et al. (2010)). Approximately 49 percent of active satellites are in LEO, 41 percent are in geostationary orbit (GEO), and the remainder are in elliptical or other orbits.

Existing legal frameworks for orbit use, such as the Outer Space Treaty, complicate the process of establishing of explicit orbital property rights and hinder cleanup efforts. For example, Article 2 of the OST forbids national appropriation or claims of sovereignty over outer space, which could be interpreted as prohibiting national authorities from unilaterally establishing orbital property rights.[4] (Gorove (1969)) As

[3]The risk of debris striking another satellite is generally orders of magnitude larger than the risk of debris deorbiting and harming consumers directly.

[4]Weeden (2010) and Weeden and Chow (2012) are skeptical of what decentralized bargaining can achieve in this setting, given the difficulties in solving other global coordination problems, like climate change, and the number of current and potential orbit users

launch and satellite costs decrease and more firms enter markets for satellite services, the debris problem is likely to get worse (Selk (2017)).

The congestibility of orbital resources has been recognized in economics as early as Sandler and Schulze's 1981 paper, **"The Economics of Outer Space"** (Sandler and Schulze (1981)). In it the authors present a series of models to analyze issues which may face a future space economy, including a static optimization program to manage orbital spectrum and positions as club goods. Their program accounts for congestion due to radio frequency interference and collision risk, but ignores debris accumulation. Despite this early recognition of economic externalities in orbit, economists have paid relatively little attention to orbital management, likely in part due to the lack of development in commercial space markets. More recent economic analyses have considered the static economic costs of inefficient electromagnetic spectrum and position use or the inefficiency of open access and voluntary debris mitigation under monopolistically competitive behavior in static or two-period settings (Macauley (1998), Macauley (2015), Adilov et al. (2015)). The economic dynamics of orbit use under perfectly competitive open access and socially optimal management over longer time horizons have not been analyzed yet.

Orbital congestion has been studied more actively in physics and engineering, beginning with Kessler and Cour-Palais' analysis of the creation of a debris belt (Kessler and Cour-Palais (1978)). This community has focused on two areas: how orbital congestion might evolve in particular orbits as the numbers of satellites and debris fragments increase, and how satellite systems and trajectories should be designed to be more robust to spectral congestion and physical collision risk. Papers in the former literature have relied on launch rate models estimated from historical trends, and have abstracted away from optimizing or forward-looking behavior by the agents launching the satellites. Gordon's observation on early models of fisheries (Gordon (1954)) comes to mind:

and their incentives. Salter and Leeson (2014) and Salter (2015) take more optimistic views of what commercial users can achieve under celestial anarchy, the former based on the idea of self-enforcing property rights and the latter based on the Coase Theorem, technological advancement, and insurance markets.

> On the whole, biologists tend to treat the fisherman as an exogenous element in their analytical model, and the behavior of fishermen is not made into an integrated element of a general and systematic 'bionomic' theory... the 'overfishing problem' has its roots in the economic organization of the industry.

Similarly, with the exception of Adilov et al. (2015), current models of Earth orbit use do not account for the economic incentives involved in launching satellites.

First, satellites are assets, and collision risk reduces their expected present value. This gives firms an incentive to avoid launching if the collision risk becomes too high, potentially stabilizing orbit use by deterring entry. However, because orbits are an open access commons, firms will launch until the collision risk makes expected profits zero, rather than stopping when expected marginal profits are zero.[5] Entry until zero profits would be socially optimal if not for the presence of endogenous collision risk. Second, debris is an accumulating stock pollutant which only matters because it increases collision risk. This gives firms an incentive to prevent debris accumulation, but only to the extent to which they are directly affected by it now and in the future. Since open access makes the expected net present value of owning a satellite zero, firms' incentives to protect the future value of their satellites is reduced. The problem of debris accumulation is made uglier by the potential for Kessler Syndrome, but it is not primarily a pollution problem.[6]

The feedback from collision risk to profits causes the equilibrium launch rate to decrease as collision risk increases. This can make the equilibrium launch rate also decrease as the debris stock increases, but

[5]The geosynchronous belt is an exception; though it is still a commons, it is not under open access. The International Telecommunication Union auctions orbital slots which national authorities then allocate. See Macauley and Portney (1984) and Jones et al. (2010) for more discussion of these mechanisms.

[6]Haveman (1973) describes the similarities and differences between open access, congestion, and pollution externalities in more detail. The key distinction between congestion and pollution is that congestion is a reciprocal externality stemming from crowding costs not being reflected in marginal use decisions, while pollution is a one-sided externality borne by agents who are different from the ones creating the pollution. Pollution is part of the problems of orbit use insofar as debris is an intertemporal externality imposed by current users on future users, but debris also imposes costs on current users, so the distinction is not clean-cut. Since debris is irrelevant in the absence of collision risk, it seems reasonable to argue that congestion, not pollution, is the heart of the problems of social cost in orbit.

the lack of feedback from future satellite operators to present satellite operators means that the equilibrium launch rate may not be responsive enough to debris accumulation to prevent Kessler Syndrome. The appropriate property rights or system of charges could make the collision risk level efficient and prevent runaway debris growth.

We build on prior economic analysis of orbital debris by extending the setting from two periods to arbitrary finite and infinite horizons and by allowing more general analytical forms of satellite and debris accumulation. We contribute to the literature on common-pool resource management in the presence of environmental risk by incorporating a capital stock that is affected by congestion of the commons and a pollution stock which increases congestion. We do not consider spectrum use explicitly, though we comment on the effects of spectrum congestion in section A.1.

Open access is characterized by too many launches and collisions and too much debris relative to the socially optimal plan, in the short-run and the long-run. While open access steady states cannot cause Kessler Syndrome and firms always face incentives to limit satellite-destroying collisions, open access may cause Kessler Syndrome while approaching a steady state. Like bioeconomic collapse in fisheries, Kessler Syndrome becomes more likely under open access as the cost of launching satellites falls, or more generally as the excess return on a satellite rises. Regions with higher decay rates or lower incidence of launch debris may have higher long-run debris levels due to a combination of short-run rebound effects and local instability of long-run equilibria. Additionally, increases in the rate of excess return earned by a satellite will tend to locally destabilize open access steady states with higher levels of debris. These results suggest that purely technical solutions aiming to limit debris growth by targeting launch debris or debris decay rates may ultimately increase the amount of orbital debris or destabilize existing equilibria. Market-based instruments or a new legal paradigm for orbit use may be essential for orbital sustainability.

2.1 Social cost in orbit

2.1.1 Orbits as an environmental resource

Orbits share characteristics of renewable and nonrenewable resources. On any timescale relevant for human decision making, the Earth will continue to exert its gravitational pull. Bodies in its gravity well will eventually either exit the well or fall to the surface. In low-Earth orbits, the decay is fast enough (on the order of hours to years) that the gravity well and paths within it are renewable on human timescales. In higher orbits, such as the geosynchronous belt, the decay is slow enough (on the order of millions of years) that the well and its paths are nonrenewable on human timescales. Between the two lie a continuum of possible paths through the well, each with their own rates of renewal. Elliptical paths through the well, such as Molniya orbits, span multiple regions of renewability.[7]

Fishing and mining offer useful intuition for the economics of orbit use. In fisheries, open access results in entry until expected profits are zero, and may drive the stock below the minimum level required for biological renewal. While fishing may become uneconomical at or past that level, once the threshold is crossed the natural dynamics of the stock will drive the fish population to zero anyway. The stock level at which fishing becomes uneconomical may also be above the minimum biological level, in which case open access will result in sustainable, if suboptimal, harvest rates. In standard open access fishery models, the lack of rights over the fish stock induces myopic fishing - though the fish stock evolves over time, fishers make a sequence of static decisions about harvest levels. Since future profits will be zero anyway due to open access, there is no incentive to conserve today for higher profits tomorrow. In contrast, a mine owner faces an inherently dynamic problem. Each decision to extract affects the amount they are able to extract in the future. As a result, a mine owner's profit-maximizing extraction path equalizes discounted returns over time. To do otherwise would be to leave money on the table in some periods.

[7] A Molniya orbit has a low perigee over the Southern Hemisphere and a high apogee over the Northern Hemisphere. Molniya orbits require less power to cover regions in the Northern Hemisphere (e.g. former Soviet Union countries) than geosynchronous orbits, due to the low incidence angles of rays from the Northern Hemisphere to geosynchronous positions.

As in an open access fishery, firms under open access launch satellites until expected profits are zero, and may fill the orbits past the maximum level at which natural decay can prevent net debris growth. While satellite launching may become uneconomical at or past that level, the natural dynamics of orbiting bodies will cause further collisions and debris growth if the number of objects in orbit is not reduced. If owning a satellite is not uneconomical before this occurs, it is likely to be uneconomical after. Orbit users equalize the expected net present value of owning a satellite across periods, similar to how a competitive mine owner sequences ore extraction.

The neoclassical growth model with pollution is another helpful reference point. Satellites are capital assets which produce a per-period payoff of π at zero marginal cost per period, but a fixed cost of F to build and launch.[8] The collision rate, $L(S,D)$, is similar to the capital depreciation rate in how it enters the law of motion for satellites. Debris is a stock of residuals from production (launch debris) and "depreciation" (collision fragments). If debris did not cause collisions ($L(S,D) = L(S)$), it would be irrelevant and only the satellite stock would need management. If the collision rate were uncoupled from even the satellite stock ($L(S) = L$), then this model would reduce to the neoclassical growth model with an irrelevant pollution stock. Open access would then be efficient.

The coupling between the satellite stock and the collision rate ($L(S)$) implies congestion in orbit under open access, since firms will not internalize the cost of the risks their satellites pose to others.[9] Adding a coupling between the debris stock and the collision rate ($L(S,D)$) adds to the congestion.

[8] This is a simplification to focus on the margin of launch decisions. Operational costs like managing receiver stations on the ground or monitoring the satellite to perform stationkeeping are incurred each period. Since the decision to launch a satellite involves forecasting the operational costs, they can be viewed in this model as having been capitalized into the fixed cost.

[9] This mirrors the difference between a competitive firm which doesn't internalize the effect of its entry on the price and a monopolist who internalizes the price effect of marginal units of production. As in other natural resource settings, an orbital monopolist acts as a "satellite conservationist" by restricting and resequencing launches to preserve the fleet.

When a firm launches a satellite, its entry to the orbit adds launch debris to the orbit. While its satellite is in orbit, the firm contributes to congestion in the orbit. If the satellite is lost in a collision, its removal reduces the risk to survivors, but the new fragments from its destruction increase the risk to survivors later. Firms ignore the congestion they cause through the debris created by their launch, the addition of their satellite to the orbit, and the new fragments created by their satellite's eventual destruction. The debris effect of launch is "dynamic" congestion imposed by a firm's entry on itself and others. The risk generated by a satellite's presence is "steady state" congestion imposed by a firm on others. The debris effects of destruction are also dynamic congestion imposed by a firm's exit on others. As long as launches create debris, marginal satellites in orbit are bundled with debris creation, so satellites and launch debris are also complements in producing social value. The essential tradeoff of the planner's problem is in balancing the lifetime value created by satellites and launch debris against the present and future congestion created by each.

2.1.2 A simple model of orbital mechanics with no aggregate uncertainty

An orbital region is a set of closed paths around a central body. When the paths are chosen to form a closed spherical shell around the central body, the orbital region is also called an orbital shell. Paths which span shells are possible and useful for some applications, but highly elliptical orbits (such as Molniya orbits) are the exception rather than the rule. In what follows, we assume all bodies are in an orbital shell. This approach is frequently used in debris modeling, e.g. Rossi et al. (1998) and Bradley and Wein (2009), though higher fidelity models use large numbers of small regions to track individual objects, e.g Liou et al. (2004), Liou and Johnson (2008), and Liou and Johnson (2009). We abstract from the composition of orbital stocks, and assume that all satellites and debris are identical.

The number of active satellites in orbit is the number of launches in the previous period plus the number of satellites which survived the previous period. The amount of debris in orbit is the amount from the previous period which did not decay, plus the number of new fragments created in collisions, plus the amount of debris in the shell created by new launches. The laws of motion for the satellite and debris stocks show this:

$$S_{t+1} = S_t(1 - L(S_t, D_t)) + X_t \tag{2.1}$$

$$D_{t+1} = D_t(1 - \delta) + G(S_t, D_t) + mX_t. \tag{2.2}$$

$L(S_t, D_t)$ is the proportion of orbiting active satellites which are destroyed in collisions, and $G(S_t, D_t)$ is the number of new fragments created by collisions between orbiting bodies. We assume that the collision rate is nonnegative, increasing in each argument, and bounded below by 0 and above by 1.[10] No satellites can be destroyed when there are none in orbit ($L(0, D_t) = 0\ \forall D_t$). In a model with aggregate uncertainty in the number of losses, $L(S_t, D_t)$ would be the expected number of losses in period t.

We assume that the number of new fragments is nonnegative, increasing in each argument, and zero when there are no objects in orbit ($G(0,0) = 0$). To derive results about the occurrence of Kessler Syndrome in section 2.2.2, we also assume that the growth in new fragments due to debris alone will eventually be greater than δ. δ is the rate of orbital decay for debris, and m is the amount of launch debris created by new satellites.

To fix concepts, a form for $L(S_t, D_t)$ is helpful. We can model the average rate at which objects of type j are struck by objects of type k as

$$p_{jk}(k_t) = 1 - e^{-\alpha_{jk} k_t}, \tag{2.3}$$

where $\alpha_{jk} > 0$ is a physical parameter reflecting the relative mean sizes, speeds, and inclinations of the object types. The rate at which satellites are destroyed is the sum of the rates at which they are struck by debris and by other satellites, adjusted for the number of satellites which are struck by both. For

[10]Firms try to avoid collisions by maneuvering their satellites when possible; the collision rate in this model should be thought of as the rate of collisions which could not be avoided, with easily avoided collisions optimized away. Collisions which could have been avoided but were not due to human error are included in this. Implicitly we are assuming that firms are operate their satellites as imperfect cost-minimizers.

satellite-satellite and satellite-debris collisions, equation 2.3 gives us

$$
\begin{aligned}
L(S,D) &= p_{SS}(S) + p_{SD}(D) - p_{SS}(S)p_{SD}(D) && (2.4)\\
&= (1-e^{-\alpha_{SS}S}) + (1-e^{-\alpha_{SD}D}) - (1-e^{-\alpha_{SS}S})(1-e^{-\alpha_{SD}D})\\
\implies L(S,D) &= 1 - e^{-\alpha_{SS}S - \alpha_{SD}D}. && (2.5)
\end{aligned}
$$

The form in equation 2.3 is convenient as it allows us to solve explicitly for the open access launch rate and is easy to manipulate. Though this form does not satisfy our assumption that $L(0,D) \equiv 0$, it can be derived by analogy to a kinetic gas model of particle collisions Letizia et al. (2017). We use these probability functional forms to parameterize $G(S_t, D_t)$ as

$$
\begin{aligned}
G(S,D) &= \beta_{SS}p_{SS}(S)S + \beta_{SD}p_{SD}(D)S + \beta_{DD}p_{DD}(D)D && (2.6)\\
\implies G(S,D) &= \beta_{SS}(1-e^{-\alpha_{SS}S})S + \beta_{SD}(1-e^{-\alpha_{SD}D})S + \beta_{DD}(1-e^{-\alpha_{DD}D})D, && (2.7)
\end{aligned}
$$

where the $\beta_{jk} > 0$ are physical parameters reflecting the mean number of fragments created by collisions between objects of type j and k, weighted by the average time each such fragment spends in the shell of interest. We use these forms where they facilitate exposition or computation, but our analytical results do not assume a specific functional form for the collision rate.

2.1.3 Satellites as a productive asset

Active satellites provide services to a variety of entities: individuals, firms, governments, research agencies, and others. The services provided tend to be information services like mobile broadband, images of the Earth, and GPS. We ignore such product differentiation, though it is an important feature of orbit use, to focus on the dynamics of the collision rate and debris growth. Adilov et al. (2015) account for differentiation in a two-period setting.

We assume that satellites are identical, infinitely lived unless destroyed in a collision, and produce a single unit of output per period valued at $\pi > 0$. The market for satellite output is perfectly competitive, so that the price per unit service is the same as the social marginal benefit from that unit of service. The rate of return on a satellite is given by $\pi/F \equiv r_s > 0$. We discuss the effects of relaxing the assumption that satellites are infinitely lived in section A.2. We assume that costs and returns are constant over time, and comment on the effect of the rate of return on a satellite varying over time in section A.3.

There are two types of agents: a profit-maximizing firm which can own up to one satellite at a time, and a fleet planner who owns all satellites. A firm which owns a satellite collects a revenue of π every period that the satellite survives. A fraction $L(S_t, D_t)$ of the orbiting satellites are destroyed in collisions every period. Since they are identical, the probability that an individual satellite survives the period is $1 - L(S_t, D_t)$. The discount factor used by all agents is $\beta = (1+r)^{-1}$, where $r > 0$ is the social discount rate. The value of a satellite, denoted $Q_t(S_t, D_t, X_t)$, is the sum of present returns and the expected discounted value of its remaining lifetime returns.

$$Q_t(S_t, D_t, X_t) = \pi + \beta(1 - L(S_t, D_t))Q_{t+1}(S_{t+1}, D_{t+1}, X_{t+1}) \tag{2.8}$$

where $X_t = \int_0^\infty x_{it} di$ is the aggregate launch rate based on each potential launcher's entry decision x_{it}. We assume for simplicity that firms cannot choose to deorbit satellites.

A firm which does not own a satellite in period t faces the decision to pay a fixed cost F and launch a satellite which will reach orbit and start generating revenues in period $t+1$, or to wait and decide again whether or not to launch in period $t+1$.[11] Assuming potential launchers are risk-neutral profit maximizers, the value of potential launcher i at period t is

$$V_{it}(S_t, D_t, X_t) = \max_{x_{it} \in \{0,1\}} \{(1 - x_{it})\beta V_{it+1}(S_{t+1}, D_{t+1}, X_{t+1}) + x_{it}[\beta Q_{t+1}(S_{t+1}, D_{t+1}, X_{t+1}) - F]\} \tag{2.9}$$

$$\text{s.t. } S_{t+1} = S_t(1 - L(S_t, D_t)) + X_t$$

$$D_{t+1} = D_t(1 - \delta) + G(S_t, D_t) + mX_t$$

The launch Bellman has an i subscript to indicate that firms may or may not choose to launch. Since satellites are identical, there is no i subscript on the value of a satellite or the expectation operator.

[11] Whether the lag reflects time-to-build or time-to-launch, there is a difference between the timescale of physical interactions in orbit and the timescale of launch decisions. The former occurs continuously, while the latter does not.

2.1.4 Open access and the equilibrium collision rate

Under open access, firms launch satellites until the value of launching is 0,

$$X_t \geq 0 : \beta Q_{t+1} = F. \tag{2.10}$$

The resulting value of a satellite is the period return plus the expected value of its survival under open access,

$$Q_t = \pi + (1 - L(S_t, D_t))F. \tag{2.11}$$

The open access condition and the satellite value can be rewritten to show that in equilibrium the flow of benefits generated by a satellite is equated with the flow of opportunity costs and expected collision costs (the marginal private costs):

$$\pi = rF + L(S_{t+1}, D_{t+1})F. \tag{2.12}$$

The Implicit Function Theorem and our assumptions on the derivatives of the collision rate and new fragment function allow us to obtain comparative statics on the launch rate from equation 3.51 without imposing a specific functional form for $L(S,D)$. We provide general analytical results in section 2.1.10, and a specific example in this section to illustrate the intuition. Let $L(S,D) = 1 - e^{-\alpha_{SS}S - \alpha_{SD}D}$. Define the following quantities:

$$\text{(Log rate of excess return:) } R = -\log(1 + r - r_s) \tag{2.13}$$

$$\text{(Launch contribution to collision rate:) } a^{-1} = \alpha_{SS} + \alpha_{SD}m \tag{2.14}$$

$$\text{(Carryover satellite stock:) } \mathcal{S}_t = S_t(1 - L(S_t, D_t)) \tag{2.15}$$

$$\text{(Carryover debris stock:) } \mathcal{D}_t = D_t(1 - \delta)) + G(S_t, D_t), \tag{2.16}$$

The forms of the first two terms come from $L(S,D)$. The first term is the negative log return differential, and is only defined when $r_s - r < 1$. Since the equilibrium collision rate under open access is $r_s - r$, this restriction imposes that some satellites survive each period. The second term is the direct effect of a new launch on the collision rate: each new launch adds 1 satellite to orbit with collision parameter

α_{SS}, and m units of debris with collision parameter α_{SD}. Modern launches typically carry more than one satellite; a launch delivering k satellites would contribute ka^{-1} to the collision rate, though we do not model ridesharing on launches. The third and fourth terms are the present stock's contributions to the future stock.

Using the form for $L(S,D)$ in equation 3.51, we can solve for the equilibrium launch rate as

$$\bar{X}(S_t,D_t) = a[R - \alpha_{SS}\mathscr{s}_t - \alpha_{SD}\mathscr{d}_t].$$

Increases in the excess return cause more firms to want to own a satellite, and hence to launch. On the other hand, strengthening the collision rate couplings (measured in α_{SS} and α_{SD}, described more generally in section 2.1.6) increases the persistence of prior stocks and reduces the scope for firms to launch new satellites. Increasing the amount of launch debris, m, increases the risk that the newly launched satellites pose to themselves, and reduces the equilibrium launch rate. Since the launch rate cannot be negative, the open access launch policy is

$$X_t = \begin{cases} \bar{X}(S_t,D_t) & \text{if } R > \alpha_{SS}\mathscr{s}_t + \alpha_{SD}\mathscr{d}_t \\ 0 & \text{otherwise.} \end{cases} \tag{2.17}$$

As $\alpha_{SS} + \alpha_{SD}m \to 0$, launches become decoupled from the collision rate and debris evolution. If this happens, the launch rate goes to infinity when $r_s > r$. This is an extreme and unrealistic case, but it highlights the role of collisions in the model: without the risk of a collision, a satellite is a perfectly safe asset that provides a higher rate of return than the risk-free rate. In reality, the risk of a collision is just one of many factors to be weighed in the expected cost of launching a satellite.

2.1.5 Marginal and average collision rates

The marginal survival rate, $1 - L - \frac{\partial L}{\partial S}S$, measures the proportion of satellites which will survive collisions given the change in the collision rate caused by new satellites. Its sign is important in understanding orbit use and the marginal external cost, in particular the case when $1 - L - \frac{\partial L}{\partial S}S \geq 0$. Rearranging, we see that this statement relates two quantities: the average survival rate ($\frac{1-L}{S}$), and the marginal collision rate due to a satellite ($\frac{\partial L}{\partial S}$). In particular, when the marginal survival rate is nonnegative, the average survival rate exceeds the marginal collision rate due to another satellite ($1-L-\frac{\partial L}{\partial S}S \geq 0 \implies \frac{1-L}{S} \geq \frac{\partial L}{\partial S}$). Since the average survival rate is a decreasing function of the average collision rate, the condition implies that the marginal collision rate due to another satellite is greater than the average collision rate, so another satellite will increase the collision rate.

The opposite statement, $1-L-\frac{\partial L}{\partial S}S \leq 0$, implies that the marginal collision rate is lower than the average collision rate (greater than the average survival rate), so another satellite would decrease the collision rate. This could be plausible in cases where the "most likely destruction" after adding another satellite is less severe than it was before - for example, if the new satellite poses a significant threat to a dangerous already-orbiting satellite, the increased likelihood of that one collision may decrease the expected number of collisions. This seems unrealistic for most applications. For the rest of this paper, we rule out this case and assume $1-L-\frac{\partial L}{\partial S}S \geq 0$.

Another way to motivate this assumption is through "careful placement": since firms launching satellites want them to survive for as long as possible, they will avoid launching in ways which are more likely to result in the satellite's destriction. As a result, the increase in the expected number of collisions from a marginal satellite should be smaller than the average probability a satellite survives. This assumption is formalized below:

Assumption 1. *(Launches increase the rate of collisions) The marginal survival rate is nonnegative,*

$$1-L-\frac{\partial L}{\partial S}S \geq 0,$$

implying that a new launch will weakly increase the collision rate.

2.1.6 Three types of physical couplings

We use the term "physical coupling" here to refer to ways that objects in the laws of motion for satellite and debris stocks are connected to each other. These couplings give some structure to the laws of motion, and drive the physical dynamics of orbit use. There are three economically relevant physical couplings between objects in orbit. If all of these couplings were turned off, there would be no problem of excess congestion or interesting dynamics in orbit.

The couplings highlight the role of debris in the economics of orbit use. Without debris, the only congestion externality can be a "steady state" one created by the present stock of satellites being too high. With debris and some of the couplings (described below), the congestion externality can be "dynamic" in the sense that the debris stock makes present congestion a function of past launch rates.

The collision rate couplings The collision rate coupling refers to the arguments of the mean collision rate function, $L(\cdot)$. When the couplings are turned off, the collision rate is exogenous: $L(\cdot) = L$. It may vary over time, but since collision rate is exogenous with the coupling turned off, there is no congestion externality. Even if open access causes Kessler Syndrome due to other couplings, the open access launch rate is efficient in the sense that the planner would produce the same outcome. With the collision rate coupling turned off, there are no consequences to excess satellite levels or debris growth.

If the collision rate is coupled with the satellite stock ($L(\cdot) = L(S)$), then there can be a "steady state" congestion externality. The strength of this coupling can be measured by $\frac{\partial L}{\partial S}$. Despite any other couplings and their effects on debris growth, the only source of inefficiency is that open access results in too many satellites being launched relative to the optimal plan. Though the other couplings may produce interesting dynamics in debris accumulation, those dynamics are irrelevant to socially optimal orbit use. Both the open access and socially optimal launch paths will then involve single-period jumps to the steady state. The optimality of the most rapid approach path (MRAP) is driven by the linearity of the objective function and the lack of persistent consequences due to satellite launches. Though the state equations make the setting dynamic, firms and the planner face a sequence of static decisions. This situation is explored further in section 2.1.7.

If the collision rate is coupled with the debris stock ($L(\cdot) = L(D)$) but not the satellite stock, there can be a "dynamic" congestion externality. The strength of this coupling can be measured by $\frac{\partial L}{\partial D}$. The presence of other couplings will influence the nature of this externality, but with just a coupling between debris and collision rate there can be persistent consequences for specific launch histories due to debris accumulation. Depending on the other couplings, the planner may no longer find the MRAP to the steady state optimal, although firms will continue to take the MRAP. As a result of this coupling and others, firms may end up overshooting the steady state, particularly if there are initially low levels of satellites and debris. The "fully coupled" case of collision rate, which is the main focus of this paper, links collision rate to both the satellite and debris stocks ($L(\cdot) = L(S,D)$), allowing for dynamic congestion where the MRAP is not optimal.

The launch debris coupling The launch debris coupling refers to the amount of launch debris created, i.e. the parameter m in the debris law of motion. The strength of this coupling can be measured by m, with higher values of m implying a stronger coupling. When $m = 0$, launches are uncoupled from the debris stock. When collision rate is coupled to the debris stock, this coupling strengthens the externality of launch, since when $m > 0$ launches create debris independent of any collision effects. However, when collision rate is coupled to the debris stock, the launch debris coupling can "stabilize" open access orbit use by forcing firms to internalize some of the persistent effects of their launch on orbital congestion.

The new fragment formation couplings The new fragment formation coupling refers to the arguments of the new fragment function, $G(\cdot)$. We assume that with any coupling, $G(\cdot)$ is a strictly increasing function of its arguments. When the coupling is turned off, the number of new fragments created is exogenous and there are no positive feedbacks between debris. With the collision rate coupling and the launch debris coupling activated, current congestion can still depend on past launch rates, but the persistence may be muted by of the lack of the new fragment formation coupling. If the new fragment function is coupled to the satellite stock, $G(\cdot) = G(S)$, then there can be persistent consequences of excess satellite levels. When the new fragment function is coupled to the debris stock, $G(\cdot) = G(D)$, there can be multiple equilibria: one with low debris and one with high debris. Of the two, only the former will be stable, while the latter will either return to the former or else explode to infinite debris. This multiplicity will exist independent of the collision rate and launch debris couplings. When the collision rate is only coupled to the satellite stock or uncoupled ($L(\cdot) = L(S)$ or $L(\cdot) = L$), the multiplicity is only in debris levels. When the collision rate is fully coupled ($L(\cdot) = L(S,D)$) and the new fragment function is coupled with only the debris stock, multiple equilibria in both satellites and debris can exist: one with low debris and high satellites, and one with high debris and low satellites. As in the previous case, only the low debris equilibrium is stable. As with the collision rate couplings, the strengths of these couplings can be measured by $\frac{\partial G}{\partial S}$ and $\frac{\partial G}{\partial D}$.

2.1.7 Steady state congestion: collisions without debris

Suppose the collision rate does not depend on the debris stock, i.e. $L_t = L(S_t)$. This could be because there is no debris, or because satellites are perfectly shielded against debris. In this case there are no interesting dynamics in the model: open access jumps to the steady state in a single period, as does the planner.

The lack of persistent consequences for specific launch or collision histories means that the congestion problem is effectively static. Present decisions do not affect future decisions. In this case, marginal satellites create a negative congestion externality if they increase the collision rate, i.e. cause a first-order stochastically dominant shift in the distribution of collisions. Satellites are always a good in this case (the marginal value of a satellite to the fleet is always positive). The linearity of the expected social welfare function in collision probabilities suggests that the planner will be indifferent to or like any shift in the collision distribution which does not increase the average collision rate.

2.1.8 Dynamic congestion: collisions with debris

When the collision rate depends on the debris stock, i.e. $L_t = L(S_t, D_t)$, debris creates persistent consequences for specific launch and collision histories. In this case open access can have interesting dynamics, and the planner's optimal satellite accumulation path will in general not be a jump to the steady state. Firms under open access still face a linear objective function and attempt to take a single-period jump to the steady state, but the delayed effect of period t launches on period $t+2$ debris means the system dynamics may not allow such a jump.

Both firms and the planner will launch more satellites when the skies are clear. While the planner will progressively decrease their launch rate to reach the optimal steady state, firms may overshoot the stable region of the state space and end up on a trajectory which results in Kessler Syndrome. This happens because firms set their launch rates in response to debris accumulation one period ahead, but no farther. Open access removes any incentive for firms to care about longer-horizon effects. Even if firms do not launch into Kessler Syndrome, they may still consistently miss the steady state. This can happen because firms will launch at rates which make the future satellite and debris stocks create collision risk greater than the excess rate of return on a satellite. When that happens, firms will stop launching, allowing the stocks to fall again.

2.1.9 Marginal external cost and the optimal launch plan

In this section, we use letter subscripts on functions to indicate partial derivatives with respect to a given argument, e.g. $L_S(S,D) \equiv \frac{\partial L(S,D)}{\partial S}$.

The fleet planner owns all of the satellites in orbit, and controls all launches to maximize the expected net present value of the satellite fleet. This is the "sole owner" benchmark used in other environmental and natural resource settings.[12] In doing so, the planner equates the marginal benefit of another satellite with its social marginal cost. The social marginal cost in this setting is the sum of the opportunity cost of the investment and the collision rate $(r + L(S_{t+1}, D_{t+1}))$, and the effect of the marginal satellite on future collisions and debris growth $(\xi(S_{t+1}, D_{t+1})/F)$. The first two terms are private marginal costs internalized by firms, and the final term is a marginal external cost not internalized by firms.

[12]For example, Gordon (1954) appeals to the sole owner's management of a fishery in showing that the competitive equilibrium is inefficient, and Scott (1955) appeals to the sole owner's management to show cases where short-run competitive equilibria can manage fisheries efficiently.

There are three important functions which are necessary to understand the marginal external cost: the marginal survival rate ($\mathscr{L}(S_t,D_t)$), the growth-launch fragment balance ($\Gamma_1(S_t,D_t)$), and the new fragments from the current stock ($\Gamma_2(S_t,D_t)$).

The marginal survival rate, introduced in section 2.1.5, is

$$\mathscr{L}(S_t,D_t) = 1 - L(S_t,D_t) - SL_S(S_t,D_t). \tag{2.18}$$

$\mathscr{L}(S_t,D_t)$ is the probability that an already-orbiting satellite will escape a collision if one more satellite is added to the fleet. In the standard growth model with an exogenous depreciation rate on capital, this would be the proportion of capital which is undepreciated. With the couplings in $L(S_t,D_t)$, $\mathscr{L}(S_t,D_t)$ is the proportion of undepreciated capital after accounting for the marginal unit's effect on the depreciation rate and capital stock.

The growth-launch fragment balance is

$$\Gamma_1(S_t,D_t) = G_S(S_t,D_t) - m\mathscr{L}(S_t,D_t). \tag{2.19}$$

$\Gamma_1(S_t,D_t)$ is the balance between the growth in debris through collisions caused by the marginal satellite and the chance a satellite which survives the marginal satellite's launch is exposed to the debris created by that launch. When $\Gamma_1(S_t,D_t)$ is positive the marginal satellite causes more collision fragments between satellites and debris through its presence than through its launch debris. When $\Gamma_1(S_t,D_t)$ is negative it is the opposite case - the marginal satellite's launch debris is a bigger threat to satellites which survive collisions than the satellite's effect on the rest of the fleet and the debris.

The number of new fragments from the current stock is

$$\Gamma_2(S_t,D_t) = 1 - \delta + G_D(S_t,D_t) + mSL_D(S_t,D_t). \tag{2.20}$$

$\Gamma_2(S_t,D_t)$ is the number of debris fragments which will enter the next period due to this period's debris. It contains two pieces: the first $(1-\delta+G_D(S_t,D_t))$ is the number of debris fragments which will not decay plus the number of new fragments which will be created in collisions between satellites and debris due to marginal units of debris, and the second ($mSL_D(S_t,D_t)$) is the increase in satellites lost due to launch debris.

With these terms defined, we can express the marginal external cost as

$$\xi(S_{t+1},D_{t+1}) = \underbrace{S_{t+1}L_S(S_{t+1},D_{t+1})F}_{\substack{\text{cost of collisions}\\ \text{caused by marginal}\\ \text{satellite } (\geq 0)}} - \underbrace{\frac{\Gamma_1(S_{t+1},D_{t+1})}{\Gamma_2(S_{t+1},D_{t+1})}S_{t+1}L_D(S_{t+1},D_{t+1})F}_{\substack{\text{relative debris effect}\\ \text{weighted cost of collisions caused}\\ \text{by marginal debris } (\lesseqgtr 0)}}. \tag{2.21}$$

The first term of $\xi(S_{t+1},D_{t+1})$ is always nonnegative, but the second term may be positive or negative. When the second term is negative the growth in new fragments due to a new satellite exceeds the risk caused by launch debris, and the planner would like to see the satellite stock reduced.

The fleet planner's problem is

$$W(S_t,D_t) = \max_{X_t \geq 0} \{\pi S - FX_t + \beta W(S_{t+1},D_{t+1})\} \tag{2.22}$$

$$\text{s.t. } S_{t+1} = S(1 - L(S_t,D_t)) + X_t \tag{2.23}$$

$$D_{t+1} = D(1-\delta) + G(S_t,D_t) + mX_t. \tag{2.24}$$

The planner's optimality condition, expressed in terms of the flow of marginal benefits and costs, is

$$\pi = rF + L(S_{t+1},D_{t+1})F + \xi(S_{t+1},D_{t+1}). \tag{2.25}$$

The above flow condition allows us to determine the socially optimal collision rate.

Proposition 1. *(Optimal collision rate) The planner launches so that the collision rate is equated to the rate of excess return net of the rate of marginal external cost, i.e.*

$$L(S_{t+1},D_{t+1}) = r_s - r - \frac{\xi(S_{t+1},D_{t+1})}{F}. \tag{2.26}$$

Proof. Rearranging equation 2.25 and dividing by F yields equation 2.26. □

Without further assumptions on the model primitives, the marginal external cost is not always positive. The possibility of a negative marginal external cost here comes from the physical function $\Gamma_1(S_t, D_t)$. $\Gamma_1(S_t, D_t)$ is the balance between the growth in debris through collisions caused by the marginal satellite ($G_S(S_t, D_t)$) and the chance a satellite which survives the marginal satellite's launch is exposed to the debris created by that launch ($m\mathscr{L}(S_t, D_t)$). When $\Gamma_1(S_t, D_t)$ is positive the marginal satellite causes more collision fragments between satellites and debris through its presence than through its launch debris. When $\Gamma_1(S_t, D_t)$ is negative it is the opposite case - the marginal satellite's launch debris is a bigger threat to satellites which survive collisions than the satellite's effect on the rest of the fleet and the debris. $\Gamma_2(S_t, D_t)$ is the number of debris fragments which will enter the next period due to this period's debris. It contains two pieces: the first $(1 - \delta + G_D(S_t, D_t))$ is the number of debris fragments which will not decay plus the number of new fragments which will be created in collisions between satellites and debris due to marginal units of debris, and the second $(mSL_D(S_t, D_t))$ is the increase in satellites lost due to launch debris. The fragment balance is weighted by the inverse number of new fragments generated by the current stock.

We assume that $m\mathscr{L}(S_t, D_t)$ is small enough that the marginal external cost is always positive in a steady state. One intuition for this assumption is that m is small. Launch integrators and vehicle providers have incentives to invest in this, as a strong launch coupling implies that a satellite faces some probability of being destroyed by the debris from its own launch. Launch services are generally competitive enough, and buyers of launch services sufficiently risk-averse, to make this reasonable.

Figure 2.1 shows how the satellite-debris dynamics shift private and social marginal cost curves over time. Increases in satellites and debris shift both costs up, with both marginal cost curves shifting up more under open access than under the optimal plan. This offers another intuition for the congestion wedge in orbit use: firms are unable to control how much marginal costs shift as well as the planner would.

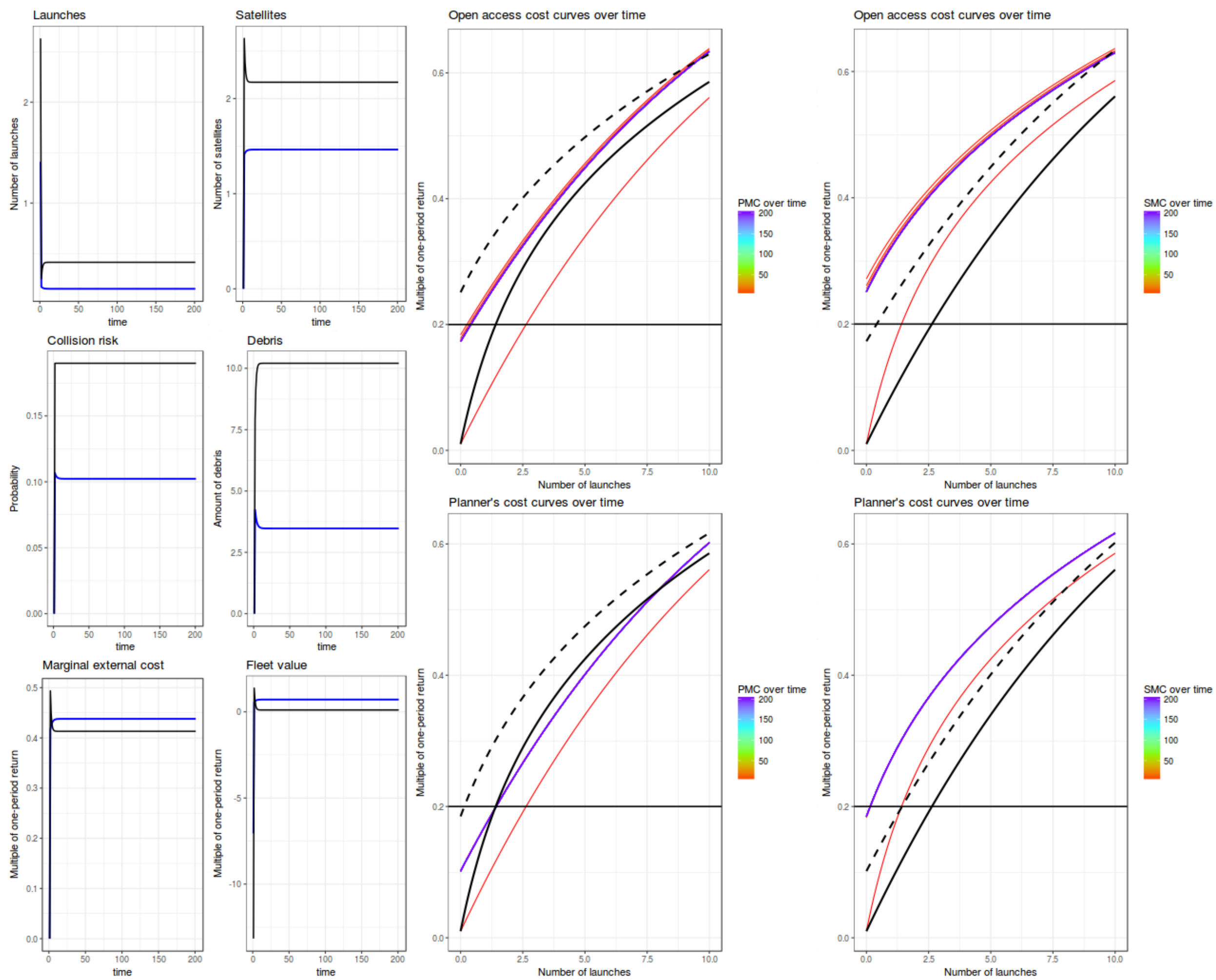

Figure 2.1: Left: time paths under open access (black) and the optimal plan (blue). Right: shifts in private and social marginal cost curves over time under each plan (open access above and the optimal plan below). The middle column of cost curves show how the private marginal cost shifts each period in color, with the initial social marginal cost shown in solid black and the final social marginal cost shown in dashed black. The right column of cost curves show how the social marginal cost shifts each period in color, with the initial private marginal cost shown in solid black and the final private marginal cost shown in dashed black. The marginal benefit of another satellite is the horizontal black line.

2.1.10 Properties of open access equilibria

In this section we explore the properties of open access equilibria. The main results are that the equilibrium collision rate is determined by the excess return on a satellite, that when the collision rate is increasing in new launches (Assumption 1) the collision rate and debris stocks are too large under open access relative to the optimal plan, and that exogenous increases in the debris decay rate or decreases in the amount of launch debris may increase the equilibrium debris stock due to rebound effects. The first two properties establish the nature of open access launching as well as the existence and magnitude of the congestion externality. The third property is relevant to policy proposals which seek to limit debris growth through technical changes, without addressing the economic incentives involved in launching satellites.

Proposition 2. *(Equilibrium collision rate) The equilibrium collision rate along the open access path is equal to the rate of excess return on a satellite over the market asset.*

Proof. Rearranging the open access equilibrium flow condition (equation 3.51) and dividing by F, the equilibrium collision rate can be written as

$$L(S_{t+1}, D_{t+1}) = r_s - r. \tag{2.27}$$

□

In the case where the economic parameters π, F, β are time-varying, Proposition 2 implies that the equilibrium collision rate in period t under open access will be determined by the values of the economic parameters in period $t-1$, i.e. the period when the newest satellites in orbit were launched. In this case, the equilibrium collision rate will not be a constant, but will still be determined by the economic fundamentals of launching satellites. Section A.3 explores this in more detail.

Proposition 2 implies that the equilibrium collision rate is increasing in the profitability of a satellite, decreasing in the cost of launching a satellite, and increasing in the discount factor. Proposition 2 does not imply that the open access equilibrium must be a steady state. Any type of dynamics in the satellite and debris stocks which keep the next-period collision rate equal to the excess return on a satellite can be an equilibrium path.

Corollary 1. *There are multiple open access equilibria, defined by the isoquant where collision rate is equal to the excess return on a satellite.*

Corollary 1 follows from the form of the equilibrium condition. Since the collision rare is equal to excess return everywhere along that isoquant, every point on the isoquant is an open access equilibrium. The initial conditions and physical dynamics determine which equilibrium is reached from a non-equilibrium state.

Corollary 2. *When it is profit-maximizing to launch, firms under open access pursue the most rapid approach path to an equilibrium.*

Corollary 2 follows from the fact that equilibrium under open access requires choosing a launch rate to equate the rate of excess return on a satellite with the next-period collision rate. In a sense, this is an assumption of the model and not a result: if firms equilibriate the collision rate in every period, then they are necessarily following paths directly from the initial conditions to the equilibrium set. Since the laws of motion are linear in the launch rate, the open access paths from the initial condition to the equilibrium set are also linear in the state space. The implications of this type of approach path are more interesting. Unless the equilibrium reached is also a steady state, MRAPs from an initial condition to the equilibrium set are not MRAPs to steady states, and result in overshooting in at least one state variable. This is explored in more detail in Proposition 8.

Corollary 3. *Along the open access equilibrium path, the collision rate is*

(1) increasing in the per-period satellite return,

(2) decreasing in the launch cost, and

(3) increasing in the discount factor.

Corollary 3 shows that policies which increase the cost or reduce the profitability of operating a satellite will reduce the equilibrium collision rate. If firms become more patient, they will be willing to tolerate a higher equilibrium collision rate. Proposition 3 shows the existence of the externality: equilibrium collision rate is too high.

Proposition 3. *(Externality) The open access equilibrium results in higher collision rate than the optimal plan.*

Proof. We denote objects along the open access path with hats, and objects along the optimal path with stars, e.g. $\hat{S}_t$ is the satellite stock in period t under open access and S_t^* is the satellite stock in period t under the optimal plan. The equilibrium collision rate is

$$L(\hat{S}_{t+1}, \hat{D}_{t+1}) = r_s - r, \tag{2.28}$$

while the socially optimal collision rate is

$$L(S^*_{t+1}, D^*_{t+1}) = r_s - r - \frac{\xi(S^*_{t+1}, D^*_{t+1})}{F}. \tag{2.29}$$

Since we assumed $G_S(S,D) > m\mathscr{L}(S,D) \ \ \forall S,D$, the marginal external cost $\xi(S^*_{t+1}, D^*_{t+1})$ is positive, and $L(S^*_{t+1}, D^*_{t+1}) < L(\hat{S}_{t+1}, \hat{D}_{t+1})$. $\square$

It is useful to know how the launch rate and debris stock respond to changes in the model parameters. Particularly policy-relevant parameters include the launch cost, the launch debris, and the decay rate. Pigouvian launch taxes have been suggested to reduce orbital debris, and would manifest as changes in the launch cost; command-and-control policies to reduce the amount of launch debris have been proposed, and reusable launch vehicles reduce the amount of launch debris; decay rates are typically decreasing as altitude increases and policies which reduce the lifespan of orbital debris can be modeled as increasing the decay rate.[13] Proposition 4 considers the effects of changes to the collision rate, satellite stock, and debris stock on the equilibrium launch rate. Proposition 5 considers the effects of changes in the launch costs, launch debris, and decay rates on the equilibrium launch rate and debris stock.

[13] Current FCC policy mandates that inactive satellites must be deorbited within 25 years of launch, which could be captured in the decay rate parameter of this model. Similar policies are encouraged by the IADC and other space agencies.

Proposition 4. *(Economic incentives to reduce congestion) Along the open access equilibrium path, the launch rate is*

(1) decreasing in the current satellite stock, and

(2) decreasing in the current debris stock if and only if the marginal survival rate is nonnegative.

The proof, shown in Appendix C, follows from applying the Implicit Function Theorem to equation 3.51, shown in the appendix. The key inequality from the proof determines how the launch rate varies with the debris stock (period t values are shown with no subscript, and period $t+1$ values are marked with a $'$, e.g. $S_t \equiv S, S_{t+1} \equiv S'$):

$$\frac{\partial X}{\partial D} < 0 \text{ if}$$

$$(S,D): \quad \underbrace{\frac{\partial L(S',D')}{\partial S'} S}_{\substack{t+1 \text{ marginal collision probability from} \\ t+1 \text{ satellites} \cdot \text{present satellite stock}}} < \underbrace{\frac{\frac{\partial L(S',D')}{\partial D'}}{\frac{\partial L(S,D)}{\partial D}}}_{\substack{\text{growth rate of} \\ \text{marginal collision rate} \\ \text{due to debris}}} \cdot \underbrace{\left(1 - \delta + \frac{\partial G(S,D)}{\partial D}\right)}_{\substack{\text{net growth in debris} \\ \text{due to debris}}} \tag{2.30}$$

Inequality 2.30 defines a set of values for (S,D) within which the launch rate is decreasing in the debris stock and outside of which the launch rate is increasing in the debris stock. The condition states that the marginal collision rate in period $t+1$ from satellites in that period across the stock of satellites in t is less than the product of the growth rate of the marginal collision rate from debris and the net growth in debris due to debris. If $1 - L - L_S S \geq 0$, as we have assumed, then $\frac{\partial X}{\partial D}$ is nonpositive. Intuitively, inequality 2.30 states that the launch rate should be decreasing in the debris stock whenever the increase in the collision rate due to debris and debris growth exceeds the increase in collision rate from marginal satellites. This is the essence of "careful placement" of new satellites.

Proposition 5. *(Short-run rebound effects) Along the open access equilibrium path, the current launch rate is*

(1) increasing in the return on a satellite,

(2) decreasing in the amount of launch debris, and

(3) increasing in the debris decay rate.

The next-period debris stock is

(1) increasing in the return on a satellite,

(2) increasing in the amount of launch debris, and

(3) decreasing in the debris decay rate.

The proof is shown in Appendix C. The potential for policies intended to reduce future debris stocks to have perverse effects arises from a combination of Propositions 4 and 5. Policies which reduce amount of time before inactive satellites must deorbit (i.e. increase the debris decay rate) can lead to better steady states, shown in Proposition 9, but the convergence may be non-monotonic. The initial reduction in debris spurs an increase in the launch rate, which may be too large to reach the steady state right away. Figure 2.2 shows an example of how a non-monotonic steady state transition caused by an increase in the decay rate can lead to higher short-run levels of debris.

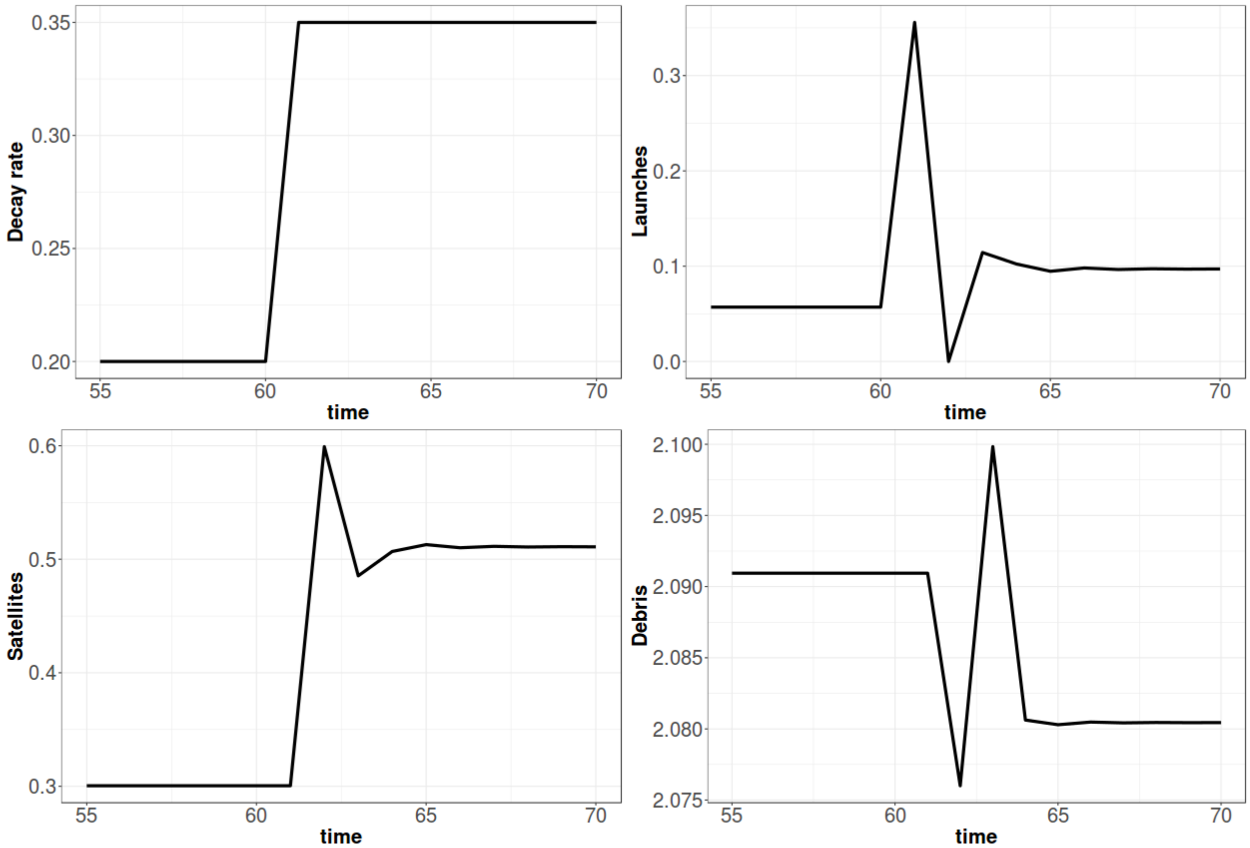

Figure 2.2: An example of an increase in the rate of debris decay, δ, causing non-monotonic steady state convergence. Though the debris stock initially falls, the launch response is large enough that the debris stock rises above its previous level before reaching the new steady state.

These effects stem from the substitutability of the rate of return on a satellite and the collision risk in the equilibrium condition (equation 3.51). Since firms will enter the commons until the expected profits are zero, reducing the risk along one margin (e.g. less launch debris) can result in firms increasing risk along another margin (e.g. more launches). Since the launch debris and decay rate effects are (a) the opposite of what a purely physical model which ignores open access would predict, and (b) they occur between steady states, we refer to them as short-run rebound effects. Proposition 9 in section 2.2.1 explores the long-run effects of the same parameter changes on debris levels. Figure 2.6 illustrates numerically how the statics results in Proposition 5 interact with the stability results in Proposition 7 across thousands of time series simulations.

Economic controls which increase the cost of launching will have the expected effects: fewer launches and less debris. The results in Proposition 5 do not account for the fact that command-and-control policies which decrease the amount of launch debris or increase the decay rate may also increase the cost of launching a satellite. Such feedbacks could reduce or remove the perverse effects. The takeaway is that in order for policies targeting launch debris generation and debris decay rates to be effective at reducing the equilibrium debris stock, they must affect the incentive to launch a satellite in the first place.

2.1.11 Action and inaction

Under open access, firms will not launch satellites if the excess return is insufficient to cover the congestion cost. Holding the launch rate constant, collisions and debris growth may drive an orbit from the action region, where some find it economical to launch, to the inaction region, where it is uneconomical for any to launch. Launch shutdown in some periods may be socially optimal. As long as Kessler Syndrome doesn't occur, the physical dynamics of the system will eventually bring satellite and debris stocks back to the action region. Figure 2.3 shows an example of action and inaction regions in the open access and optimal launch policies, and figure 2.5 shows the associated fleet values.[14]

Definition 1. *(Action region) The action region is the set of satellite and debris levels for which open access results in launches, i.e.*

$$(S_t, D_t) : L(S_{t+1}, D_{t+1}) \leq r_s - r.$$

[14]Figure 2.4 shows time paths where open access bounces between the regions and the planner does not.

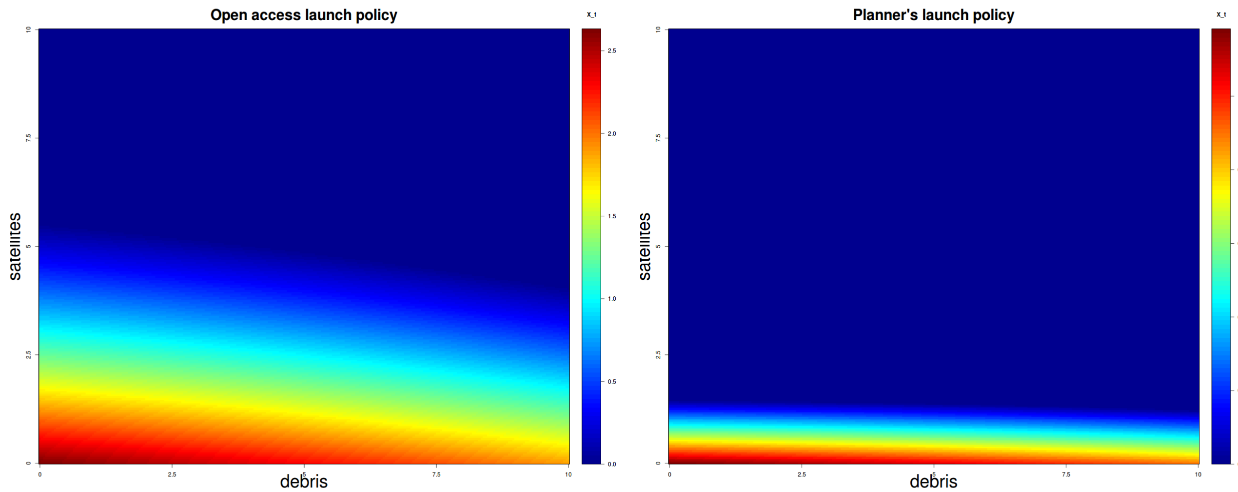

Figure 2.3: An example of the action and inaction regions under open access (left) and the optimal plan (right). Open access does not restrict launches in response to satellite or debris stocks as quickly as the planner or as severely.

The inaction region is the complement of the action region. Under open access, the inaction region is where the collision rate is greater than the excess return on a satellite. In the inaction region, aggregate dynamics are driven entirely by the collision rate, decay rate, and new fragment production (the "physical dynamics"). The laws of motion become

$$S_{t+1} = S_t(1 - L(S_t, D_t)) \tag{2.31}$$

$$D_{t+1} = D_t(1 - \delta) + G(S_t, D_t). \tag{2.32}$$

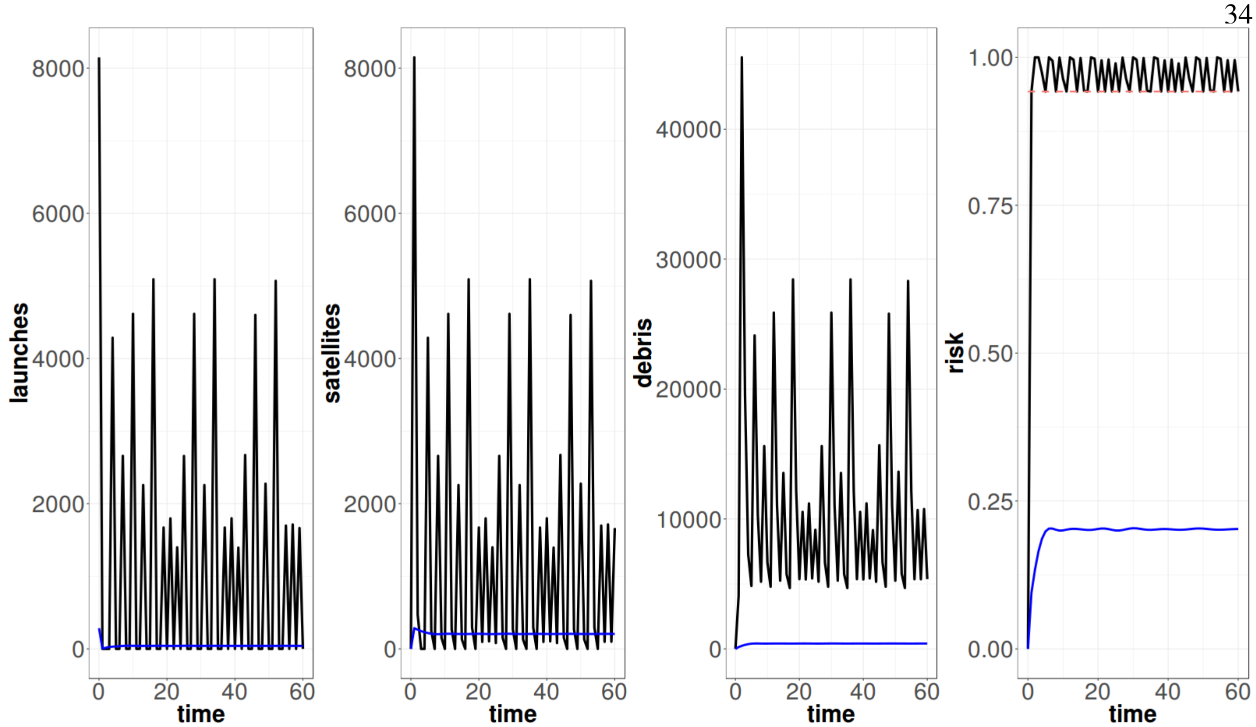

Figure 2.4: An example of open access time paths which bounce between action and inaction regions (black). The planner (blue) does not find this optimal.

One possible steady state under these laws is to have no satellites or debris. Before that happens the stocks would again reach the action region, and the zero objects steady state will not be reached. The other possibility is that, as the satellite stock is driven to zero, the debris stock reaches the threshold at which Kessler Syndrome occurs. Though open access reduces launches in response to exogenous increases in the steady state debris level, the point on the equilibrium isoquant reached by a MRAP from the initial conditions may be in a place where the physical dynamics lead to the Kessler region.

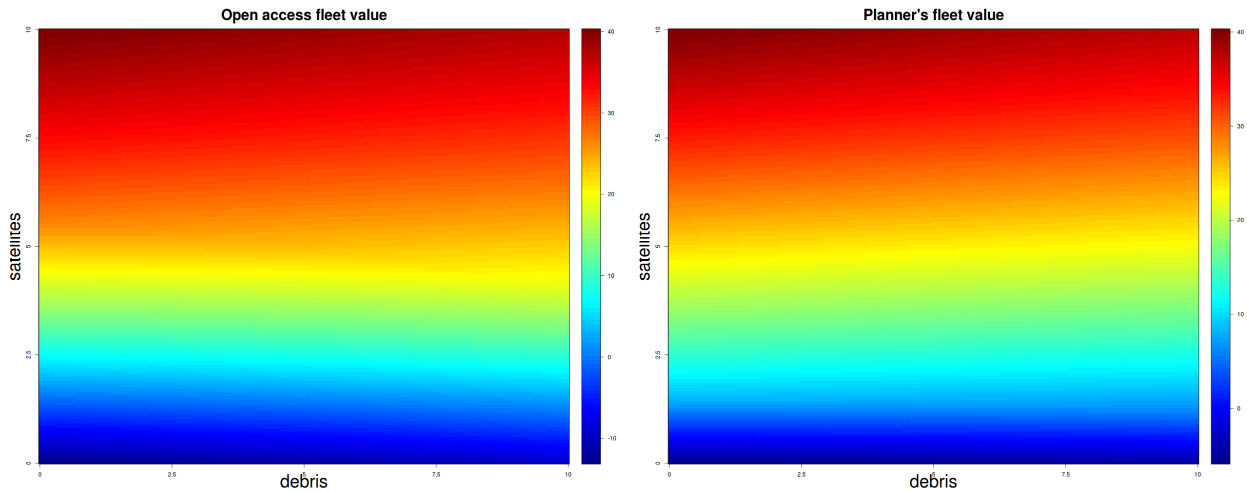

Figure 2.5: An example of the fleet value functions under open access (left) and the optimal plan (right). The planner's relatively muted launch policy creates a smaller region of negative profits and a larger region of positive profits. Losses under the open access plan are steeper, larger fleets less valuable, and higher debris stocks more costly than under the planner's policy.

2.2 The economic dynamics of orbital mechanics

Having described the nature of open access equilibria, we now turn to their dynamic properties. Two issues of first-order policy concern are the long-run properties of open access equilibria, and the conditions under which open access will prevent or cause Kessler Syndrome. In this section, we focus on understanding these issues by analyzing open access steady states and by formally characterizing Kessler Syndrome and its relation to open access equilibria. We also consider the effects of other dynamic issues affecting orbit use, namely time-varying economic parameters and space weather. Our main results are that there can be multiple steady states of varying stability, and that increases in the excess return on a satellite shift the equilibrium set outward and bring it close to the region of the state space where Kessler Syndrome occurs. Allowing rates of return to vary over time does not change these properties, though it highlights the role of launch costs as a source of satellite value and suggest that Pigouvian launch fees will need to account for this in order to be efficient. Space weather does not change the properties of the equilibrium set, though it shifts its location and in doing so may cause Kessler Syndrome.

2.2.1 Properties of open access steady states

The distinction between a "steady state" and an "equilibrium" is relevant here. The laws of motion for satellites and debris may be stationary under constant non-equilibrium launch rates, a case frequently analyzed in the physics and engineering literature. Another possibility is for firms to launch until zero profits in every period, while the satellite and debris stocks vary over time. As long as the collision rate in each such period is equal to the excess return on a satellite, the firms are in an open access equilibrium. What we are interested in here are states where both conditions are true: the physical aggregates are stationary over time and firms are earning zero economic profits. We refer to these as "open access steady states".

In an open access steady state, the collision rate must be equal to the excess return on a satellite ($L(S,D) = L(S',D') = r_s - r$) along with the usual conditions that the aggregate variables be stationary:

$$(S,D) : L(S,D) = r_s - r, \tag{2.33}$$

$$\begin{aligned} S' = S &\implies S = (1 - L(S,D))S + X \\ &\implies X = L(S,D)S, \end{aligned} \tag{2.34}$$

$$\begin{aligned} D' = D &\implies D = (1-\delta)D + G(S,D) + mX \\ &\implies \delta D = G(S,D) + mX. \end{aligned} \tag{2.35}$$

Equations 2.33, 2.34, and 2.35 define the open access steady states. If there is no launch debris ($m = 0$), then steady states only require balancing natural debris decay (δD) against debris growth due to objects already in orbit ($G(S,D)$). As m increases, the effect of replacement satellites (reflected in mX) on the steady state level of debris also increases, requiring a lower rate of debris growth due to objects already in orbit.

Proposition 6. *(Multiplicity) There can exist multiple open access steady states.*

The proof is shown in Appendix C. The exact open access steady state reached is determined by the physical and economic dynamics. Not all open access steady states are locally stable. When there are multiple stable open access steady states, the one reached will depend on the initial conditions. Shocks to the debris stock, like weapons tests by national authorities, could move the system around in the same basin or to a different one.

Proposition 7. *(Local stability) Open access steady states may be locally unstable.*

The proof is shown in Appendix C. The key equation from the proof is

$$\frac{\partial \mathscr{Y}}{\partial D}(D) = -\delta - \frac{L_D(S,D)}{L_S(S,D)}(G_S(S,D) + m(r_s - r)) + G_D(S,D). \tag{2.36}$$

The stability of an open access steady state depends on three factors, shown on the right hand side of equation 2.36. The first is the decay rate (δ), which can ensure stability if it is high enough. Physically, higher decay rates indicate a region with greater renewability. The second is the equilibrium effect of new satellites on debris growth. The effect is increasing in the strength of the collision rate debris coupling (L_D), the new fragment satellite coupling (G_S), and the launch debris coupling (m). These are all launch byproducts deter future launches through debris creation and reduce the number of satellites which can be sustained. The effect is decreasing in the strength of the collision rate debris coupling (L_S), which determines the effect of new satellites on the collision rate through their presence in orbit. The third is the strength of the new fragment debris coupling, which creates the potential for local instability if it is large enough. Physically, it measures the strength of the positive feedback between debris. The same coupling creates the possibility of Kessler Syndrome.

Open access steady states with less debris are more likely to be locally stable if the rate of excess return ($r_s - r$) is not too high, or if the coupling between new fragments and active satellites (G_S) is large enough. It depends linearly on the convexity of the new fragment satellite coupling (G_{SD} and G_{SS}), and nonlinearly on the strength and convexity of the collision rate couplings (L_S, L_D, L_{SD}, and LDD). Short-run dynamics of the type described in propositions 4 and 5 can interact with local instability to prevent the system from reaching a stable steady state, as shown in figure 2.6.

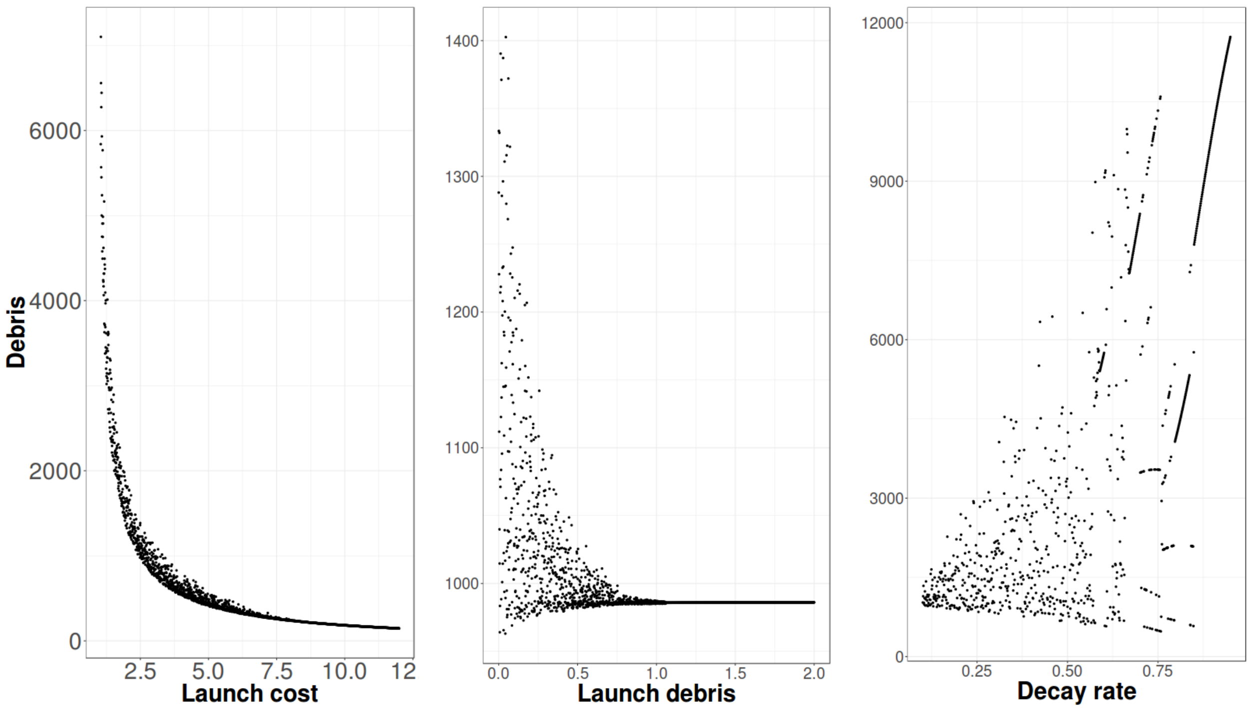

Figure 2.6: Each point is a long-run ($t = 1000$) debris level for different values of F, m, and δ under open access. The spreads indicate the presence and amplitudes of cycles in debris stocks.

Proposition 8. *(Overshooting) For each open access steady state, there are two disjoint sets in the action region. Paths from one will initially overshoot the open access steady state debris level, and paths from the other will initially overshoot the open access steady state satellite level.*

Proof. (Sketch) The full proof is shown in Appendix C. The intuition is that there is a limited set of points which can reach the open access steady state in one period, and that all other points in the action region are on paths which overshoot at least one state variable in approaching an steady state via the equilibrium isoquant. □

Proposition 8 highlights two properties of open access. First, open access attempts to immediately equilibriate the collision rate by moving to the isoquant where it is equal to the excess return, even if the point on the isoquant that can be reached in one period is not a steady state. This is analogous to a driver speeding to reach their destination and forgetting to decelerate as they approach. Second, this "uninternalized acceleration" problem is more severe at low levels of debris than at higher levels, provided the debris level is below the set leading directly to the steady state. In the driving analogy, this is the effect of starting location on the uninternalized acceleration: when the driver is farther away, they must accelerate more and go faster to reach their destination in the same amount of time, meaning that forgetting to decelerate has higher consequences. Economically, open access induces firms to ignore longer-term environmental consequences of their scramble to take advantage of cleaner orbits. Overshooting can cause Kessler Syndrome if the portion of the equilibrium isoquant reached is in the basin of the Kessler region. Figure 2.7 shows a case where open access and the planner reach steady states.

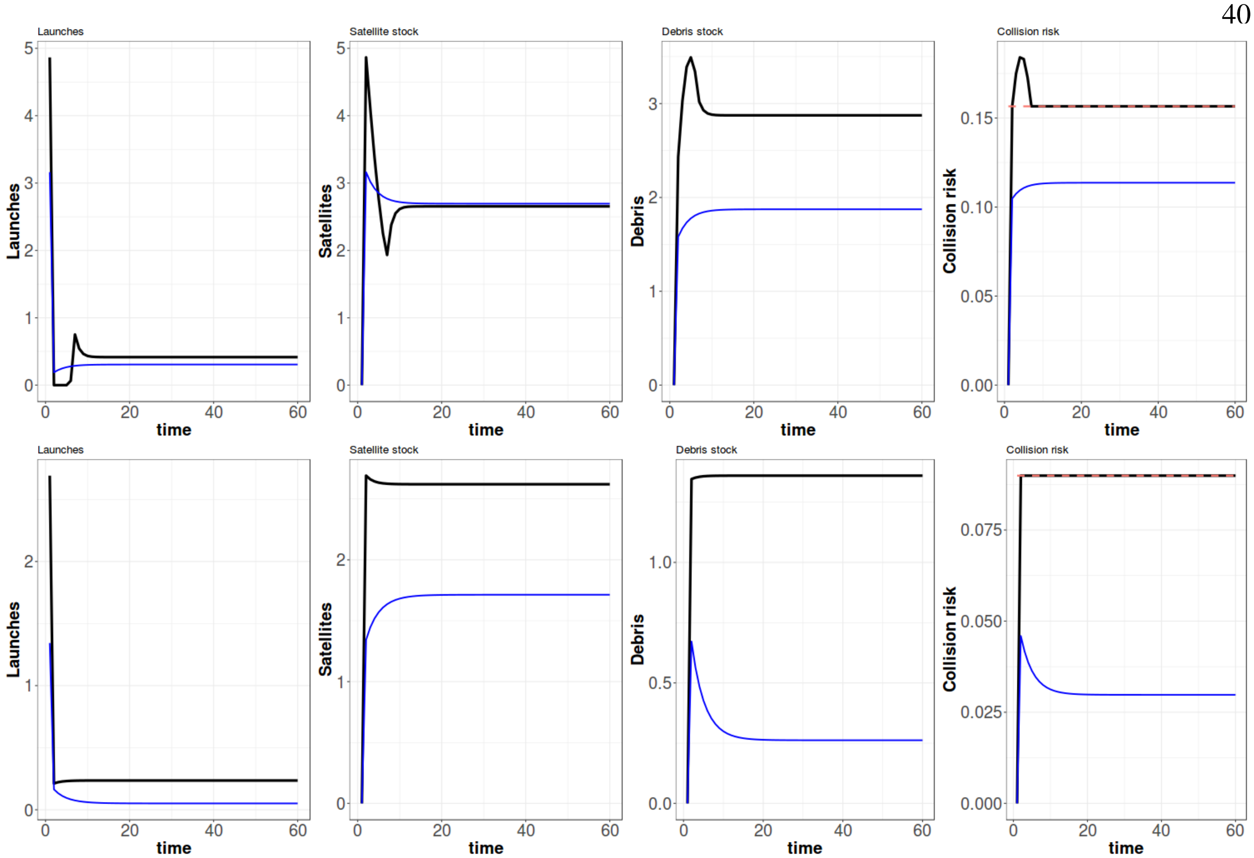

Figure 2.7: Examples paths which reach the steady state under open access (black) and the optimal plan (blue). Open access is not as responsive as the planner to delayed debris accumulation, which can result in firms under open access launching into the inaction region (left).

The overshooting becomes more severe as the initial condition and physical dynamics bring firms to points in the action region which are farther from the line segment leading to the open access steady state. To see this, note that all open access launch plans are parallel lines in the satellite-debris space. Points farther from the line segment leading to the open access steady state in question will therefore end up farther from that steady state. Given the potential multiplicity of open access steady states, however, these points may lead to a different open access steady state. Since they are all on the same isoquant, reaching one of many steady states implies that other steady states have been overshot. This result suggests some scope for orbit use stabilization policies which simply impose a fixed cap on the number of launches per period, since slowing the rate at which firms approach the equilibrium isoquant can reduce the amount of overshooting. Figures 2.8 and 2.9 show examples of overshooting in the open access and planner's phase diagrams.

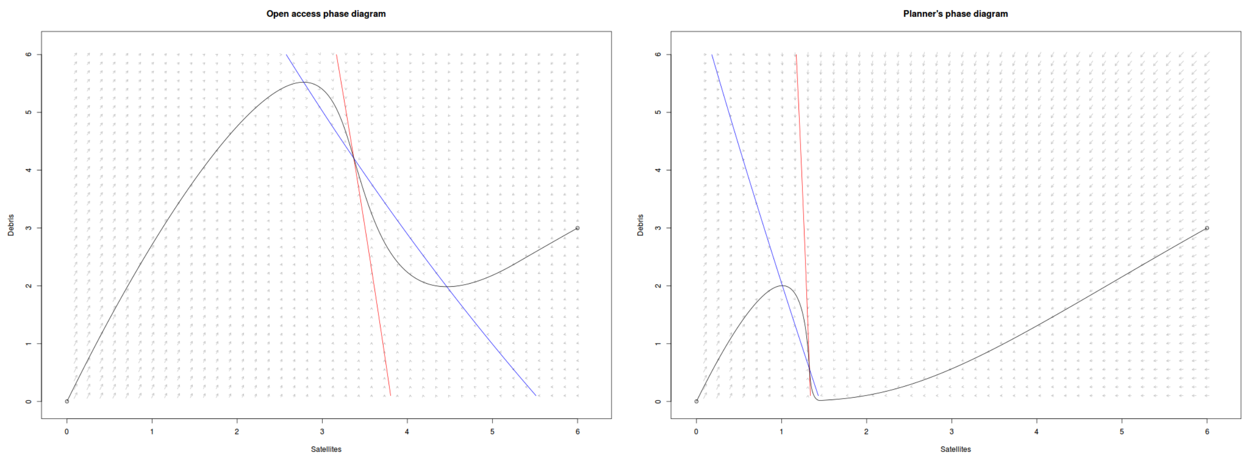

Figure 2.8: Left: open access phase diagram of the satellite-debris system with trajectories. Right: optimal plan phase diagram of the satellite-debris system. The blue line is the debris nullcline, the red line is the satellite nullcline, and the black lines are trajectories from $(0,0)$ and $(6,3)$.

In figure 2.8, both open access and the planner overshoot when starting from no satellites and no debris. In figure 2.9, open access overshoots when starting from no satellites and no debris, but the planner doesn't.

Proposition 9. *(Long-run debris levels) Open access steady state debris levels are*

(1) increasing in the return on a satellite and the amount of launch debris if and only if the open access steady state is locally stable, and

(2) decreasing in the decay rate if and only if the open access steady state is locally stable.

The proof is shown in Appendix C. When the open access steady state is locally stable, long-run debris levels have the expected signs under small changes in return rates, decay rates, and launch debris levels. These signs are reversed when the open access steady state is locally unstable. Unstable steady states will not be where the system ends up in the long-run, so a thought experiment starting from one of these states is on questionable footing. Instead, consider a steady state near the boundary of stability. For such steady states, the effects of technology shocks affecting the return on a satellite or the amount of launch debris produced, or sunspots affecting the decay rate (see section A.1), can be ambiguous. If a combination

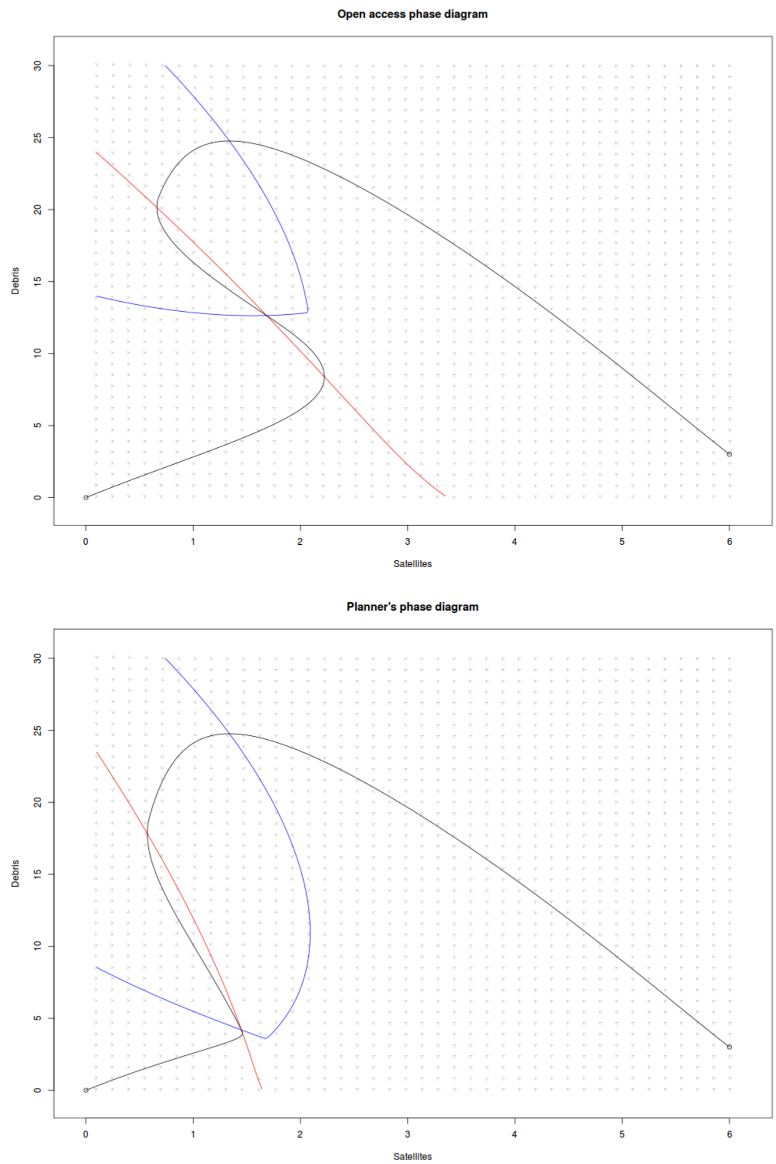

Figure 2.9: Left: open access phase diagram of the satellite-debris system with trajectories. Right: optimal plan phase diagram of the satellite-debris system. The blue line is the debris nullcline, the red line is the satellite nullcline, and the black lines are trajectories from $(0,0)$ and $(6,3)$. The parameterization is the same as figure 2.8 except for the numbers of new fragments created in collisions (stronger new fragment couplings with satellites and debris in this figure).

of shocks pushes an open access steady state over the boundary into instability - for example, an increase in debris due to military activities coinciding with sunspot activity - the system will move to a different steady state with more debris. These types of long-run effects can matter if the short-run dynamics preclude moving smoothly between nearby stable steady states. In such cases the short-run dynamics and the number and stability of open access steady states will determine the long-run behavior of the system.

2.2.2 Kessler Syndrome

The occurrence of Kessler Syndrome is a key concern in managing orbit use. If open access can prevent Kessler Syndrome, regulating orbit use is not as important from an environmental perspective. Even though orbit use will be inefficient it won't cause irreversible environmental damage. On the other hand, if open access can cause Kessler Syndrome, orbit use management becomes more urgent.

In this section we formally define Kessler Syndrome and establish some properties of the debris threshold beyond which it occurs. Our main result, Proposition 10, is that open access debris levels are increasing in the excess return on a satellite while the Kessler threshold is constant, implying that sustained increases in the return on a satellite can cause Kessler Syndrome under open access. Though the Kessler threshold is defined purely in terms of the system's physics, the occurrence of Kessler Syndrome depends critically on the economics of orbit use.

Assumption 2. *(Debris growth) The growth in new fragments due to debris is larger than the decay rate for all levels of the satellite stock greater than some level $\bar{D} > 0$,*

$$\bar{D} : G_D(0,D) > \delta \ \forall D > \bar{D}.$$

Assumption 2 and $G(S,D)$ being increasing in both arguments, there is a unique threshold $D^\kappa \geq \bar{D}$ above which Kessler Syndrome occurs. Figure 2.10 shows an example of an open access path which causes Kessler Syndrome. Kessler Syndrome occurs when the number of new fragments created by collisions between debris exceeds the amount which decays in a single period. For regimes where this condition doesn't hold at any level of debris, Kessler Syndrome is impossible.

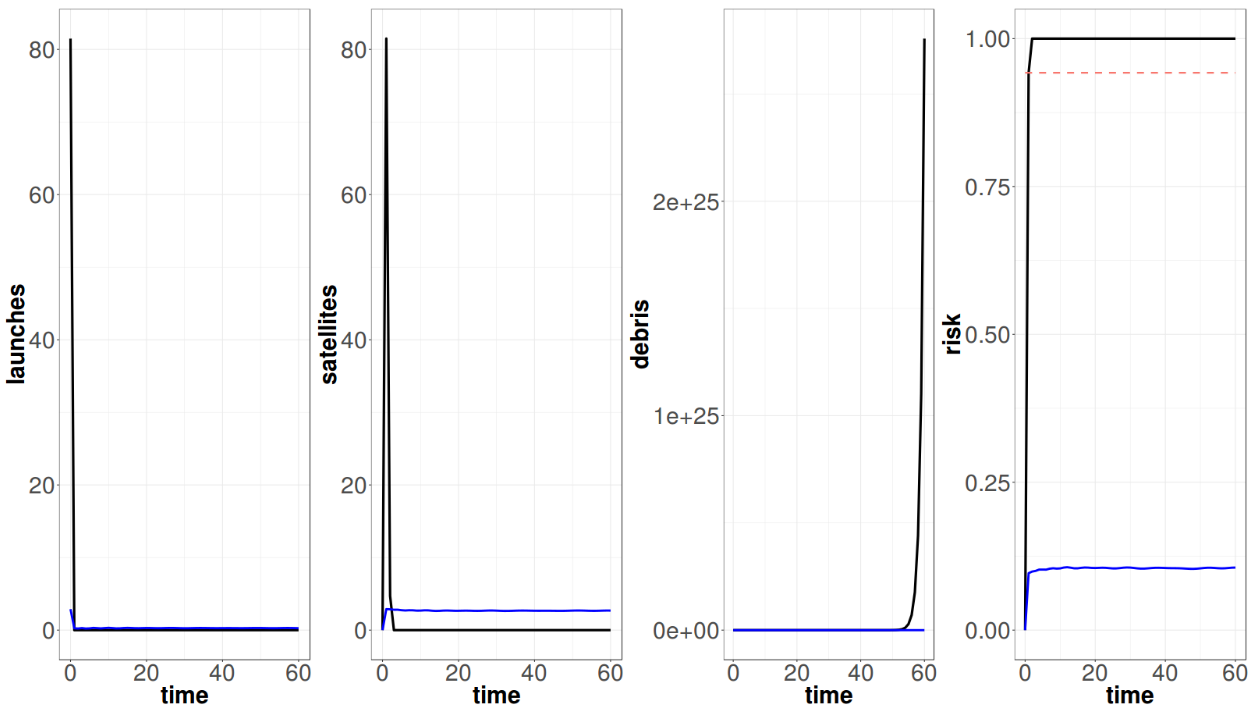

Figure 2.10: An example of open access causing Kessler Syndrome (black) when the planner would avoid it (blue).

Definition 2. *(Kessler Syndrome) The Kessler region is the set of debris levels for which cessation of launch activity and immediate deorbit of all active satellites cannot prevent continued debris growth, i.e.*

$$D^{\kappa} : G(0, D^{\kappa}) > \delta D^{\kappa}.$$

Kessler Syndrome has occurred when the debris stock enters the Kessler region.

Our assumption that the new fragments function is monotonically increasing in both arguments, and the lack of debris removal technologies, imply that Kessler Syndrome is an absorbing state.

The Kessler threshold is entirely determined by the physical properties of the orbit. As the open access steady state debris level rises, say due to a change in the decay rate, the open access launch rate decreases (Proposition 9). The open access steady state debris level is by definition less than or equal to the Kessler threshold, so the open access launch rate is decreasing as steady state debris levels approach the threshold. However, this does not imply that open access will never cause Kessler Syndrome; open access launch policies may overshoot the steady state due to debris growth in period $t+1$ (Proposition 8). While steady states are inconsistent with Kessler Syndrome, open access equilibria need not be.

Corollary 4. *The Kessler threshold is increasing in the decay rate.*

This is an intuitive property of the physical system: as its absorptive capacity grows, the level at which its capacity is overwhelmed also grows.

Proposition 10. *(Kessler Syndrome and open access) If there are sustained increases in the return on a satellite, open access will eventually cause Kessler Syndrome.*

Proof. The result follows from Proposition 5 and the fact that the Kessler threshold is independent of the economics of satellite use. Since D^{κ} is constant with changes in r_s, and the open access debris level $\hat{D}_{t+1}$ is increasing in r_s, sustained increases in r_s will eventually push $\hat{D}_{t+1}$ over D^{κ}. $\square$

Proposition 10 shows that profit motives are not necessarily sufficient to prevent Kessler Syndrome, and may in fact cause it. While the open access steady state will avoid the Kessler region by definition, open access equilibria need not. Because of this, when the excess return on a satellite rises, Kessler Syndrome may occur in the transition from the original steady state to the new one.[15]

2.2.3 The risk to LEO

LEO is one of the fastest growing segments in the satellite industry today, particularly for smaller satellites with fewer guidance and control systems (Brodkin (2017), Selk (2017), Dvorsky (2018)). LEO users typically have shorter planning horizons than GEO users and face lower costs. Propositions 4, 5, 9, and 10 reinforce the conclusions from physical models of debris growth: LEO is at the highest risk of Kessler Syndrome (Liou and Johnson (2008), Kessler et al. (2010)). Our results show that this is not just a feature of the physics given current use patterns, but of the economics underlying those use patterns. While higher decay rates in LEO may protect it against Kessler Syndrome, Proposition 10 shows that the economic properties of LEO - particularly the low cost of access - work in the opposite direction. Attempts by national authorities to unilaterally prevent satellites from being launched face a leakage problem similar to the one faced by national authorities attempting to unilaterally control pollution emitted by mobile capital - firms can leave the regulated area and launch from an unregulated one.[16]

Larger collision or fragmentation parameters are known to increase the chance of Kessler Syndrome (e.g., Rossi et al. (1998), Kessler and Anz-Meador (2001), Liou (2006)). Adilov et al. (2015) have shown that lower launch costs will increase the equilibrium amount of debris. We reaffirm these findings, and show that lower launch costs will also increase the equilibrium collision rate. An increase in the collision and fragmentation parameters could be caused by cost-minimizing satellite launchers opting to launch cheaper

[15] Though we express this result elsewhere in the paper in terms of "chance" or "likelihood", these are only informal uses for expository purposes. There is no probability measure defined anywhere in the model, so no sense in which one outcome can be "more likely" than another.

[16] Dvorsky (2018) documents what may be the first instance of launch leakage: a California startup denied launch permission by the FAA went outside the FAA's jurisdiction and purchased a launch from an Indian launch provider instead. The FAA denied permission on the grounds that the startup's satellites were too small to be effectively tracked and would increase the risk of unavoidable collisions and debris growth.

satellites with fewer guidance and control systems and less shielding. A reduction in launch costs could occur independent of changes in satellite characteristics, for example as more firms enter the launch market and drive launch prices down.

These issues are not as pronounced in GEO in part because property rights have been defined for GEO slots. Our model suggests that other characteristics of GEO, in particular the high cost of access and the low natural decay rate, create additional economic incentives to reduce debris accumulation and collision risk there.

2.3 Conclusion

There are potentially serious problems associated with open access to Earth's orbits. In this paper we present the first long-run economic model of orbit use, explore the nature of the congestion externality associated with launching satellites, and study policy implications of this externality. We highlight three key messages for readers concerned for the future of Earth's orbits.

First, too many firms will launch satellites because they won't internalize the risk they impose on other orbit users. Though profit maximizing satellite owners have incentives to reduce launches as risk of a collision grows, they do not respond to debris growth or the risk of collisions optimally. Second, rebound effects can result in higher debris levels when the rate of natural renewal is high and when the debris created by a launch is low. Combined with the low cost of access, these effects suggest that lower regions of LEO may end up more congested than higher regions. Third, Kessler Syndrome is more likely under open access as the excess return on a satellite rises. As launch costs fall and new commercial satellite applications become viable, open access will be more likely to cause Kessler Syndrome despite launch reductions in response to orbital congestion.

Economists tend to focus on property rights, corrective taxes, or other market-based mechanisms to solve externality problems. While these mechanisms may prevent Kessler Syndrome and ensure efficient orbit use, orbital management policies should be designed in light of the unique physical features of this resource. Orbits are a global commons and will likely require global policy solutions. These solutions may be enforced by states, arise as self-enforcing agreements between private actors, or some combination of the two. The problems of social cost in orbit, of interacting congestion and pollution with endogenous regime change, are similar to the problems of climate change. The fact that economic policy solutions to climate change have proven difficult to build consensus around and implement suggests enacting policy solutions to orbital externalities may not be easy either.

Chapter 3

Economic Principles of Space Traffic Control

Open access to common-pool resources tends to cause resource overuse or stock collapse (Gordon, 1954; Hardin, 1968; Dietz et al., 2003). Despite awareness of this fact for over one hundred years, new and existing common-pool resources are still plagued by open access problems (Stavins, 2011). Open access orbit use has led to the accumulation of orbital debris, from nonoperational satellites to nuts, bolts, and propellant fuel particulates. Collisions between orbiting bodies can shatter satellites into thousands of dangerous high-velocity fragments, some of which may be too small to track. Runaway debris growth, known as Kessler Syndrome, threatens to render high-value orbits unusable for decades or centuries (Kessler and Cour-Palais, 1978). As technology makes satellites cheaper to launch and more reliable, firms are planning to launch thousands of satellites into already-congested orbits. The need for policies to manage orbital congestion is more pressing than ever. While commons management problems have been studied extensively (Weitzman, 1974; Ostrom, 1999; Newell and Pizer, 2003; Costello et al., 2008), engineers, economists, and policymakers know little about how space traffic should be managed and debris removal technologies should be employed. In this chapter I answer two fundamental questions of space traffic control. First, what do optimal space traffic control policies look like? Second, how should active debris removal be employed? The key insights of this chapter are that space traffic control policies should target satellites in orbit rather than satellite launches, and that satellite owners must pay for debris removal for it to reduce equilibrium collision risk.

In this chapter, I derive economic principles of space traffic control policy in the first dynamic model of satellite launch and ownership with physical uncertainty over collisions and positive feedbacks in debris growth. I highlight the key policy design constraints imposed by open access and show how the use of active debris removal technologies will affect equilibrium collision risk and debris growth. I show that despite uncertainty over the risk of catastrophic collisions, the traditional "prices vs. quantities" question is moot. Price or quantity policies can achieve first-best outcomes because both regulators and firms are equally uncertain about the collision risk. The key design issue is whether the regulator's policy targets satellites in orbit (for example, a satellite tax) or the act of launching satellites (for example, a launch tax).[1] In the setting I study, regulating satellites in orbit achieves higher expected social welfare than regulating the act of launching satellites. Regulating satellite launches instead of satellites in orbit creates rents to satellite ownership and induces suboptimal spikes in equilibrium collision risk just before the policy takes effect. Satellite launch controls are also limited in their ability to induce deorbits, and optimal satellite launch controls have unfavorable dynamic properties. Contrary to predictions from non-economic models of orbit use, active debris removal may reduce the debris stock without affecting equilibrium collision risk. If satellite owners receive debris removal for free, more launchers will enter to take advantage of the cleared space. For active debris removal to reduce equilibrium collision risk, satellite owners must bear the cost of removal.

Prior analyses have quantified the costs and benefits of mitigating and reducing debris and collision risk (Liou and Johnson, 2008, 2009; Bradley and Wein, 2009; Ansdell, 2010; Schaub et al., 2015; Macauley, 2015), noted that open access and the common-pool nature of orbits make rational actors ignore their effects on other orbit users (Merges and Reynolds, 2010; Weeden and Chow, 2012; Adilov et al., 2015; Salter, 2015), and considered necessary legal and institutional features that an orbit use management policy framework ought to have (Weeden, 2010; Weeden and Chow, 2012; Akers, 2012).[2] I build on these analyses

[1] While I study principles relevant to orbital regulation, I do not explicitly analyze the problem of a global orbit use regulator tasked with creating an efficient and self-enforcing international agreement on orbit use. Such analysis is eminently important to the problem of orbit use, but beyond my scope in this chapter.

[2] The legal issues of debris removal are non-trivial. International space law gives satellite operators ownership of their debris even after the satellite's lifetime, forcing potential salvage operations to negotiate with each individual satellite or fragment owner for

by formally modeling open access incentives with a realistic dynamic structure which reveals feedbacks between the environment and orbit-users. This structure allows me to identify issues with launch taxes and publicly-provided debris removal not visible in earlier studies which did not include maximizing behavior or realistic dynamics. My results show that ignoring these issues can result in welfare losses as rational orbit-users attempt to capture the rents created by launch taxes or dissipate the rents created by debris removal.[3]

I contribute to the literature on common-pool resource management and orbit use in three ways. First, I present the first economic analysis of orbit use management policy under open access and aggregate physical uncertainty. Prior models of orbit use with policy recommendations such as Bradley and Wein (2009), Weeden and Chow (2012), Macauley (2015), and Adilov et al. (2015) do not simultaneously account for open access, forward-looking investment behavior and dynamics, and physical uncertainty in collision risk while developing policy prescriptions. Accounting for these features together reveals novel insights, such as the fact that regulating satellites in orbit is preferable to regulating satellite launches since satellite controls affect only the private marginal cost of launching while launch controls affect both the private marginal benefit and cost of launching.[4]

rights to remove debris. For small fragments, attributing ownership and negotiating removal may be infeasible. A substantial legal and engineering literature has considered these issues and potential solutions, for example Carroll (2009); Merges and Reynolds (2010); Ansdell (2010). Many of these scholars have suggested amending existing legal frameworks to allow salvage and debris removal bounties, so as to incentivize negotiations between debris owners and would-be debris removers. Yet the march of technology continues despite legal uncertainty, and debris removal technologies are being developed and tested, for example Pearson et al. (2010).

[3]To be clear, the need for orbit use management policy is not uniform across all orbits. Formal orbit allocation procedures exist in the geosynchronous (GEO) belt (Macauley, 1998; Jehn et al., 2005). But in low-Earth orbit (LEO), no such procedures exist. Orbital management must be done indirectly through spectrum management by national authorities (such as the Federal Communications Commission in the US), or directly through non-binding guidelines from international agencies (such as the Inter-Agency Debris Committee).

[4]There are similarities between the orbit use problem and other global commons or congestible resource problems, such as regulating atmospheric carbon dioxide or controlling road traffic congestion. Satellite controls are analogous to congestion pricing, while launch controls are analogous to road access tolls or gasoline taxes. Directly pricing congestion is more effective at controlling

Second, I explicitly incorporate dynamic feedbacks and physical uncertainty in orbit use and study their economic effects. This allows my model to be augmented with high-fidelity engineering models and provide policymakers with quantitative policy design guidance. Prior models of orbit use such as Macauley (2015) and Adilov et al. (2015) focus on qualitative properties of orbit use. While Macauley (2015) estimates tax and rebate values for a range of space traffic control policies, the values are derived in a two-period framework which obscures the effects of dynamic feedbacks, physical uncertainty, and open access launch behavior. Accounting for these effects is necessary to provide real-time quantitative guidance for optimal or second-best policies.

Third, I present the first economic analysis of the effects of active debris removal on orbit use, accounting for profit-maximization and open access. Prior models of active debris removal such as Liou and Johnson (2009), Bradley and Wein (2009), Carroll (2009), and Ansdell (2010) do not account for economic behavior in studying the physical or legal issues in debris removal. Muller et al. (2017) accounts for profit motives when deriving a lower bound on the value of debris removal, and Klima et al. (2016) accounts for physical dynamics and strategic behavior between debris removers when analyzing debris removal, but neither account for open access launch behavior or compare the resulting outcomes to first-best outcomes. My modeling framework accounts for profit-maximization and open access alongside physical dynamics and identifies previously-unknown issues in orbit use management, such as the relevance of how debris removal is financed to equilibrium outcomes and the inefficiency in cooperative removal plans created by open access launching.[5] My modeling framework can also provide quantitative guidance regarding optimal debris removal policy and the size of open access distortions to debris removal.

The remainder of the chapter is organized as follows. In section 3.1 I describe institutional details

marginal road use decisions than pricing access or fuel.

[5]My analysis can be thought of as a "best-case" bound on the effects of debris removal: even if the strategic issues in debris removal identified by Klima et al. (2016) could be overcome, open access makes the cooperative removal plan continue to be inefficient.

of orbit use and present the basic definitions and modeling framework. In section 3.2 I analyze orbit management policies and the use of debris removal technologies, and present my key results: Proposition 15, that stock controls achieve greater expected social welfare than flow controls, and Proposition 16, that satellite owners must pay for debris removal if the technology is to reduce equilibrium collision risk. I show proofs of these and a few other economically important results in the main text (the rest are in the Appendix, section B). Finally, I conclude in section 3.3 with discussion of the results and thoughts on the future of commercial orbit use.

3.1 Essentials of Orbit Use Policy and Technology

3.1.1 Defining "space traffic control"

One of the central challenges of space traffic control is how to define "space traffic control". Nicholas Johnson, a scientist at NASA, has proposed an aim of space traffic control: "...the goal of space traffic management is to minimize the potential for (radio frequency) or physical interference at any time" Johnson (2004). The radio frequency interference problem is relatively tractable and being handled by existing institutions (Jones et al., 2010). The physical interference problem, essentially collision avoidance, is more difficult from technical and legal perspectives. In GEO, space traffic control is "position control": since satellites in GEO have very low speeds relative to each other, traffic control is as simple as spacing satellites far enough apart that they are unlikely to collide or cause radio frequency interference. In the current regulatory regime, the International Telecommunications Union assigns frequency blocks and geostationary "slots" to national authorities. These authorities are then free to assign their frequencies and slots to entities within their jurisdiction as they see fit, and are also responsible for enforcing responsible spectrum use. In the United States, this is handled by the FCC.[6]

[6] Readers interested in more detail about the history and institutions of space traffic control are referred to Johnson (2004); Jones et al. (2010). Technical proposals for mass removal are discussed in Klinkrad and Johnson (2009), Weeden (2010) discusses the legal challenges, and Tkatchova (2018) examines the potential for markets in debris removal.

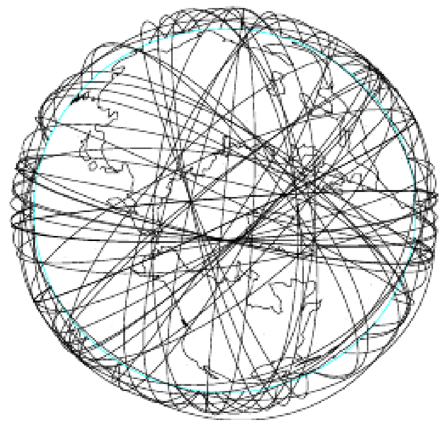

Figure 3.1: *Orbits of 56 cataloged satellites with mean altitudes between 700-710km. Source: Johnson (2004).*

Space traffic control in LEO is harder than in GEO. Satellites in LEO are constantly in motion with respect to each other and have little or no control over their trajectories. Notions like "keep-out zones" are impractical since satellites may only occasionally or accidentally pass through them, and concepts like "rules of the road" raise the question of how a road is to be defined in LEO. Figure 3.1 shows the orbits of 56 cataloged satellites with mean altitudes of 700-710 kilometers, and makes the inaptness of road, sea, and air analogies clear. The growth in LEO use has motivated calls for broader notions of space traffic control which encompass non-GEO regimes. There are currently no international regulatory agencies which coordinate launches and satellite placements to manage debris growth and collision risk; the extent of management policies currently is a patchwork of national regulations and non-binding international guidelines. Table 3.1 shows the breakdown of currently-operational satellites by location of launch site to emphasize the international dimension of the problem. Figure 3.2 shows the growth in orbit use from active satellites and debris, as well as the increase in competition to provide commercial launch services.

For this chapter, I define space traffic control as policies or technologies intended to manage the probability of collisions between active satellites and other bodies. This definition encompasses satellite path as well as debris growth management. Any space traffic control policy, including command-and-control regulations, can be characterized as a price or quantity control, such as a tax or a quota. If the effect

Table 3.1: Currently-operational satellites by origin, orbit class, and orbit type as of April 30, 2018

Breakdown of operating satellites					**Total**
by Country of origin	United States: 859	Russia: 146	China: 250	Other: 631	1,886
by Orbit Class	LEO: 1,186	MEO: 112	Elliptical: 40	GEO: 548	1,886
Breakdown of US satellites by Owner Type	Civil: 20	Commercial: 495	Government: 178	Military: 166	859

Source: Union of Concerned Scientists (2018).

of a policy is to raise the cost or limit the availability of satellite launch, I label it as a "flow" control. If the effect is to raise the cost of operating a satellite or constrain the allowed number of satellites in orbit, I label it as a "stock" control. The existing patchwork of policies includes both flow controls intended to manage launch capacity and prevent launches from interfering with air traffic, and stock controls intended to manage spectrum congestion. While most existing literature on space traffic control focuses on controlling the trajectories of objects in orbit, I focus on controlling the number of objects in orbit. Brief consideration will show that the former implies the latter. I treat debris removal separately because the technology is not yet commercially available, so analysis of a world without debris removal is more immediately relevant to policy design.

3.1.2 A simple model of orbital mechanics with aggregate uncertainty

Following the approach in Chapter 2, I consider the evolution of orbital stocks in an arbitrary spherical shell around the Earth, referred to as the "shell of interest". I consider two types of fictitious agents: a social planner who launches and owns all satellites in orbit to motivate optimal satellite launch and debris removal plans, and a global regulator who manages all satellites launched or in orbit to motivate policy choice.

The notation follows that of the deterministic (no uncertainty) model in Chapter 2, with one change:

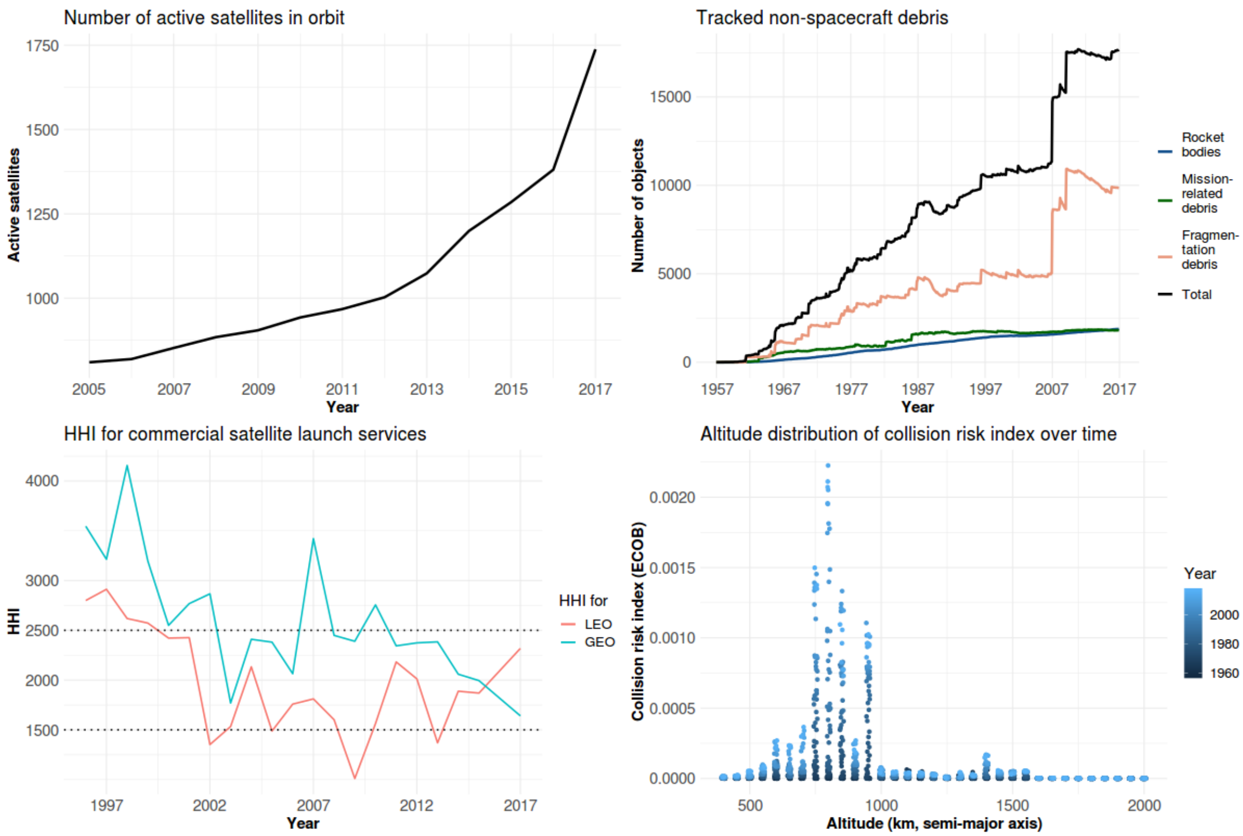

Figure 3.2: *Trends in orbit use.*

Upper left panel: Number of active satellites in orbit per year since 2005.

Upper right panel: Monthly tracked non-spacecraft debris. These do not include derelict satellites which were not deorbited.

Lower left panel: Herfindahl-Hirschman Index for commercial launch services to low-Earth orbit and geostationary orbit.

Lower right panel: Evolution over time of the spatial distribution of ECOB collision risk index in low-Earth orbit. The large spike between 500-1000km is driven by a combination of commercial activity and China's 2007 anti-satellite missile test.

Sources: Union of Concerned Scientists (2018), NASA Orbital Debris Program Office (2017), and Letizia et al. (2018).

collisions are now represented by the random variable ℓ_t, the proportion of satellites which will be lost in collisions at the end of period t (the collision rate). $G(S_t, D_t, \ell_t)$ is the number of new debris fragments generated due to all collisions between satellites and debris. ℓ_t has similar properties as $L(S_t, D_t)$: it is nonnegative and bounded below by 0 and above by 1; no satellites can be destroyed when there are none in orbit ($S_t = 0 \implies \ell_t \equiv 0$); as $S_t \to \infty$ or $D_t \to \infty$, $\ell_t \to 1$ due to physical crowding (unless there are no satellites in orbit).

The number of active satellites in orbit in t is the number of launches in the previous period plus the number of satellites which survived the previous period. The amount of debris in orbit is the amount from the previous period which did not decay, plus the number of new fragments created in collisions, plus the amount of debris in the shell created by new launches. Formally,

$$S_{t+1} = S_t(1 - \ell_t) + X_t \tag{3.1}$$

$$D_{t+1} = D_t(1 - \delta) + G(S_t, D_t, \ell_t) + mX_t. \tag{3.2}$$

As in the deterministic case, I assume that the number of new fragments is nonnegative, increasing in each argument, and zero when there are no objects in orbit ($G(0,0,0) = 0$).

The most important source of uncertainty in orbit management is uncertainty over the proportion of satellites lost to collisions in a given period, ℓ_t.[7] However, the growth in debris objects is also uncertain, so why is uncertainty in $G()$ not treated similarly? The answer is uncertainty in orbit is of economic interest only insofar as it affects active satellites. Uncertainty in the position and interactions between objects in orbit is not economically relevant to orbit management. To see this, consider a counterfactual world where active satellites could not be affected by debris. In this case the uncertainty in debris growth would be

[7] Note that there are also sources of economic uncertainty in orbit use. The future trajectory of launch costs, demand for orbits from other industries, the decisions made by other operators to harden their satellites, and many other factors in orbit use are uncertain. I focus on physical uncertainty because of its first-order relevance to satellite survival and the usability of the orbital environment. Including sources of economic uncertainty would likely change the result in Proposition 11 to favor price or quantity instruments according to the relative slopes of marginal benefit and marginal cost curves (Weitzman, 1974).

irrelevant for orbit management, because the regulator's interest in orbit management is in controlling the value generated by active satellites. Thus, uncertainty in this model is represented by ℓ_t not because debris growth is known perfectly, but because uncertainty over orbital stocks only matters to the extent to which it affects active satellites. Definition 3 describes the revelation of ℓ_t.

Definition 3. *(Symmetric physical uncertainty) The collision rate in period t is revealed to all agents after any debris removal decisions are made but before any launch decisions are made for the period.*

"Symmetric physical uncertainty" means that launching firms and the regulator all know how many satellites will be lost in period t before acting, but not which satellites. On the other hand, satellite owners engaging in debris removal actions don't know how many satellites will be lost until after their removal action. The timing reflects three features of orbit use: (1) while conjunction alerts may be issued to affected operators up to a few days before an anticipated collision, longer-term forecasts of the collision environment are inherently probabilistic; (2) satellite owners who wish to remove debris will attempt to do so before the collisions are unavoidable; and (3) firms choosing whether or not to launch satellites can anticipate satellite owners' removal actions, in part because conjunction alerts are issued publicly to all satellite operators near the affected region so as to better coordinate avoidance maneuvers.

The symmetry between firms and the regulator is practically plausible: firms and regulatory agencies all have access to the same types of information about the position of orbital bodies[8], and can run similar calculations to predict the motion of orbital bodies from given position data. The US Department of Defense makes orbital object data fine enough to perform high-fidelity conjunction analysis on specific satellites available for nominal fees, while aggregate patterns can be modeled using data the Department of Defense makes publicly available. The European Space Agency makes similar data publicly available, albeit at

[8]State actors, particularly national security agencies, may have different information than other agents, breaking the symmetry. As long as the regulator(s) are equally ignorant of this information as firms, the symmetry holds. In some cases the regulator may actually have an informational advantage over firms. This suggests a question not pursued here: what are the limits to a regulator's ability to manage orbital congestion without revealing state secrets?

lower fidelity for academic and hobbyist use. Most of these analyses are probabilistic in nature. Since satellites in the model are all identical, the identity of the satellites lost doesn't matter, and the probability any specific satellite is lost is the same as the aggregate rate.

I assume that ℓ_t has a conditional density $\phi(\ell_t|S_t, D_t)$. With physical uncertainty, only the density of ℓ_t is determined by S_t and D_t, so I explicitly include the draw of ℓ_t as an argument of $G()$. The expected value at the end of period t of a function $f(\ell_{t+1})$ is

$$E_t[f(\ell_{t+1})] = \int_0^1 f(\ell_{t+1})\phi(\ell_{t+1}|S_{t+1}, D_{t+1})d\ell_{t+1}. \tag{3.3}$$

$E_t[\ell_{t+1}]$ is the stochastic analog to $L(S_{t+1}, D_{t+1})$ during period t. This motivates an assumption that the expected collision rate be "increasing" in the number of satellites and amount of debris, in the sense that an increase in either satellites or debris results in a new distribution which first-order stochastically dominates the old one. Assumption 3 states this precisely.

Assumption 3. *(The collision rate is increasing in satellites and debris) An increase in either satellites or debris results in a new distribution which first-order stochastically dominates the old one, that is,*

$$\int_0^k \ell\ \phi(\ell|S+\varepsilon_S, D+\varepsilon_D)d\ell \geq \int_0^k \ell\ \phi(\ell|S, D)d\ell \quad \forall \varepsilon_S, \varepsilon_D \geq 0,\ \forall k \in (0,1),$$

with strict inequality for some $\varepsilon_S > 0$ *or* $\varepsilon_D > 0$.

To reduce notational burden, I suppress the conditioning variables where it is clear from context, though I sometimes make them explicit in proofs. For brevity, I refer to $E_t[\ell_{t+1}]$ as the "collision risk".

3.1.2.1 Functional forms for the collision risk and number of new fragments

For simulations in this chapter, I use functional forms for the collision risk and new fragment formation based on engineering model in Bradley and Wein (2009):

$$E[\ell|S,D] = \min\{\alpha_{SS}S^2 + \alpha_{SD}SD, 1\}, \tag{3.4}$$

$$G(S,D,\ell) = \begin{cases} \beta_{SS}\left(\frac{S}{S+D}\right)\ell S + \beta_{SD}\left(\frac{D}{S+D}\right)\ell S + \beta_{DD}\alpha_{DD}D^2 \text{ if } S+D>0 \\ 0 \text{ if } S+D=0 \end{cases} \tag{3.5}$$

which satisfy all the properties described above. α_{SS}, α_{SD}, and α_{DD} are positive constants which can be derived from an ideal gas model and descriptions of the shapes and sizes of the subscripted object types. They are often referred to as "intrinsic collision probabilities" in engineering studies. β_{SS}, β_{SD}, and β_{DD} are positive constants describing the mean "effective" (that is, adjusted for size and time spent in the shell of interest) number of fragments created in collisions between the subscripted object types. These can be calculated from descriptions of the material compositions of the objects colliding, their relative velocities, and masses. They are often referred to as "fragmentation parameters" in engineering studies. These forms are used to generate figures and simulations, but not for analytical results.

Economically, the expected collision risk can be thought of as a matching function which matches active satellites to debris and other active satellites. The form in equation 3.4 implies that matching between active satellites and debris or other active satellites exhibits "thick market effects": one more active satellite or unit of debris increases the ease with which all active satellites are matched with other orbital bodies. On the other hand, the form in equation 2.5 in Chapter 2 implies that the matching exhibits "thin market effects". I explore the implications of the two different types of market effect assumptions on debris removal plans in section D.8 of the Appendix.

3.1.2.2 Kessler Syndrome

Kessler Syndrome is a central concern in orbit use management. If open access can prevent Kessler Syndrome, regulating orbit use is not as important from an environmental perspective. Even though orbit use will be inefficient it will not cause irreversible environmental damage. On the other hand, if open access can cause Kessler Syndrome, orbit use management is more urgent.

In this section I formally define Kessler Syndrome and establish some properties of the debris threshold beyond which it occurs. Open access debris levels are increasing in the excess return on a satellite while the Kessler threshold is constant, implying that sustained increases in the return on a satellite can cause Kessler Syndrome under open access. Though the Kessler threshold is defined purely in terms of the system's physics, the occurrence of Kessler Syndrome depends critically on the economics of orbit use.

Assumption 4. *(Debris growth) The growth in new fragments due to debris is larger than the decay rate for all levels of the debris stock greater than some level $\bar{D} > 0$,*

$$\bar{D} : G_D(0, D, \ell) > \delta \ \forall D > \bar{D} \ \ \forall \ell.$$

Due to assumption 4 and $G(S, D, \ell)$ being increasing in all arguments, there is a unique threshold $D^\kappa \geq \bar{D}$ above which Kessler Syndrome occurs. Past this threshold, the number of new fragments created by collisions between debris exceeds the amount which decays in a single period. For regimes where this condition doesn't hold at any level of debris, Kessler Syndrome is impossible. Such regimes are likely to be at extremely low altitudes, possibly sub-orbital. For all of the simulations shown in this chapter, Kessler Syndrome is possible.

Definition 4. *(Kessler Syndrome) The Kessler region is the set of debris levels for which cessation of launch activity and immediate deorbit of all active satellites cannot prevent continued debris growth, that is,*

$$D^\kappa : G(0, D^\kappa, \ell) > \delta D^\kappa \ \ \forall \ell.$$

Kessler Syndrome has occurred when the debris stock enters the Kessler region.

Without active debris removal technologies, Kessler Syndrome is an absorbing state. Once Kessler Syndrome occurs in the model, the debris stock grows without bound. In reality, the fragments would eventually pulverize each other into small fragments and either find stable orbits or decay back to the Earth, but this process could take centuries or millennia (Kessler and Cour-Palais (1978)).

3.1.3 The economics of open access and optimal orbit use without debris removal

With the environment and physical considerations described, I turn to the optimization problems facing orbit-using agents. I compare the equilibrium resulting from individual firms' launch decisions under open access to the outcome of the optimal launch plan followed by the fleet planner. The main point of this section is the comparison between equations 3.11 and 3.15, presented in figure 3.3.

An infinitely-lived firm which owns a satellite collects a return of π every period that the satellite survives and applies a discount factor of $\beta = \frac{1}{1+r}$ to future revenues. A fraction ℓ_t of the orbiting satellites are destroyed in collisions every period. The realization of ℓ_t is revealed to all agents just after satellites launched in $t-1$ have reached orbit, but before any launch decisions in t are made. Thus, ℓ_t is known before launch decisions in t, but ℓ_{t+1} is unknown. Expectations in period t are therefore taken over realizations of ℓ_{t+1}, as shown in equation 3.3. Since the satellites are identical, the probability that an individual satellite survives the period is $(1-\ell_t)$, and the probability it is destroyed is ℓ_t. If the satellite is destroyed, the firm will once again face the decision of launching or not. The value of a satellite in period t is

$$Q(S_t,D_t,\ell_t,X_t) = \pi + \beta[(1-\ell_t)E_t[Q(S_{t+1},D_{t+1},\ell_{t+1},X_{t+1})] + \ell_t E_t[V_i(S_{t+1},D_{t+1},\ell_{t+1},X_{t+1})]] \tag{3.6}$$

A firm which does not own a satellite in period t faces the decision to pay a fixed cost F and launch a satellite which will reach orbit and start generating revenues in period $t+1$, or to wait and decide again whether or not to launch in period $t+1$. Once the firm decides to launch, it must wait one period before the satellite will reach orbit and begin producing returns. Assuming potential launchers are risk-neutral profit maximizers, the value of potential launcher i at period t is

$$\begin{aligned} V_i(S_t,D_t,\ell_t,X_t) = \max_{x_{it}\in\{0,1\}} &\{(1-x_{it})\beta E_t[V_i(S_{t+1},D_{t+1},\ell_{t+1},X_{t+1})] \\ &+ x_{it}[\beta E_t[Q(S_{t+1},D_{t+1},\ell_{t+1},X_{t+1})] - F]\} \end{aligned} \tag{3.7}$$

$$S_{t+1} = S_t(1-\ell_t) + X_t$$

$$D_{t+1} = D_t(1-\delta) + G(S_t,D_t,\ell_t) + mX_t$$

$$\ell_{t+1} \sim \phi(\ell|S_{t+1},D_{t+1})$$

Under open access, firms launch until profits are zero:

$$X_t > 0: \; V_i(S_t, D_t, \ell_t, X_t) = 0 \tag{3.8}$$

$$\implies \beta E_t[Q(S_{t+1}, D_{t+1}, \ell_{t+1}, X_{t+1})] = F. \tag{3.9}$$

The value of a satellite is then

$$Q(S_t, D_t, \ell_t, X_t) = \pi + (1 - \ell_t)F, \tag{3.10}$$

and the equilibrium collision risk is

$$E_t[\ell_{t+1}] = r_s - r, \tag{3.11}$$

where $r_s = \frac{\pi}{F}$ is the one-period rate of return on a single satellite.[9]

With the open access equilibrium launch rate characterized, I move on to characterize the fleet planner's optimal launch rate. Recall that the planner owns all satellites in orbit, and can launch as many as they would like each period. The fleet planner maximizes the expected net present value of the entire fleet. Their problem is

$$W(S_t, D_t, \ell_t) = \max_{X_t \geq 0} \{\pi S_t - F X_t + \beta E_t[W(S_{t+1}, D_{t+1}, \ell_{t+1})]\} \tag{3.12}$$

$$\text{s.t.} \;\; S_{t+1} = S_t(1 - \ell_t) + X_t \tag{3.13}$$

$$D_{t+1} = D_t(1 - \delta) + G(S_t, D_t, \ell_t) + m X_t. \tag{3.14}$$

The planner launches so that the loss rate is equated to the rate of excess return net of the marginal external cost (ξ_{t+1}), that is,

$$E_t[\ell_{t+1}] = r_s - r - \frac{E_t[\xi(S_{t+1}, D_{t+1})]}{F}. \tag{3.15}$$

where $E_t[\xi(S_{t+1}, D_{t+1})]$ is the marginal external cost of a satellite launch. For the results in this chapter, it suffices to assume that the marginal external cost is weakly positive for all S_{t+1} and D_{t+1} along the optimal path, and strictly positive for some values of S_{t+1} and D_{t+1}. Figure 3.3 illustrates the differences between open access and optimal policies in the deterministic setting.

[9] The open access equilibrium is Markov perfect: conditional on the state of the game, no launching firm can profit by deviating to "wait" and no waiting firm can profit by deviating to "launch".

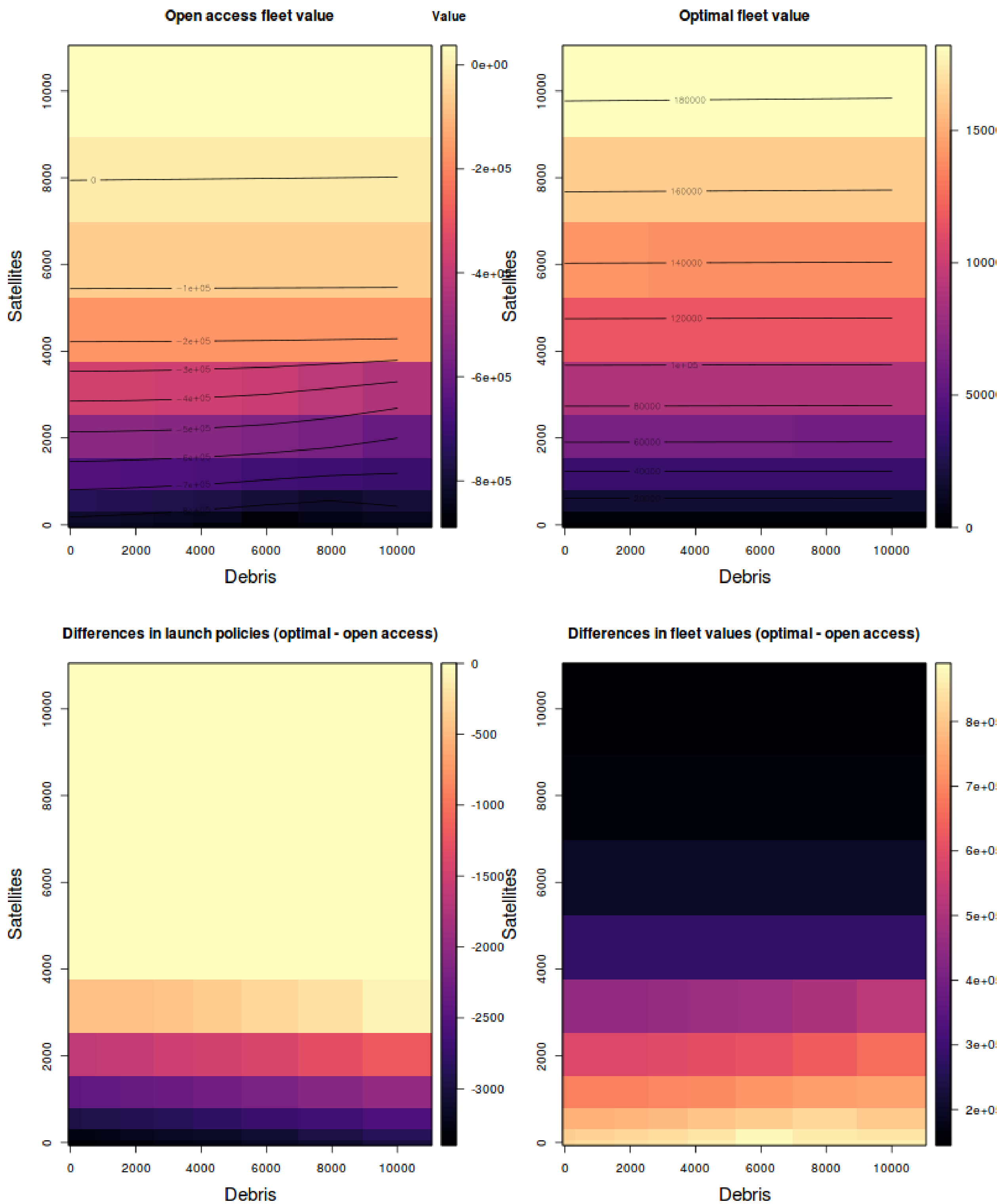

Figure 3.3: *An example of the gap between open access and optimal launch policies, with the corresponding gap in fleet values.*
The planner launches fewer satellites in every state than open access firms would. The value gap is maximized when (a) there is no debris and (b) the planner would stop launching satellites but open access firms do not.

3.1.4 Debris removal technologies

Orbital debris is dangerous to active satellites in part because debris objects cannot be maneuvered and often do not transmit their location to ground stations. Active satellites, on the other hand, tend to do both, making collision avoidance maneuvering easier. Active satellites also tend to have some guidance and control systems which allow them to be deorbited remotely, if necessary. Debris objects tend not to have such systems, because they are fragments of a satellite, non-responsive to ground operator commands, or out of fuel and incapable of further maneuvers. Active debris removal technologies are those which can interact with debris objects and deorbit them. They are contrasted with passive removal, which involves measures like setting a satellite on a path which will result in its deorbit in a specified timeframe.

Active debris removal technologies are being developed, but have not yet been commercially deployed. Some of these technologies involve specialized removal satellites which use the Earth's magnetic field for propulsion and deploy nets, harpoons, or tethers (for example, Pearson et al. (2010)) to either deorbit debris or recycle the materials for in-space manufacturing. Ground-based lasers are another candidate technology to deorbit debris.

I assume no new satellites are required to implement removal, which can be interpreted in two ways: that the removal technology is ground-based; or that the satellites required are already in orbit and can never be destroyed or lost. Including the requirement that new satellites be used for removal complicates the model in interesting and relevant ways that are beyond my scope here. I also assume that only satellite owners can purchase debris removal.

With the ability to remove debris from orbit, satellite owners can remove clearly-dangerous pieces of debris before they impact their satellites. The remaining collisions will be caused by errors in debris risk assessments, satellite trajectory forecasts, and collisions which were deemed too costly to avoid. To reflect this in the model, I adjust the timing of when ℓ_t is revealed when debris removal technologies are present. Satellite owners purchase R_t total units of removal before ℓ_t is revealed, with the aim of changing the distribution of ℓ_t until the marginal private benefit of removal equals the marginal private cost. After removal has been purchased, ℓ_t is drawn from a distribution conditioned on S_t and $D_t - R_t$ (instead of just S_t and D_t) and revealed to all satellite owners and prospective launchers. The launchers then decide whether or not to launch.[10]

With debris removal before collisions, the laws of motion and distribution of the collision rate become

$$S_{t+1} = S_t(1-\ell_t) + X_t \tag{3.16}$$

$$D_{t+1} = (D_t - R_t)(1-\delta) + G(S_t, D_t - R_t, \ell_t) + mX_t \tag{3.17}$$

$$\ell_t \sim \phi(\ell_t | S_t, D_t - R_t). \tag{3.18}$$

Expectations before removal in t are indicated by $\tilde{E}_t[\cdot]$ and treat ℓ_t as a random variable, while expectations after removal in t are indicated by $E_t[\cdot]$ and treat ℓ_t as known. The expected collision risk before removal is effected is

$$\tilde{E}_t[\ell_t] = \int_0^1 \ell_t \phi(\ell_t | S_t, D_t - R_t) d\ell_t. \tag{3.19}$$

Potential launchers in t have the same expectations as before: they are aware of ℓ_t, and treat ℓ_{t+1} as uncertain. Formally,

$$E_t[\ell_{t+1}] = \int_0^1 \ell_{t+1} \phi(\ell_{t+1} | S_{t+1}, D_{t+1} - R_{t+1}) d\ell_{t+1} = \tilde{E}_{t+1}[\ell_{t+1}]. \tag{3.20}$$

Though $E_t[\ell_{t+1}] = \tilde{E}_{t+1}[\ell_{t+1}]$, I use separate notation so that the subscript on the expectation operator indicates the period in which the agent forms the expectation, and the tilde above the expectation operator indicates whether the expectation is formed before or after is drawn and revealed. $E_t[\ell_{t+1}]$ is an expectation formed in t after ℓ_t is drawn and revealed, $\tilde{E}_t[\ell_t]$ is an expectation formed in t before ℓ_t is drawn and revealed.

[10] In reality, the timing of satellite launches and debris removals will not be this clearly separated. However, potential launchers will be able to anticipate satellite owners' debris removal demands, and where possible structure their launches to take advantage of these efforts.

Table 3.2: Types of orbit management policies

	Quantity control	**Price control**
Flow control	Launch permits, capacity constraints,...	Launch fees, monopoly pricing,...
Stock control	Orbit use leases, path rights,...	Orbit use taxes, collision liability,...

3.2 Space Traffic Control

3.2.1 Policy without active debris removal

Space traffic control policies restrict either the number of satellites launched to or the number of satellites in an orbit in a given window of time. As described earlier, I refer to policies restricting the number of launches in a given period as *flow controls*, and policies which restrict the number of satellites in orbit in a given period as *stock controls*. Stock controls entail an explicit or implicit payment made every period that the satellite is in orbit. The payment gives the satellite owner the right to keep their satellite in orbit that period. Flow controls entail a payment made once when the satellite is launched. The payment gives the satellite launcher the right to launch in that period. Table 3.2 gives some examples of each type of control policy. Both types of controls are currently in place around the world - the FAA's launch permit system is a flow control for launches from the United States, while the ITU's minimum spacing requirements for satellites in GEO are a stock control for GEO use.[11] Existing controls tend to be implemented as quantities, as in the two examples given, but could also be implemented as prices, for example, a launch or satellite tax.

Quantity restrictions imply price restrictions and vice versa. In many settings, either mode can generate equivalent social welfare. Weitzman (1974) establishes that the equivalence can break down in the presence of regulatory uncertainty over the firm's marginal cost of production. Whether a price or quantity instrument should be preferred in such settings depends on the relative slopes of the marginal benefit and

[11]These requirements tend to focus on launch capacity and spectrum interference, rather than the risk of collisions and debris growth. The framework developed here applies to orbit use management policies regardless of their intent.

marginal cost curves. This is not the case for orbits, where the main source of uncertainty comes from the motion of physical objects which are in principle observable by all actors. Unlike the regulatory problems considered in Weitzman (1974) and Newell and Pizer (2003), the firm has no additional information about the motion of orbital bodies for the regulator to harness through instrument design.

The distinction between stock and flow controls is relevant to a broad class of economic management problems. To encourage renewable energy generation, a regulator may weigh investment (stock) vs production (flow) tax credits (Aldy et al., 2018). To manage public infrastructure a regulator may weigh investment in damage abatement (flow) vs quality restoration (stock) (Keohane et al., 2007).[12]

In the absence of informational or administrative constraints on the regulator, the preferred instrument is that which most directly targets the externality-generating activity (Sandmo, 1978). In the renewable energy case, production tax credits can encourage renewable energy generation more effectively than investment tax credits.[13] In orbit, stock controls dominate flow controls because the collision risk externality is driven by the number of objects in orbit rather than the number of objects launched in a period. Because stock controls directly target the incentive to own a satellite while flow controls target the incentive to launch a satellite, they present satellite owners and launchers with different incentives. These differing incentives drive a wedge between their abilities to manage orbital congestion.

Stock and flow controls can often be made equivalent in the sense that one can be capitalized or annuitized to the same present value cost as the other. However, they have different effects on the incentive

[12]Keohane et al. (2007) consider the use of stock and flow controls to manage the quality of a resource, but their use of "stock control" is slightly different due to the setting considered. In their setting, "stock controls" refer to policies which restore the stock of a deteriorating resource. Here, the term refers to limiting the stock of a commodity which deteriorates the resource. Keohane et al. (2007)'s use of "flow controls" is closer to the use of the term here: they consider abating the flow of pollutants into the environment, and I consider controlling the flow of satellites into orbit.

[13]Provided capacity is not a binding constraint, production effort is costly, and the production function is not characterized by decreasing returns to scale, as described in Aldy et al. (2018) and Parish and McLaren (1982).

to launch or own a satellite. Imposing a fee at launch increases the cost of entering the orbital commons, penalizing entrants while increasing the rents accruing to incumbents in orbit. Imposing a recurring fee while the satellite is in orbit reduces the rents of satellite ownership without restricting entry, treating entrants and incumbents equally. These differing incentives can lead to welfare differences between stock and flow modes of orbit control. To show how stock and flow controls affect the decision to launch a satellite, consider two cases with price-based controls. In the first, a stock control is levied on satellite owners. In the second, a flow control is levied on satellite launchers. I assume the regulator can commit to future policies, so that $t+1$ values are known to firms with certainty.[14]

The decision to launch under a stock control: Let the price that a satellite owner pays in t be p^s_t. Firms deciding whether to launch or not in period t will account for their anticipated regulatory burden as they drive the profits of launching a satellite down to zero. The marginal benefit from owning a satellite in $t+1$ must therefore equal not only the opportunity cost of the launch, but also the direct regulatory cost of the stock control. Formally,

$$X_t : \beta E_t[Q(S_{t+1}, D_{t+1}, \ell_{t+1}, X_{t+1})] = F \tag{3.21}$$

$$E_t[Q(S_{t+1}, D_{t+1}, \ell_{t+1}, X_{t+1})] = \pi - p^s_{t+1} + (1 - E_t[\ell_{t+1}])F \tag{3.22}$$

$$\implies \pi = rF + E_t[\ell_{t+1}]F + p^s_{t+1}. \tag{3.23}$$

The decision to launch under a flow control: Let the price that a satellite launcher pays in t be p^f_t. Firms deciding whether to launch or not in period t will account for the regulatory burden of launching.

[14]Stock and flow controls both require forward guidance, since announced or anticipated $t+1$ values affect the launch rate in t. But whereas flow controls require forward guidance regarding the entire time path of control values, stock controls only require forward guidance about the next-period control value. Even without commitment, the regulator faces no incentive to deviate from a previously-announced stock control rule. This may not be the case for flow controls. Because anticipated changes in the flow control rule can cause launching firms to "bunch" and attempt to launch either just before a price increase or just after a price decrease, the regulator has an incentive to make flow control policy changes a surprise. Such surprises would change private expectations of the control policy path. In environmental economics, Newell et al. (2005) consider the tradeoffs between commitment and discretion in stabilizing quantity-policy prices. This tradeoff is analyzed in more depth in the monetary policy literature; Svensson (2003) provides a comprehensive discussion.

Since they know that future launchers will face a similar regulatory burden, they will consider how the future flow control price will affect the open access value of a satellite. The marginal benefit of owning a satellite in $t+1$ must therefore equal the opportunity cost of the launch, which includes the flow control price they pay and the forgone interest. However, the marginal benefit of owning a satellite in $t+1$ now includes not only the direct revenues the satellite generates but also the additional expected value from the flow control levied on $t+1$ launchers. Formally,

$$X_t : \beta E_t[Q(S_{t+1}, D_{t+1}, \ell_{t+1}, X_{t+1})] = F + p_t^f \tag{3.24}$$

$$E_t[Q(S_{t+1}, D_{t+1}, \ell_{t+1}, X_{t+1})] = \pi + (1 - E_t[\ell_{t+1}])F + (1 - E_t[\ell_{t+1}])p_{t+1}^f \tag{3.25}$$

$$\implies \pi + (1 - E_t[\ell_{t+1}])p_{t+1}^f = rF + E_t[\ell_{t+1}]F + (1+r)p_t^f. \tag{3.26}$$

Figure 3.4 illustrates equations 3.23 and 3.26. Although imposing either type of control can reduce the equilibrium number of launches, flow controls raise the private marginal benefit of launching along with the private marginal cost. Stock controls, on the other hand, affect only the marginal cost of launching. This is the core intuition for why stock controls are preferable to flow controls for managing orbital congestion.

Leakage issues and legal hurdles: Both types of controls face leakage issues. Flow controls implemented by regional launch providers may suffer "launch leakage", while stock controls implemented by regional regulatory agencies may suffer "mission control leakage". Similar leakage issues have been studied extensively in the environmental and public economics literatures, for example Fowlie (2009); Fischer and Fox (2012); Böhringer et al. (2017). Though these issues are relevant to effective policy implementation, analyzing them is beyond the scope of this chapter. Within the US context, leakage is often limited by restrictions on dual-use technologies, e.g. ITAR. While firms may attempt to circumvent these restrictions by incorporating elsewhere, those with US citizens as shareholders are vulnerable to efforts to "pierce the corporate veil", making the firm's attempts to circumvent dual-use restrictions the responsibility of its shareholders.

The legal hurdles to implementing stock controls may also be higher than those for flow controls,

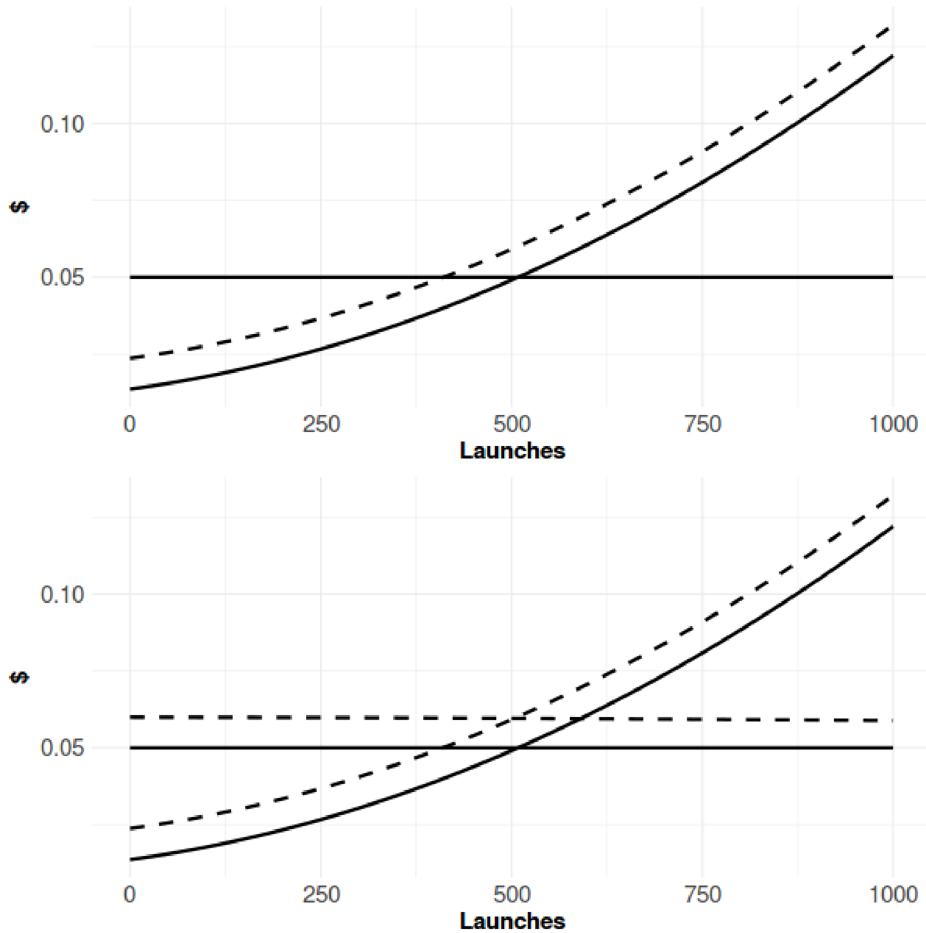

Figure 3.4: *Private marginal benefits and costs of launching a satellite under stock (left) and flow (right) controls.*

Left panel: The horizontal solid line is the marginal benefit of launching a satellite, while the upward-sloping solid line is the marginal cost. The dashed line indicates the effect of imposing a stock control: the marginal cost is increased, lowering the equilibrium number of satellites launched.

Right panel: The horizontal solid line is the marginal benefit of launching a satellite, while the upward-sloping solid line is the marginal cost. The dashed lines indicate the effects of imposing a flow control: the period t control raises the marginal cost of launching in t, but the entry restriction of the period $t+1$ control raises the marginal benefit of launching in t. Corollary 5 establishes that a constant price will increase the marginal cost by more than it increases the marginal benefit, but the net effect size may be very small.

since they require a legal framework in which the right to exclude agents from an orbit can be held and enforced. Such a framework would have to be globally agreed-upon and potentially self-enforcing. I do not consider the prospects of such an agreement in this chapter, although similar issues have been studied extensively in economics generally and environmental economics specifically, for example Telser (1980); Barrett (2005, 2013).

3.2.1.1 Using stock and flow space traffic control policies

In this section, I formally describe some properties of stock and flow controls and how they should be used to manage space traffic. The first property is price-quantity equivalence: under symmetric physical uncertainty, a stock or flow control can be implemented as a price or quantity and achieve equivalent expected social welfare. This allows me to consider price or quantity implementations interchangeably. I then show how stock and flow controls should be used to limit launches, and consider the implications of these details for optimal control values. I follow this by showing how the launch rate responds to the initiation of a stock or flow control, and how a regulator could use those controls to induce firms to deorbit already-orbiting satellites and stop launching new ones. These properties are used in the following section to establish that regulating orbit use through stock controls achieves higher expected social welfare than using flow controls.

Proposition 11. *(Price-quantity equivalence) Under symmetric physical uncertainty, price and quantity implementations of stock controls are equivalent, as are price and quantity implementations of flow controls.*

Proof. I show the result for stock controls first, and then for flow controls.

Stock controls: I refer to price-based stock controls as satellite taxes, and quantity-based stock controls as satellite permit quotas. Let the launch rate under a satellite tax be $\tilde{X}_t$, and the permit price under a permit quota be $\tilde{p}_{t+1}$.

Under a satellite tax, the number of satellites launched will be

$$\tilde{X}_t : \pi = rF + E_t[\ell_{t+1}]F + p_{t+1} \tag{3.27}$$

Under a binding satellite permit quota, firms will purchase permits and launch satellites until the price of a permit is

$$\tilde{p}_{t+1} : \pi = rF + E_t[\ell_{t+1}]F + \tilde{p}_{t+1} \tag{3.28}$$

For a given state vector (S_t, D_t, ℓ_t) and a chosen price p_{t+1}, the monotonicity of $E_t[\ell_{t+1}]$ ensures that equation 3.27 determines a unique value of $\tilde{X}_t$. For the same state vector and $X_t = \tilde{X}_t$, the monotonicity of $E_t[\ell_{t+1}]$ ensures that $\tilde{p}_{t+1} = p_{t+1}$ solves 3.28.

Flow controls: I refer to price-based flow controls as launch taxes, and quantity-based flow controls as launch permit quotas. Let the launch rate in t under a launch tax be $\tilde{X}_t$, and the permit price in $t+1$ under a permit quota be $\tilde{p}_{t+1}$.

Under a launch tax, the number of satellites launched will be

$$\tilde{X}_t : \pi = rF + E_t[\ell_{t+1}]F + (1+r)p_t - (1 - E_t[\ell_{t+1}])p_{t+1} \tag{3.29}$$

Under a binding launch permit quota, firms will purchase permits and launch satellites until the price of a permit is

$$\tilde{p}_{t+1} : \pi = rF + E_t[\ell_{t+1}]F + (1+r)p_t - (1 - E_t[\ell_{t+1}])\tilde{p}_{t+1} \tag{3.30}$$

For a given state vector (S_t, D_t, ℓ_t) and a chosen price p_{t+1}, the monotonicity of $E_t[\ell_{t+1}]$ ensures that equation 3.29 determines a unique value of $\tilde{X}_t$. For the same state vector and $X_t = \tilde{X}_t$, the monotonicity of $E_t[\ell_{t+1}]$ ensures that $\tilde{p}_{t+1} = p_{t+1}$ solves 3.30. $\square$

With access to commitment, a regulator using a flow control sets either the future number of permits or their price (X_{t+1} or p_{t+1}) in order to influence the launch rate today (X_t). Raising p_{t+1} in t raises the marginal benefit of launching a satellite today, but lowers it tomorrow. The use of flow controls requires the regulator to trade off the future launch disincentive of raising p_{t+1} against the current launch incentive it creates. The regulator's true instrument with a flow control is not the price of the control itself, but the *change* in price between periods. Rather than a price mapping to a quantity, here it is a (real) change in price which maps to a quantity and vice versa. The regulator can set any initial flow control price so long as they commit to a path of control prices based on equation 3.26. A similar penalty-rebate structure appears in the mining flow control studied in Briggs (2011), where incentivizing mine owners to mine less in t requires a lower Pigouvian tax in period $t+1$.

Note that stock control prices must be positive to reduce launches in any given period, while flow control prices need not be positive to do the same. Along positive price paths the flow control is an entry restriction while along negative price paths it is an entry subsidy. Current restrictions deter current entry, but future restrictions deter future entry and boost the rents accruing to incumbents, incentivizing current entry. Current subsidies encourage current entry, but future subsidies encourage future entry and reduce the rents accruing to incumbents, incentivizing firms to delay entry. In either case, the regulator is able to use the change in flow control prices to rearrange satellite launches over time.[15]

The need to commit to a flow control path makes terminal conditions economically relevant to their use. If the regulator plans to use the flow control for only a limited duration, after which the orbits will be under open access again, the flow control price path will decrease over time until it is zero in the period where open access is restored.[16] A flow control which attempts to ensure optimality with no planned phase-out will be forced to follow an exploding price path, positive or negative, as the regulator attempts to balance present and future incentives and disincentives without causing launchers to "bunch" suboptimally

[15]Technically, the flow control structure creates a problem if the loss rate drawn in t is large enough that the expected loss rate in $t+1$ is one. If this happens, there is no future launch control price which can satisfy equation 3.26. Lemma 7 in section D.6 of the Appendix shows this formally. If the regulator wishes to control the launch rate in periods where the expected future loss rate is one, they must break their earlier commitment and adjust p_t until the launch rate is where they want it to be. Since I assume the regulator cannot break their commitment, in this case there is simply no time-consistent flow control which can affect the launch rate. This holds for both price and quantity implementations. That said, if potential launchers expect their satellite to be destroyed after one period, they will only launch in the unrealistic edge case where one period of returns from a satellite exceeds the cost of launching. It is more likely that potential launchers would rather not launch if their satellites are expected to survive only one period.

[16]Why might a regulator want to do this? Open access launching tends to overshoot the open access steady state, potentially ending up in the Kessler region. A regulator who wished to prevent this without committing to optimality may therefore impose a flow control until the risk of overshooting is sufficiently reduced.

in any period while the control is active. This property of the price path is formally established in section D.6 of the Appendix. The credibility of such a price path is doubtful, but beyond my scope here.

Limiting launches with stock and flow controls:

While stock controls are straightforward - raise the price to reduce launches - flow controls are subtler. To limit launches in t, the flow control price in $t+1$ should be lowered instead of raised. The intuition for this can be seen in Figure 3.4 and in equation 3.26, where the price of a flow control in $t+1$ enters the launch decision in t as a marginal benefit rather than a marginal cost. This has implications for the design of optimal controls: an optimal stock control equates the $t+1$ control price with the expected marginal external cost in $t+1$, while an optimal flow control makes the expected real difference in t and $t+1$ control prices equal to the negative of the expected $t+1$ marginal external cost.

Lemma 1. *(Launch response to stock and flow controls) The open access launch rate is*

- *decreasing in the future price of a stock control;*
- *decreasing in the current price and increasing in the future price of a flow control.*

Proof. See Appendix section B. □

Corollary 5. *The shift in marginal cost of owning a satellite due to an increase in the flow control price is greater than the prior shift in marginal benefit due to the entry restriction.*

Proof.

$$r > 0 \implies \left|\frac{\partial X_t}{\partial_t p}\right| > \left|\frac{\partial X_t}{\partial p_{t+1}}\right| \tag{3.31}$$

$$\implies 1 + r > 1 - E[\ell_{t+1}], \tag{3.32}$$

which is true because $\ell_{t+1} \in [0,1]$ by definition. □

Committing in t to raising the flow control price in $t+1$ raises the marginal benefit of owning a satellite before $t+1$, when the new price comes into effect and raises the marginal cost of launching a satellite. This increases the number of launches in t and reduces the number in $t+1$. On the other hand, committing in t to lowering the flow control price in $t+1$ reduces the marginal benefit of owning a satellite in $t+1$, when the new price comes into effect and lowers the marginal cause of launching a satellite. This reduces the number of launches in t and increases the number in $t+1$. This "launch bunching" is absent in stock controls.

Optimal control policies: Making a stock control optimal is simple: set the price equal to the expected marginal external cost of another satellite. Letting p_t^s be the value of the stock control in period t,

$$p_{t+1}^s = E_t[\xi(S_{t+1}, D_{t+1})] \tag{3.33}$$

will make launchers behave as the planner would command.

Making a flow control optimal is more complicated. The value of the control in the previous period must be taken into account to balance intertemporal launch incentives. The expected survival rate must be accounted for as well, as it determines the expected rent due to entry restriction the firm will realize. Formally, letting p_t^f be the value of the flow control in period t,

$$p_{t+1}^f = \frac{(1+r)p_t^f - E_t[\xi(S_{t+1}, D_{t+1})]}{1 - E_t[\ell_{t+1}]} \tag{3.34}$$

is required. Perhaps counterintuitively, the expected marginal external cost of another satellite must be *subtracted* from the future value of the stock control. This is because the future flow control price represents a benefit to current launchers, rather than a cost, as seen in Figure 3.4. Figure 3.7 shows examples of optimal stock and flow control policies.

Initiating control: Along an interior equilibrium path, stock and flow controls can both be optimal. This raises the questions of whether the equivalence holds when a control is first put into place, and whether boundaries (periods when a control is implemented from an open access status quo, or when a control is used to shut down all launches) present any challenges to either control type. Proposition 12 shows that the equivalence does not hold at initiation boundaries. Figure 3.5 illustrates Proposition 12.

Proposition 12. *(Smoothness at boundaries) Stock controls can be initiated without letting the launch rate exceed the open access launch rate. Flow controls cannot be initiated without forcing the launch rate to exceed the open access launch rate.*

Proof. See Appendix section B. □

As described in Lemma 1 and Figure 3.4, positive flow control prices first shift the marginal benefit of owning a satellite before the control is implemented upwards, then shift the marginal cost of owning a satellite after the control is implemented upwards. As described in Corollary 5, the increase in marginal cost is necessarily greater than the increase in marginal benefit, so if the flow control price is kept stationary after the increase, there will be fewer launches per period than before. However, more than the equivalent open access number of firms will launch just before the flow control price is raised to capture its rents.

Price and quantity stock controls are equally easy to use to halt all launching. As a price, the control value is simply raised until no firm wants to launch. As a quantity, the control value is simply frozen at whatever number of satellites in orbit us desired. Using flow controls for this purpose is trickier, with quantities being more intuitive than prices. A quantity flow control can be used to halt all launching by setting the number of allowed launches to zero. Though expectations of a launch shutdown may induce firms to launch earlier, the mechanics are described in Lemma 1 and bunching can be mostly avoided with careful attention to the entire path of allowed launch quantities (bunching before the period when control is implemented is unavoidable). Using a price flow control is not as simple as raising the price once, however, since (a) in the period before the price increase firms will want to launch to capture the rents from restricted entry the following period, and (b) if the difference in the flow control price between periods is constant launching may resume. To avoid bunching and maintain shutdown, the regulator must instead *lower* the flow control price in the period when launch shutdown is desired, and then commit to an ever-decreasing sequence of prices which will eventually go to negative infinity (the reasoning is described in Lemma 6). The credibility of such price paths is questionable.

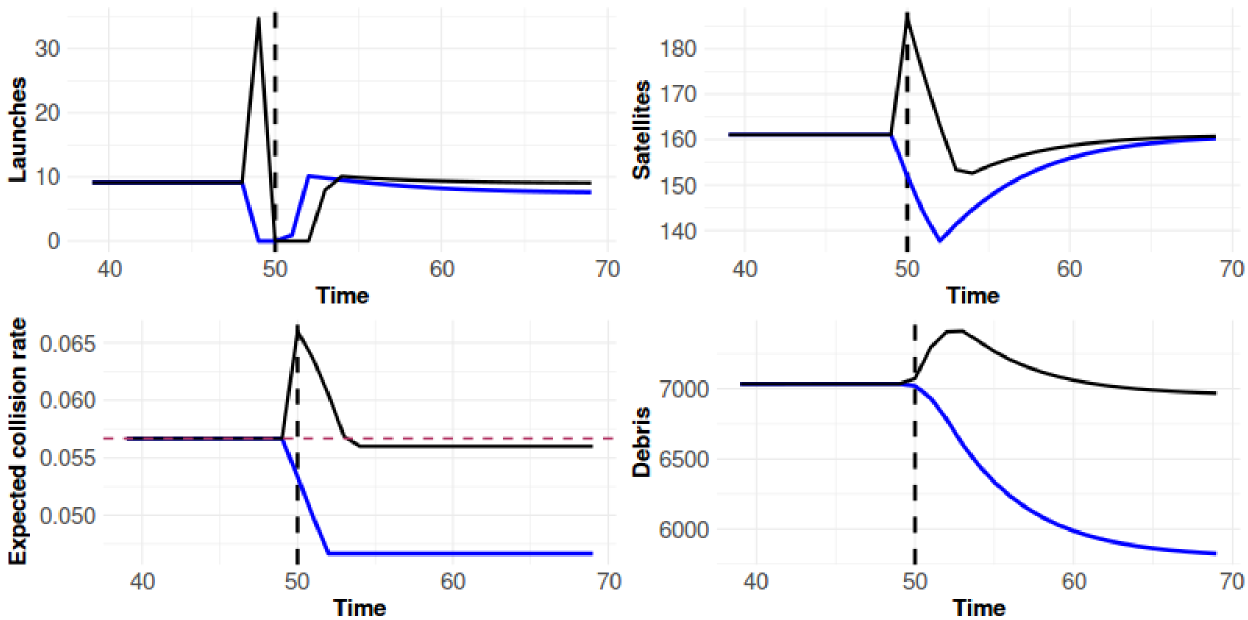

Figure 3.5: *The effects of introducing a generic constant stock (blue line) or flow control (black line).* The purple dashed line shows the equilibrium collision risk under open access. Introducing a stock control smoothly reduces the expected collision risk and debris stock, while introducing a flow control forces both to jump above the open access levels before they are reduced. The assumption of constancy is made for exposition, and is not important to the result.

Inducing satellite owners to deorbit:

Satellite owners often have the option to deorbit their satellite if it becomes too expensive to operate[17]. In this section only, I include the deorbit option for satellite owners to consider whether stock and flow control policies can induce deorbits. The firm's net payoff from deorbiting their satellite is $V^d \lesseqgtr 0$. V^d includes any liquidation revenues (for example, from selling mission control equipment) or costs (for example, costs of damage to people or property during the deorbit). Firms decide whether or not to deorbit after ℓ_t is revealed. A firm which decides to deorbit doesn't claim the revenues from being in orbit that period. Formally,

$$Q(X_t, S_t, D_t, \ell_t) = \max\{\pi + (1-\ell_t)F, V^d\} \tag{3.35}$$

Satellites that are in the process of being deorbited may still collide with each other or be struck by debris. Denoting the number of satellites deorbited as Z_t, the laws of motion with deorbit are

$$S_{t+1} = (S_t - Z_t)(1-\ell_t) + X_t \tag{3.36}$$

$$D_{t+1} = D_t(1-\delta) + G(S_t, D_t, Z_t, \ell_t) + mX_t \tag{3.37}$$

$$\ell_t \sim \phi(\ell_t | S_t, D_t). \tag{3.38}$$

The satellites which are deorbited but still destroyed in collisions ($Z_t \ell_t$) are included in D_{t+1}. Firms choose to deorbit if the payoff from deorbit exceeds the payoff from remaining in orbit:

$$\text{Deorbit if } V^d > \pi + (1-\ell_t)F \tag{3.39}$$

$$\ell_t > 1 + \frac{\pi - V^d}{F}. \tag{3.40}$$

$V^d < \pi$ is a no-arbitrage condition: it ensures that firms can't pump money out of an orbit by repeatedly launching satellites and deorbiting them as soon as they reach orbit.[18] The no-arbitrage condition implies that flow controls can't force firms to deorbit with positive flow control prices, described in Proposition 13.

[17] In the early 2000s, lack of profitability nearly induced the operators of the Iridium constellation to deorbit their satellites. Iridium SSC ultimately went bankrupt, but was able to find a consortium of buyers who kept the constellation in orbit. Modern cubesats are often launched without sufficient guidance and control capabilities to initiate deorbit. Their trajectories are typically planned so that they will naturally deorbit within a few years of their launch.

[18] Since $\ell_t \in [0,1]$, a firm will never deorbit its satellite if it isn't required. Since firms here own only one satellite each, it would

Proposition 13. *(Controlling the rate of deorbit) Stock controls with positive prices can make satellite owners deorbit their satellites and induce net deorbits (more deorbits than launches). Flow controls with positive prices cannot make satellite owners deorbit their satellites or induce net deorbits.*

Proof. See Appendix section B. □

Intuitively, making it costlier for firms to launch new satellites cannot make already-orbiting satellites less valuable. This is why flow controls are unable to induce deorbits, at least with positive prices. Flow controls with negative prices may or may not be able to induce deorbits, depending on parameter values and the number of new entrants induced.

3.2.1.2 Risks and policy choice

In this section, I consider how the choice of stock or flow control mode will affect the equilibrium collision risk and the probability of Kessler Syndrome. I establish that stock controls generate weakly higher expected fleet values than flow controls over arbitrary horizons. Due to the smoothness properties described in Proposition 12, both collision and Kessler risks are increased when a flow control is initiated but not when a stock control is initiated. Fundamentally, the stock of objects in orbit is the source of orbit use externalities, not the flow of objects into orbit. Because of this, it is intuitive that stock controls are better tools to manage space traffic than flow controls. An analogy to road traffic control is instructive here. Road traffic congestion is, in theory, better regulated by congestion pricing levied than by congestion-based road

be economically strange for them to throw away the potential for future profits by deorbiting. When firms own multiple satellites they may decide to deorbit one satellite to preserve others. If the satellites depreciate or technology improves, they may decide to deorbit and replace a satellite. In both cases the no-arbitrage condition $V^d < \pi$ would still apply. Suppose a constellation owner with depreciating satellites will deorbit one of their satellites, j, at time $\bar{t}_j$ in response to the depreciation and collision risks between j and the other satellites in their constellation. Accounting for its marginal external cost on other satellites outside the constellation, j should be deorbited at time $t^*_j \lesseqgtr \bar{t}_j$. Whether the socially optimal deorbit time is sooner or later than the privately optimal deorbit time will depend on the relation between the effects of satellites and launch debris on collision risk and debris growth. If the marginal effect on collision risk of a satellite exceeds the marginal effect of new launch debris, $t^*_j < \bar{t}_j$. A satellite-specific stock control with price p_{jt} would be able to ensure that $t^*_j = \bar{t}_j$ for all j in each constellation. In the case of Iridium, π fell below V^d and (some of) its operators believed it would stay there.

access tolls since congestion pricing affects marginal road use decisions throughout the road while access tolls affect marginal road use decisions only at points of entry.

Proposition 14. *(New stock controls reduce risk and debris, a new flow controls increase them) The equilibrium expected collision risk, the expected future debris stock, and the probability of Kessler Syndrome will*

- *decrease when a generic stock control is introduced;*
- *increase when a generic flow control is introduced.*

Proof. The effect of introducing a control on the equilibrium collision risk: Suppose a control is scheduled to be introduced at date t. The equilibrium collision rate under open access in period $t-2$ is

$$\hat{X}_{t-2} : E_{t-2}[\ell_{t-1}] = r_s - r. \tag{3.41}$$

In general, the equilibrium expected future collision rate is an increasing function of the current launch rate. Proposition 12 establishes that $X_{t-1} < \hat{X}_{t-1}$ (the equivalent uncontrolled open access launch rate in $t-1$) if the control scheduled to be introduced in t is a stock control. Similarly, Proposition 12 establishes that $X_{t-1} > \hat{X}_{t-1}$ if the control scheduled to be introduced in t is a flow control. Thus, introducing a generic stock control must reduce the equilibrium expected future collision rate, while introducing a generic flow control must increase it.

The effect of introducing a control on the expected future debris stock: The debris stock in t is an increasing function of both the collision rate and launch rate in $t-1$:

$$D_t = (1-\delta)D_{t-1} + G(S_{t-1}, D_{t-1}, \ell_{t-1}) + mX_{t-1}. \tag{3.42}$$

Since the launch rate decreases when a stock control is introduced, D_t is mechanically reduced due to the reduction in launch debris (mX_{t-1}). Similarly, when a flow control is introduced, the increased launch rate increases D_t through launch debris. Even without launch debris, the same conclusion holds in the following period because the expected collision risk is an increasing function of the launch rate. Formally, suppose

$m = 0$:

$$E_{t-1}[D_{t+1}] = (1-\delta)D_t + E_{t-1}[G(S_t, D_t, \ell_t)] \tag{3.43}$$

Because the expected number of new fragments is linear in probabilities,

$$E_{t-1}[D_{t+1}] = (1-\delta)D_t + G(S_t, D_t, E_{t-1}[\ell_t]) \tag{3.44}$$

$$\implies \frac{\partial E_{t-1}[D_{t+1}]}{\partial p_t} = \frac{\partial G(S_t, D_t, E_{t-1}[\ell_t])}{\partial p_t} \tag{3.45}$$

$$= \frac{\partial G(S_t, D_t, E_{t-1}[\ell_t])}{\partial E_{t-1}[\ell_t]} \frac{\partial E_{t-1}[\ell_t]}{\partial X_{t-1}}. \tag{3.46}$$

Both terms on the right-hand side of the final line are positive: the expected number of new fragments formed in collisions is increasing in the expected number of collisions, and the expected number of collisions is increasing in the number of satellites launched the previous period.

The effect of introducing a control on the probability of Kessler Syndrome: The probability of Kessler Syndrome occurring in t, given information in $t-1$, is

$$Pr_t(D_t(\ell_{t-1}) > D^{\kappa}). \tag{3.47}$$

We have already established that for any ℓ_{t-1}, D_t will decrease if a stock control is implemented in t, and increase if a flow control is implemented in t. D^{κ} is a function of the physical parameters of the orbit, and is unaffected by economic controls. Consequently, introducing a stock control must reduce the probability of Kessler Syndrome, while introducing a flow control must increase the probability of Kessler Syndrome. □

The result in Proposition 14 is one of the main reasons why a regulator should prefer stock controls to flow controls. If the imposition of a control raises the equilibrium collision risk, it is possible that it may also cause Kessler Syndrome. The economic intuition for this effect is simple. Flow controls generate rents for firms who already own satellites. Imposing a flow control therefore creates an incentive for marginal launchers to become satellite owners before the flow control is imposed. One way around this would be to levy a flow control with no prior notice. I do not consider this possibility, as it would force firms to form expectations over the regulator's possible actions. These expectations may be difficult for the regulator to elicit truthfully, as firms might anticipate the regulator's desire to act in unexpected ways to reduce firms' profits. Such expectations would make it difficult to implement regulatory policy effectively. Stock controls sidestep this issue by focusing on satellite owners. Forward-looking launchers internalize their expected costs due to the control, and can be appropriately disincentivized against launching without distorting the incentives of current satellite owners.

The relative advantage of stocks vs. flows:

The question of ultimate interest to a regulator is likely one of policy choice: "which type of instrument is better, and why?" The results so far - particularly Proposition 14 - suggest that stock controls should be preferred to flow controls along generic paths. Proposition 15 compiles the results so far to answer the policy choice question along optimal paths. Since stock controls can be initiated without losing control of the launch rate and induce deorbits when necessary, they can achieve first-best outcomes in every state of the world. Flow controls cannot. Even if interior launch rates are optimal forever and no deorbits are ever required, flow controls will achieve less social welfare than stock controls when they are put into place.

Proposition 15. *(The relative advantage of stocks vs. flows) The expected social welfare under an optimal stock control strictly exceeds the expected social welfare under an optimal flow control for an arbitrary horizon where a control must be initiated, used to stop all launches, or used to force net deorbits.*

Proof. The fleet welfare from both controls can be equal along interior equilibrium paths. However, when the flow control is initiated, Proposition 12 shows that it will cause the launch rate to exceed the uncontrolled open access launch rate whereas a stock control would not. Proposition 14 shows that the launch bunching from initiating a flow control will also cause the risk of Kessler Syndrome to increase. In those periods, stock controls will achieve strictly greater expected social welfare than flow controls.

Proposition 13 shows that flow controls may not be able to induce net deorbits (never with positive prices and only possibly with negative prices), while stock controls can always do so. Therefore, for arbitrary paths with positive prices where the regulator must either initiate control, shut down orbital access, or induce net deorbits, stock controls achieve strictly greater expected social welfare than flow controls. □

Proposition 15 is fairly straightforward, and may even understate the advantages of stock controls over flow controls. From a computational perspective, optimal flow controls are much harder to implement than optimal stock controls because they require attention to the entire control time path. Lemma 6 in the Appendix shows that price-based flow controls must have an exploding price path to balance the launch incentives and disincentives described in Lemma 1 and Figure 3.4. One solution to this may be to use a quantity flow control, such as a launch permit quota system. However, the regulator must still commit to a time path of quantity policies when the flow control is implemented, cannot prevent launch bunching before the policy goes into effect, and cannot induce deorbits. Stock controls face none of these issues. A one-period-forward forecast of the marginal external cost is sufficient, which would have been required anyway under a flow control. The regulator faces no commitment issues and can precisely control the number of satellites in orbit at any given time.

3.2.1.3 Optimal space traffic control policies

Before finishing my discussion of stock and flow controls, I illustrate optimal control policy functions by simulation. For clarity and computational tractability, I use deterministic simulations where $E_t[\ell_{t+1}] = L(S_{t+1}, D_{t+1})$. Figure 3.7 shows an example of optimal stock and flow control policies as a satellite tax and a launch tax. Figure 3.6 shows the underlying satellite stocks, debris stocks, and launch rates used to compute Figure 3.7.

The point of figures 3.6 and 3.7 is to show the qualitative properties of optimal stock and flow controls. While both types of tax vary with the marginal external cost of launching a satellite ($E_t[\xi(S_{t+1}, D_{t+1})]$), only the satellite tax varies positively with the marginal external cost. This is a convenient feature for applying stock controls: the behavior of an optimal stock control is more intuitive than that of an optimal flow control. The reasoning behind this behavior is described in in Lemma 1 and Figure 3.4.

3.2.2 Active debris removal and open access

I now turn to the effects of active debris removal technologies on orbit use. My main result, Proposition 16, shows that while active debris removal can mechanically reduce the debris stock no matter

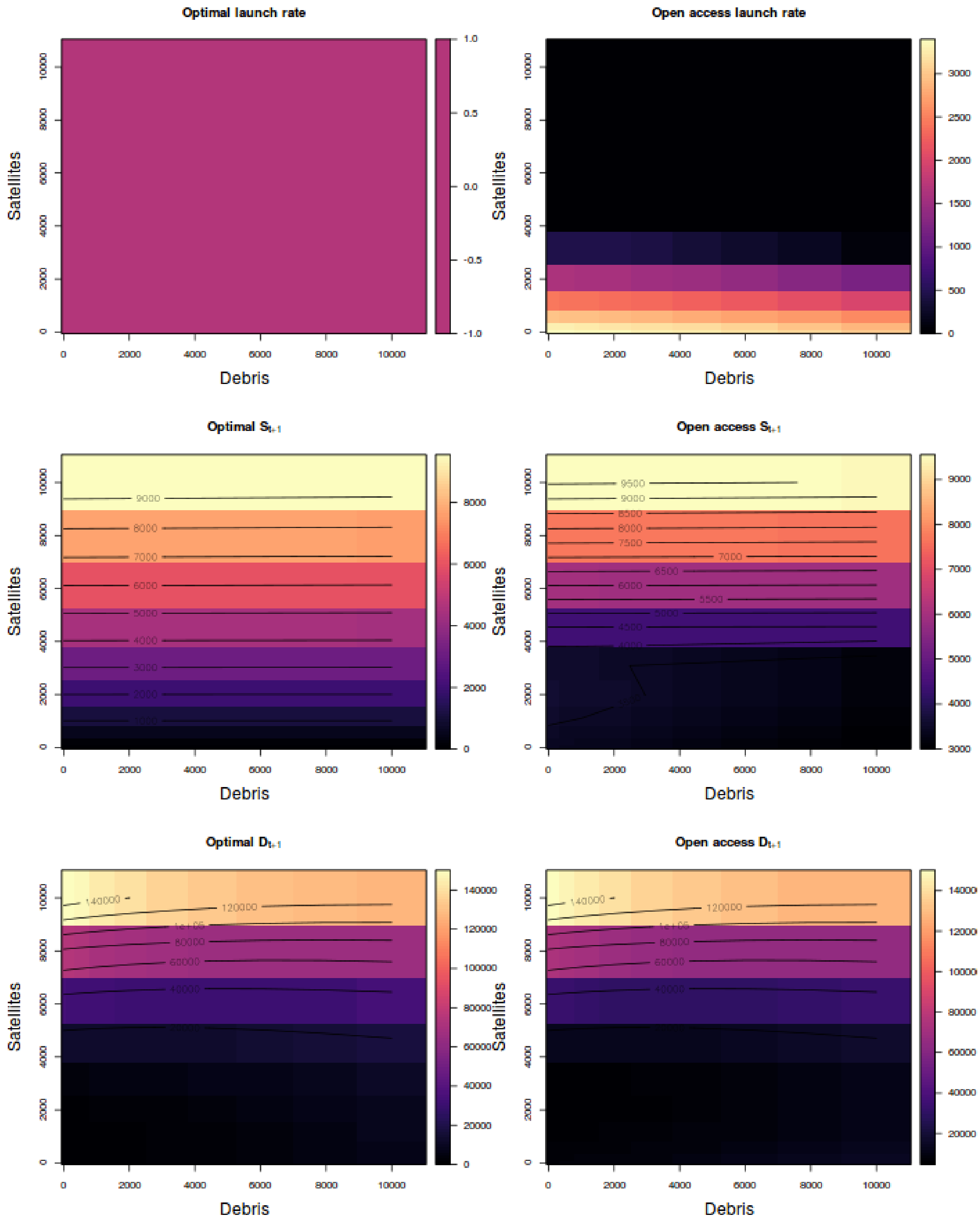

Figure 3.6: *Optimal and open access stocks and launch rates.*
Left column: Optimal launch rate (X_t), next-period satellite stock (S_{t+1}), and next-period debris stock (D_{t+1}).
Right column: Optimal launch rate (X_t), next-period satellite stock (S_{t+1}), and next-period debris stock (D_{t+1}).
The per-period return on a satellite is normalized to 1, the discount factor is set to 0.95, and the launch cost is set to 10. The open access next-period satellite stock is small but not zero in the upper right of the figure, while the optimal next-period satellite stock is zero there.

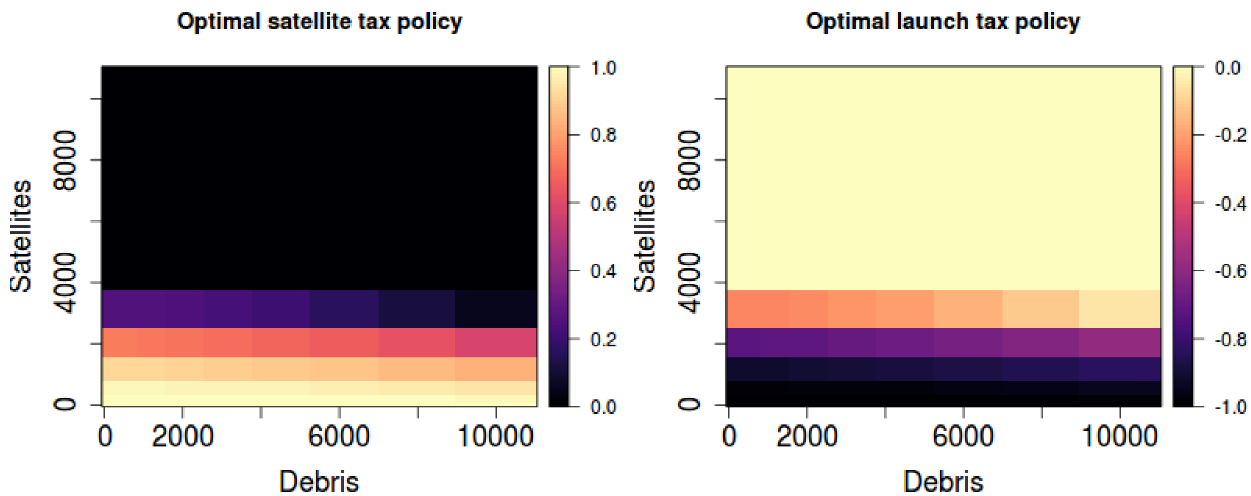

Figure 3.7: *Optimal space traffic control policies.*
Upper panels: The collision risk in $t+1$ ($E_t[\ell_{t+1}]$) under the optimal launch plan (left) and open access launch plan (right).
Lower left panel: An optimal satellite tax (stock control) in $t+1$.
Lower right panel: An optimal launch tax (flow control) in $t+1$. The tax in period t is normalized to 0.
The tax rates should be read as multiples of the per-period satellite return, which is normalized to 1.

how it is financed, it can only reduce the equilibrium risk of satellite-destroying collisions to the extent that satellite owners pay for debris removal. I show this in two steps. First, I show that exogenously provided removal which is free to satellite owners will reduce the debris stock but increase the satellite stock. The increase in the satellite stock will exactly offset the decrease in risk from debris removal, leaving the equilibrium collision risk unchanged. Then, I consider a case where exogenous debris removal involves a mandatory fee paid by satellite owners. I show that as the fee goes to zero, the collision risk returns to the original open access level.

Lastly, I show how endogenously chosen debris removal purchased by cooperative satellite owners reduces the debris stock, collision risk, and risk of Kessler Syndrome, while also allowing more firms to launch satellites. These results depend on some auxiliary properties of cooperative debris removal and open access launching with debris removal, shown in the Appendix, section D.9. Though the jointly-optimal launch and removal plan is analytically complicated, I simulate the fleet planner's launch and removal plans and compare them to the launch and removal plans under open access and cooperative removal. The simulations show that cooperative decentralized removal plans are similar to the planner's removal plans, though the launch plans differ more substantially. While the cooperative decentralized removal plan matches the regulator's in many regions of the state space, the firms do not begin debris removal as quickly as the planner would. Open access launchers respond to satellite and debris in orbit very differently than the planner: while the planner effectively ignores debris in choosing launch rates, open access firms are disincentivized by debris and not sufficiently responsive to satellites, at least until the open access launchers expect cooperative satellite owners to begin debris removal. Economically, open access launchers do not work with satellite owners to coordinate their launch activity due to their inability to secure property rights over orbits. Their launch behavior prevents cooperative satellite owners from internalizing the full value of debris removal, thus inducing the satellite owners to begin debris removal later than the planner would.

3.2.2.1 An economic model of active debris removal

With debris removal technology available, satellite owners purchase R_{it} units of removal from a competitive debris removal sector. The price of a unit of removal is c_t. The amount of debris removed is nonnegative and cannot exceed the total amount of debris in orbit. Since firms are identical, each satellite owner will choose the same level of removal, making the total amount of debris removed $R_t = S_t R_{it} \leq D_t$. The maximum amount that an individual satellite owner could choose to remove is D_t/S_t. To consider the best-case outcomes of debris removal, I focus on cooperative removal plans between satellite owners. I establish an economically intuitive necessary and sufficient condition for cooperation to be locally self-enforcing in section D.7 of the Appendix. The condition is stated below in Assumption 5.

Assumption 5. *(Making cooperation locally self-enforcing) For any non-zero cooperative removal plan, the change in the equilibrium collision risk before debris removal is greater than the ratio of the removal price to the launch cost, that is,*

$$\left.\frac{\partial \tilde{E}[\ell_t | S_t, \bar{D} - R_t]}{\partial \bar{D}}\right|_{\bar{D}=D_t} > \frac{c_t}{F} \quad \forall R_t \in [0, \bar{D}].$$

While the validity of Assumption 5 will need to be evaluated empirically for specific technologies and orbital regimes, I assume it always holds in the analysis below.[19]

The value of a satellite after debris has been removed and ℓ_t has been drawn is

$$Q_i(S_t, D_t, \ell_t, X_t) = \pi + \beta[(1-\ell_t)\tilde{Q}_i(S_{t+1}, D_{t+1}) + \ell_t E_t[V_i(S_{t+1}, D_{t+1}, \ell_{t+1}, X_{t+1})]]. \tag{3.48}$$

[19]For example, cooperation may be enforced by a grim trigger mechanism under which any deviation by any firm results in no debris removal at all.

The value of a satellite owner who purchases debris removal before the loss is

$$\tilde{Q}_i(S_t,D_t) = \max_{0\leq R_{it}\leq D_t/S_t} \{-c_t R_{it} + \tilde{E}_t[Q_i(S_t,D_t,\ell_t,X_t)]\} \tag{3.49}$$

$$\text{s.t. } \ell_t \sim \phi(\ell_t|S_t,D_t-R_t)$$

$$Q_i(S_t,D_t,\ell_t,X_t) = \pi + \beta[(1-\ell_t)\tilde{Q}_i(S_{t+1},D_{t+1}) + \ell_t E_t[V_i(S_{t+1},D_{t+1},\ell_{t+1},X_{t+1})]]$$

$$S_{t+1} = S_t(1-\ell_t) + X_t$$

$$D_{t+1} = (D_t - R_t)(1-\delta) + G(S_t,D_t-R_t,\ell_t) + mX_t.$$

The value of a launcher is

$$V_i(S_t,D_t,\ell_t,X_t) = \max_{x_{it}\in\{0,1\}} \{(1-x_{it})\beta E_t[V_i(S_{t+1},D_{t+1},\ell_{t+1},X_{t+1})] + x_{it}[\beta\tilde{Q}_i(S_{t+1},D_{t+1}) - F]\} \tag{3.50}$$

$$\text{s.t. } \tilde{Q}_i(S_t,D_t) = \max_{0\leq R_{it}\leq D_t/S_t} \{-c_t R_{it} + \tilde{E}_t[Q_i(S_t,D_t,\ell_t,X_t)]\}$$

$$\ell_t \sim \phi(\ell_t|S_t,D_t-R_t)$$

$$Q_i(S_t,D_t,\ell_t,X_t) = \pi + \beta[(1-\ell_t)\tilde{Q}_i(S_{t+1},D_{t+1}) + \ell_t E_t[V_i(S_{t+1},D_{t+1},\ell_{t+1},X_{t+1})]]$$

$$S_{t+1} = S_t(1-\ell_t) + X_t$$

$$D_{t+1} = (D_t - R_t)(1-\delta) + G(S_t,D_t-R_t,\ell_t) + mX_t.$$

Under a generic launch plan, the decision to remove debris is dynamic: removal today will impact the amount of debris tomorrow through the number of satellite destructions and the number of debris-debris collisions. Under open access, the value of a satellite tomorrow will always be driven down to the current value of the launch cost, so the future benefits of removal will never accrue to today's satellite owners. This makes the removal decision under open access static: the only benefit of debris removal internalized by satellite owners today is the way that it changes the probability that their satellite is destroyed. Even though the cost of removal is linear, nonlinearity in the coupling between the debris stock and the collision rate can yield an interior solution to the removal decision. This simplifies analysis of cooperative removal plans given open access launch plans.

Open access launching: Under open access, firms will launch satellites until the value of launching is zero:

$$\forall t, X_t : \; V_i(S_t, D_t, \ell_t, X_t) = 0 \tag{3.51}$$

$$\implies \beta \tilde{Q}(S_{t+1}, D_{t+1}) = F \tag{3.52}$$

$$\implies Q_i(S_t, D_t, \ell_t, X_t) = \pi + (1 - \ell_t)F. \tag{3.53}$$

Taking R_t as fixed, and assuming that launchers plan to choose R_{it+1} optimally when they are satellite owners, the flow condition determining the launch rate is

$$\pi = rF + \tilde{E}_{t+1}[\ell_{t+1}]F + R_{it+1}c_{t+1}. \tag{3.54}$$

This can be rewritten to yield the equilibrium collision risk,

$$\tilde{E}_{t+1}[\ell_{t+1}] = r_s - r - \frac{c_{t+1}}{F} R_{it+1}. \tag{3.55}$$

Equation 3.55 states that the equilibrium collision risk will be equal to the excess return on a satellite ($r_s - r$) minus the rate of total removal costs the launchers will face when they become satellite owners ($\frac{c_{t+1}}{F} R_{it+1}$). If there were no removal technology, $R_t = 0 \; \forall t$, and the equilibrium collision rate would be equal to the excess return on a satellite. In any period t, the decisions to launch and to remove debris are undertaken by different firms. Potential launchers make the launch decisions, while current satellite owners make the removal decisions. Once they become satellite owners, launchers will face the removal decision. While open access makes satellite owners myopic, satellite launchers remain forward-looking.

Cooperative private debris removal: Profit maximizing cooperative satellite owners will demand debris removal until their marginal benefit from removal equals its marginal cost. Under open access to orbit, the first-order condition for an interior solution to the maximization problem in system of equations 3.49 is

$$R_{it} : c_t = \frac{\partial \tilde{E}_t[\ell_t]}{\partial D_t} S_t F, \tag{3.56}$$

with the second-order condition

$$R_{it} : -\frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t^2} S_t^2 F < 0. \tag{3.57}$$

Intuitively, open access removes any potential future benefit or cost from debris removal. Satellite owners will not get to reap any benefits from increasing Q_{it+1} because today's potential launchers will enter and capture them. Equation 3.56 therefore states that under open access launching, satellite owners will purchase debris removal until the price of a unit of removal (c_t) is equal to the static private marginal benefit ($\frac{\partial \tilde{E}_t[\ell_t]}{\partial D_t} S_t F$). That benefit has three pieces: the value of their satellite next period, F; the number of owners who will make the same removal decision, S_t; and the change in the probability that their satellite is destroyed at the end of the period, $\frac{\partial \tilde{E}_t[\ell_t]}{\partial D_t}$.

Under open access firms launch until zero profits, while satellite owners remove debris until marginal benefits equal marginal costs. But current launchers are future satellite owners. If they could not coordinate as launchers, how can they do so as satellite owners? The answer is property rights. International space law gives satellite launchers ownership of any objects they put into space, even after their useful life is over. As a result, satellite owners must either purchase or exercise rights to specific pieces of debris in order to remove them. This allows satellite owners to coordinate debris removal. I assume they do so in a cooperative and efficient manner to focus on the best-case scenario for active debris removal. I ignore both the complications of decentralized bargaining between many parties and the difficulties of attributing ownership to specific small pieces of debris. Transaction and information costs associated with debris removal are relevant to policy design and implementation, but beyond my scope here.

3.2.2.2 Exogenous debris removal for free and for a mandatory fee

To develop intuition for how debris removal can reduce the equilibrium collision risk, consider a setting where $\bar{R}$ units of debris are removed from orbit every period by a regulator. Such policies are advocated for by some in the space debris literature, for example Bradley and Wein (2009) and Akers (2012). If the removal is costly to satellite owners, it is because the regulator forces them to pay a fixed fee of $\bar{c}$ per unit removed. Denote the equilibrium collision risk with exogenous removal for a mandatory fee as $E_t^R[\ell_{t+1}]$, and the equilibrium collision risk with no removal as $E_t[\ell_{t+1}]$. The open access equilibrium condition for forward-looking launchers is then

$$\pi = rF + \tilde{E}^R_{t+1}[\ell_{t+1}]F + \bar{R}\bar{c}, \tag{3.58}$$

while in the absence of removal, firms would launch until

$$\pi = rF + \tilde{E}_{t+1}[\ell_{t+1}]F. \tag{3.59}$$

Inspecting the two equations above reveals that, as the mandatory removal fee approaches zero, the equilibrium collision risk with removal approaches the equilibrium collision risk without removal. Proposition 16 shows this formally.

Proposition 16. *(Satellite owners must pay for collision risk reduction) Any debris removal technology will reduce the equilibrium collision risk if and only if:*

(1) some amount of debris is removed, and

(2) satellite owners pay for the removal.

Proof. Let the equilibrium collision risk with debris removal be $\tilde{E}^R_{t+1}[\ell_{t+1}]$. The amount of debris removed per satellite owner is $\bar{R}$, and the per-unit cost to satellite owners is $\bar{c}$. From equation 3.58,

$$\tilde{E}^R_{t+1}[\ell_{t+1}] = r_s - r - \frac{\bar{R}\bar{c}}{F}. \tag{3.60}$$

From equation 3.11, the equilibrium collision risk without debris removal is[20]

$$\tilde{E}_{t+1}[\ell_{t+1}] = r_s - r. \tag{3.61}$$

If no debris is removed, $\bar{R} = 0$. If some debris is removed but satellite owners pay nothing for it, $\bar{c} = 0$. In either case, $\tilde{E}^R_{t+1}[\ell_{t+1}] = \tilde{E}_{t+1}[\ell_{t+1}]$. If and only if some debris is removed ($\bar{R} > 0$) and satellite owners pay something for it ($\bar{c} > 0$), $\tilde{E}^R_{t+1}[\ell_{t+1}] < \tilde{E}_{t+1}[\ell_{t+1}]$.

[20]While the expectation in equation 3.11 looks slightly different from the expectation in equation 3.61, $\tilde{E}_{t+1}[\ell_{t+1}]$ is the same integral as $E_t[\ell_{t+1}]$. The difference in notation is described in equation 3.20.

More generally, for any positive amount of debris removal, the equilibrium collision risk reduction is increasing in the amount that satellite owners pay for debris removal:

$$\forall \bar{R} > 0, \quad \tilde{E}_{t+1}[\ell_{t+1}] - \tilde{E}^{R}_{t+1}[\ell_{t+1}] = \frac{\bar{R}\bar{c}}{F}, \tag{3.62}$$

$$\frac{\partial(\tilde{E}_{t+1}[\ell_{t+1}] - \tilde{E}^{R}_{t+1}[\ell_{t+1}])}{\partial \bar{c}} = \frac{\bar{R}}{F} > 0. \tag{3.63}$$

As firms pay less and less for debris removal, the equilibrium collision risk with debris removal smoothly approaches the equilibrium collision risk without debris removal:

$$\lim_{\bar{c}\to 0} \tilde{E}^{R}_{t+1}[\ell_{t+1}] = \lim_{\bar{c}\to 0}(\pi - rF - \bar{R}\bar{c}) = \pi - rF = \tilde{E}_{t+1}[\ell_{t+1}]. \tag{3.64}$$

□

Since the debris stock will be lower due to removal, the launch rate will be higher with free exogenous removal than it would under open access with no removal. Economically, free removal clears up space for new launchers to enter the orbit. This case highlights the mechanism through which active debris removal can reduce the equilibrium collision risk: not by mechanically reducing the amount of debris in orbit, but by reducing the excess return of a satellite. This mechanism also acts in the case with endogenous debris removal, as shown in Proposition 17. Despite this mechanism, the launch rate may be larger with debris removal than without. While the reduction of excess return on a satellite will lower the launch rate, the reduction in debris will increase the launch rate.

This example also highlights the main reason why active debris removal can reduce collision risk: not because it removes debris, but because it approximates a stock control. This suggests that controls on debris removal could be more effective than flow controls on satellites at reducing collision risk. Figure 3.8 illustrates the differences between exogenous debris removal for free and for a mandatory fee. When satellite owners choose how much debris to remove, however, this type of approach must account for current satellite owners' and launchers' responses to the price of removal. These responses are discussed in Propositions 8 and 24.

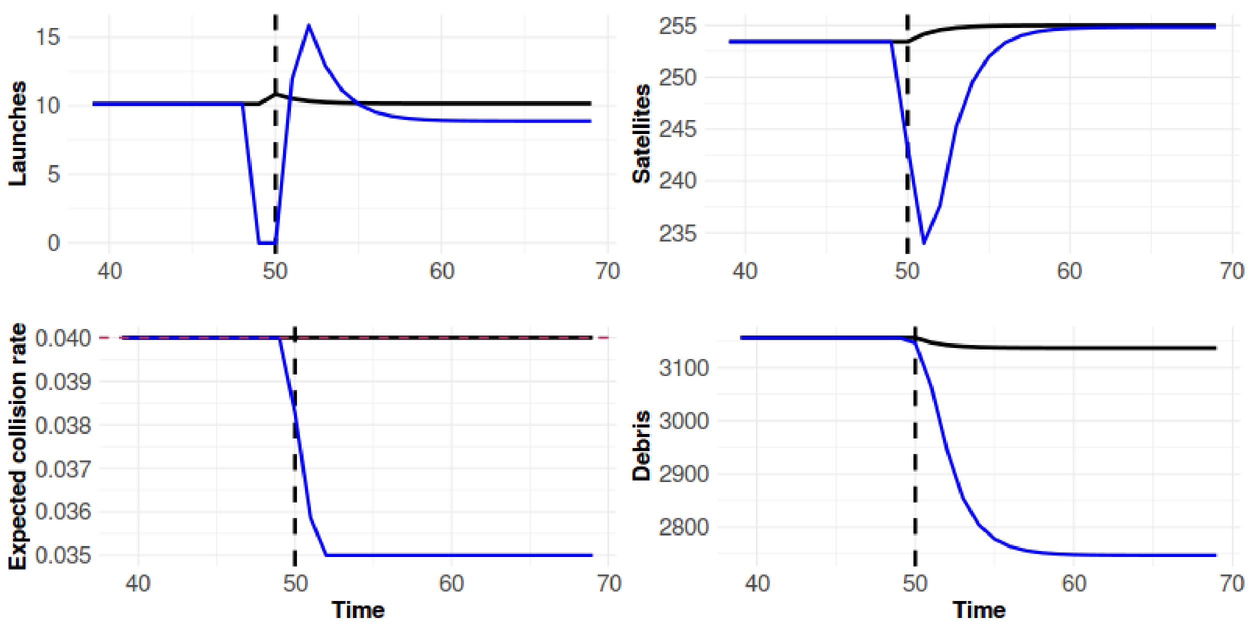

Figure 3.8: *The effects of exogenous removal for free (black line) or a mandatory fee (blue line). When debris removal is provided to satellites owners for free, potential launchers respond by launching more satellites - even though the debris stock falls, the equilibrium collision risk remains unchanged. The equilibrium collision risk will fall when active debris removal is an option if and only if it is costly to satellite owners. In the case of costly debris removal, the launch rate falls to zero until the expected collision risk is no longer above the new equilibrium level. The dashed red line shows the equilibrium collision risk under open access.*

3.2.2.3 Endogenous debris removal financed by satellite owners

Figure 3.9 illustrates the effects of introducing active debris removal paid for by cooperative satellite owners. Unlike when removal is provided exogenously, endogenous removal can induce more firms to launch satellites before the technology becomes available. Despite the cost of cooperating with others and paying for removal, lower expected collision costs due to debris removal and lower individual contributions due to additional firms paying for removal makes it optimal for potential launchers to enter the orbit.

The introduction of debris removal technologies affects equilibrium orbital stocks as well as open access launch incentives. I explore the properties of cooperative private debris removal demands and open access launching further in the Appendix, section D.9. Two of these - the uniqueness of the cooperatively-optimal post-removal debris stock and the potential for a "dynamic virtuous cycle" of debris removal - are driven by the incentives of satellite owners given open access. The third result describes intuitive physical and economic conditions under which the demand for satellite ownership by satellite launchers will be decreasing in the launch cost. Violations of these conditions may be plausible depending on the values of physical parameters.

3.2.2.4 Cooperative removal and open access risks

Ultimately, policymakers considering active debris removal technologies will want to know how debris removal will affect collision and Kessler Syndrome risks. In this section, I examine how active debris removal which is costly to satellite owners will change the equilibrium collision risk, the equilibrium future debris stock, and the equilibrium future probability of Kessler Syndrome. The results of this section, Propositions 17 and 18, establish that the use of active debris removal can reduce both equilibrium collision risk and the risk of Kessler Syndrome. While the risk of Kessler Syndrome can be mechanically reduced by removing debris no matter how removal is financed, reducing the equilibrium collision risk requires satellite owners to finance debris removal.

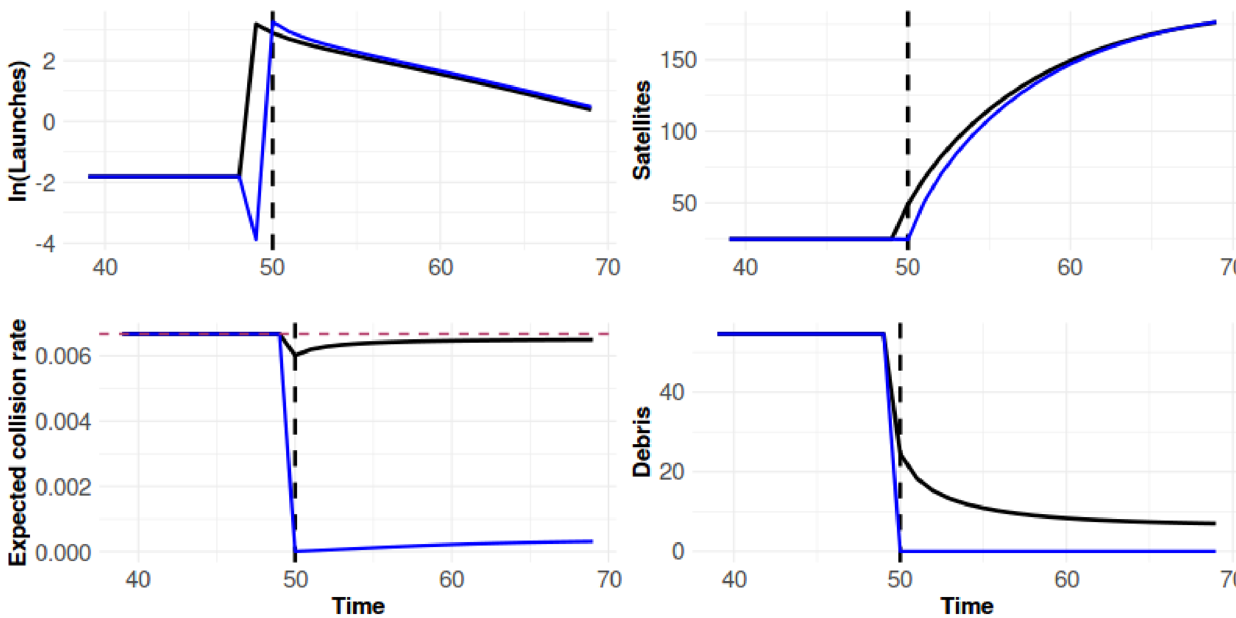

Figure 3.9: *The effects of endogenously chosen cooperative debris removal (blue line) and exogenous removal for a mandatory fee (black line). The exogenous removal path in the exogenous case is set equal to endogenous removal path. Endogenous removal reduces both the equilibrium collision risk and the debris stock more effectively than exogenous removal, even if the same removal schedule is used. The endogenous removal schedule and launch response involves completely cleaning the orbit initially, and keeping the orbit relatively clean after. The same removal schedule provided exogenously induces firms to launch earlier than they would if they chose the schedule. The dashed red line shows the equilibrium collision risk under open access.*

Proposition 17 extends Proposition 16 by considering the time path of collision risk when debris removal is introduced and when debris removal is ongoing. The intuition is similar to that of Proposition 16. Proposition 18 relies on some auxiliary properties of debris removal, shown in section D.9 of the Appendix. The key intuition is that the cooperatively-optimal level of post-removal debris is a constant. As a result, even if there is an increase in the number of satellites due to debris removal, the risk of Kessler Syndrome will be reduced because firms will continuously purchase removal to keep the debris stock at its new, lower level.

Costly debris removal services can reduce the equilibrium collision risk: The removal of debris, all else equal, should reduce the collision rate. Whether it reduces the equilibrium collision risk depends on how potential satellite launchers respond to this reduction in risk. If debris removal spurs enough new entry, debris removal may result in higher collision rates. The logic seems plausible: if debris removal should reduce the collision rate, then more firms would be able to take advantage of the cleaner orbit and should therefore launch.

This logic would be correct but for an important detail: firms which launch satellites at t will become firms which own satellites at $t+1$. As they decide whether to launch or not, forward-looking firms account for their expected debris removal expenses as satellite owners. If the firm anticipates wanting to purchase debris removal services once its satellite is on orbit, and these services are costly, the introduction of the technology must reduce the excess return from launching the satellite. Since open access equates the expected collision risk with the excess return, the reduction in excess return also reduces the expected collision risk. However, if debris removal is introduced as a free service which potential launchers anticipate not paying for, the equilibrium collision risk will remain at the earlier open access level.

Proposition 17. *(ADR can reduce collision risk) The introduction of costly debris removal services in period t will reduce the equilibrium collision risk in t if and only if*

(1) it is costly to remove debris in t, and

(2) it is privately optimal for cooperative satellite owners to remove some amount of debris in t.

Ongoing active debris removal will reduce the equilibrium collision risk if and only if individual cooperative debris removal expenditures increase from period t to $t+1$. Formally,

$$E_{t-1}[\ell_t] - E_t[\ell_{t+1}] > 0 \iff c_{t+1}R_{it+1} > c_t R_{it}.$$

Proof. The proof of this result is similar to the proof of Proposition 16, so I omit it from the main text. See Appendix section B. □

At first, the collision risk will decrease because of the expenditure that current launchers anticipate making once they are satellite owners. The fact that debris removal directly removes orbital debris is incidental to this risk reduction. As in the exogenous case, open access drives new launchers to take advantage of the newly-cleared space by launching satellites until the risk is the same as it was before removal because available. Subsidies for debris removal to satellite owners may increase the equilibrium amount of debris removed, but would not affect the equilibrium collision risk. Open access will still dissipate rents from orbit use; subsidized debris removal would only tilt the combination of new satellites and debris which equilibriates the system toward new satellites. Ongoing debris removal can only keep the collision rate below the no-ADR open access collision rate if and only if it is costly to potential launchers. If potential launchers anticipate that the cost will reduce, or that it will not be optimal to purchase removal as satellite owners, the collision rate will return to the no-ADR open access level.

Cooperative costly debris removal will reduce the equilibrium Kessler syndrome probability:

Since preventing Kessler Syndrome is one of the key motivations for developing active debris removal technologies, it is natural to wonder if debris removal will achieve this goal. Since Kessler Syndrome is caused by the amount of debris exceeding a threshold, debris removal in t will reduce the probability of Kessler Syndrome in $t+1$ if it is certain to reduce the $t+1$ debris stock. More precisely, the change in probability of Kessler Syndrome in $t+1$ is equal to the probability that the change in the $t+1$ debris stock due to removal in t is positive, plus the product of the change in expected collision risk due to removal and the original probability of Kessler Syndrome. Since the change in the $t+1$ debris stock due to removal is negative with probability one, debris removal will reduce future probability of Kessler Syndrome.

Proposition 18. *(Debris removal will reduce the future probability of Kessler Syndrome) Debris removal in t will reduce the probability of Kessler Syndrome in $t+1$.*

Proof. Kessler Syndrome will occur in $t+1$ when

$$D_{t+1} - D^{\kappa} > 0, \tag{3.65}$$

where D^{κ} is the Kessler threshold. Suppose that Kessler Syndrome has not already occurred ($D_t - D^{\kappa} < 0$). Under the probability density for satellite-destroying collisions in t ($\phi(\ell_t|S_t, D_t - R_t)$), the probability in t of Kessler Syndrome in $t+1$ is

$$Pr_t(D_{t+1} - D^{\kappa} > 0|S_t, D_t - R_t) = \int_0^1 \mathbb{1}(D_{t+1}(\ell_t) - D^{\kappa} > 0)\phi(\ell_t|S_t, D_t - R_t)d\ell_t. \tag{3.66}$$

The change in this probability due to an increase in R_t is

$$\begin{aligned}
\frac{dPr_t(D_{t+1}(\ell_t) - D^{\kappa} > 0|S_t, D_t - R_t)}{dR_t} &= \frac{\partial}{\partial R_t}\int_0^1 \mathbb{1}(D_{t+1}(\ell_t) - D^{\kappa} > 0)\phi(\ell_t|S_t, D_t - R_t)d\ell_t && (3.67)\\
&= \int_0^1 \mathbb{1}(\frac{\partial D_{t+1}(\ell_t)}{\partial R_t} > 0)\phi(\ell_t|S_t, D_t - R_t)d\ell_t \\
&\quad + \int_0^1 \mathbb{1}(D_{t+1}(\ell_t) - D^{\kappa} > 0)\frac{\partial \phi(\ell_t|S_t, D_t - R_t)}{\partial R_t}d\ell_t && (3.68)\\
&= Pr_t\left(\frac{\partial D_{t+1}(\ell_t)}{\partial R_t} > 0|S_t, D_t - R_t\right) \\
&\quad - \left.\frac{\partial Pr_t(D_{t+1}(\ell_t) - D^{\kappa} > 0|S_t, \tilde{D} - R_t)}{\partial \tilde{D}}\right|_{\tilde{D} = D_t}. && (3.69)
\end{aligned}$$

The first term in equation 3.69 is zero because debris removal reduces the future debris stock for any draw of the collision rate, and the form of the second term in equation 3.69 follows from Lemma 2. To see that the first term is zero, define the open access launch rate as an implicit function $X_t = X(S_t, D_t - R_t, \ell_t)$ defined by equation 3.54. Then, differentiate D_{t+1} with respect to R_t:

$$\begin{aligned}
D_{t+1}(\ell_t) &= (D_t - R_t)(1-\delta) + G(S_t, D_t - R_t, \ell_t) + mX_t \\
\implies \frac{\partial D_{t+1}(\ell_t)}{\partial R_t} &= -\left[1 - \delta + \frac{\partial G(S_t, D_t - R_t, \ell_t)}{\partial D_t} + m\frac{\partial X_t}{\partial D_t}\right].
\end{aligned}$$

From Proposition 22 and applying the Implicit Function Theorem to equation 3.54,

$$\frac{\partial X_t}{\partial D_t} = -\frac{\frac{\partial R_{it+1}}{\partial D_{t+1}}(1-\delta+\frac{\partial G(S_t,D_t-R_t,\ell_t)}{\partial D_t})(1-\frac{\partial R_t}{\partial D_t})}{\frac{\partial E_t[\ell_{t+1}]}{\partial S_{t+1}}F + m\left(\frac{\partial E_t[\ell_{t+1}]}{\partial D_{t+1}}F + \frac{\partial R_{it+1}}{\partial D_{t+1}}c_{t+1}\right) + \frac{\partial R_{it+1}}{\partial S_{t+1}}c_{t+1}}$$

$$= 0 \text{ whenever } R_{it} > 0 \because \frac{\partial R_t}{\partial D_t} = 1 \text{ from Proposition 22.}$$

Therefore,

$$\frac{\partial D_{t+1}(\ell_t)}{\partial R_t} = -\left[1-\delta+\frac{\partial G(S_t,D_t-R_t,\ell_t)}{\partial D_t}\right] < 0 \ \ \forall \ell_t \in [0,1].$$

Since the statement holds for all possible realizations of ℓ_t,

$$Pr_t\left(\frac{\partial D_{t+1}(\ell_t)}{\partial R_t} > 0|S_t, D_t - R_t\right) = 0.$$

The change in the probability of Kessler Syndrome due to a change in debris removal is then

$$\frac{dPr_t(D_{t+1}(\ell_t) - D^\kappa > 0|S_t, D_t - R_t)}{dR_t} = -\left.\frac{\partial Pr_t(D_{t+1}(\ell_t) - D^\kappa > 0|S_t, \tilde{D} - R_t)}{\partial \tilde{D}}\right|_{\tilde{D}=D_t}. \tag{3.70}$$

The right hand side of equation 3.70 is the negative of the change in the probability of Kessler Syndrome from the shift in the distribution of collision rates which a marginal amount of debris would cause. It is not precisely the same as the effect of another unit of debris, since the debris argument of $D_{t+1}(\ell_t)$ is held constant while the debris argument of $\phi(\ell_t|S_t, D_t - R_t)$ is increased slightly. By Assumption 3, increasing the amount of debris in orbit will shift the conditional density of the collision rate toward 1. The fact that $\mathbb{1}(D_{t+1}(\ell_t) - D^\kappa > 0)$ is at least weakly increasing in ℓ_t, combined with Lemma 4, the change in probability must be at least weakly positive. So, debris removal must reduce the probability of Kessler Syndrome:

$$\left.\frac{\partial Pr_t(D_{t+1}(\ell_t) - D^\kappa > 0|S_t, \tilde{D} - R_t)}{\partial \tilde{D}}\right|_{\tilde{D}=D_t} \geq 0 \tag{3.71}$$

$$\implies \frac{dPr_t(D_{t+1}(\ell_t) - D^\kappa > 0|S_t, D_t - R_t)}{dR_t} \leq 0. \tag{3.72}$$

□

Overall, debris removal technologies financed by satellite owners will make orbits safer, reducing both equilibrium collision risk and the risk of Kessler Syndrome. These gains in safety come from satellite owners financing the debris removal. Subsidized or publicly provided debris removal cannot reduce the equilibrium collision risk. Though it may reduce the equilibrium debris stock, subsidized or publicly provided debris removal may not reduce the equilibrium risk of Kessler Syndrome unless the agency providing debris removal commits to preventing the debris stock from exceeding a fixed level. Cooperative removal was shown to achieve this in Proposition 22. The value of debris removal technologies depends critically on the economic institutions under which the technologies are used.

3.2.2.5 Optimal removal and launch plans

Finally, I consider the jointly-optimal debris removal and satellite launch plans to compare with the cooperative removal and open access launch plans. The main point of this section is to show that even with cooperative debris removal financed by satellite owners, open access launching is not socially optimal. Proposition 21 establishes that a constrained planner cannot improve on the cooperative removal plan given open access, so this section is an exercise in determining how large a distortion open access launching creates. The unconstrained fleet planner coordinates removals and launches, taking advantage of the fact that they will be able to remove any unwanted debris before the collision rate is drawn. Their problem at the start of period t is

$$\tilde{W}(S_t, D_t) = \max_{R_t \in [0, D_t]} \{-c_t R_t + \tilde{E}_t[W(S_t, D_t - R_t, \ell_t)]\} \tag{3.73}$$

$$\begin{aligned}
\text{s.t. } W(S_t, D_t - R_t, \ell_t) &= \max_{X_t \geq 0} \{\pi S_t - F X_t + \beta \tilde{W}(S_{t+1}, D_{t+1})\} \\
\ell_t &\sim \phi(\ell_t | S_t, D_t - R_t) \\
S_{t+1} &= S_t(1 - \ell_t) + X_t \\
D_{t+1} &= (D_t - R_t)(1 - \delta) + G(S_t, D_t - R_t, \ell_t) + m X_t.
\end{aligned}$$

The planner faces the same timing of information as firms do: at the beginning of a period, before ℓ_t has been revealed, they choose how much debris they will remove. Based on their removal decision, the draw of ℓ_t is revealed. Then they decide how much they will launch. The program in system of equations 3.73 shows this decision-making process at the beginning of a period. Their jointly-optimal removal and launch plans must equate the social marginal costs and benefits of removing debris before ℓ_t is known and of launching satellites once ℓ_t is known. Formally,

$$R_t^* : c_t = -\left\{ \tilde{E}_t \left[\frac{\partial W(S_t, D_t - R_t, \ell_t)}{\partial D_t} \right] + \frac{\partial \tilde{E}_t [W(S_t, D_t - R_t, \ell_t | S_t, D - R_t)]}{\partial D} \bigg|_{D = D_t} \right\} \tag{3.74}$$

$$X_t^* : \frac{F}{\beta} = \frac{\partial \tilde{W}(S_{t+1}, D_{t+1})}{\partial S_{t+1}} + m \frac{\partial \tilde{W}(S_{t+1}, D_{t+1})}{\partial D_{t+1}} \tag{3.75}$$

An optimal removal plan exists if the sum of the objects inside the curly brackets on the right side of equation 3.74 is positive. I assume that the marginal post-removal value of debris is negative ($\frac{\partial W(S_t, D_t - R_t, \ell_t)}{\partial D_t} < 0$) along optimal paths, making both terms on the right hand side individually negative. The negativity of the second term follows from Lemma 4. This is sufficient to make the right side of equation 3.74 positive.

An optimal launch plan exists if the sum of the objects on the right side of equation 3.75 is positive. I assume that along optimal paths, the marginal pre-removal value from another satellite is positive and the marginal pre-removal value from another piece of launch debris is negative (a loss). I also assume that the pre-removal gain from another satellite is larger than the pre-removal loss from another piece of launch debris ($\frac{\partial \tilde{W}(S_{t+1}, D_{t+1})}{\partial S_{t+1}} \geq 0$, $\frac{\partial \tilde{W}(S_{t+1}, D_{t+1})}{\partial D_{t+1}} < 0$, $\frac{\partial \tilde{W}(S_{t+1}, D_{t+1})}{\partial S_{t+1}} > -m \frac{\partial \tilde{W}(S_{t+1}, D_{t+1})}{\partial D_{t+1}}$).

To compare the qualitative properties of optimal removal and launch plans with cooperative removal and open access launch plans, I simulate cases of the cooperative removal and open access launch plans (system of equations 3.50) and the optimal removal and launch plan (system of equations 3.73). Figure 3.10 shows these plans and the associated value functions.

Comparing Figure 3.10 with Figure 3.3 shows that the open access launch plan with debris removal is similar to the plan without debris removal. When cooperative satellite owners will not remove debris, open access launchers reduce launching as the expected collision rate increases, stopping all launching when the expected collision rate exceeds the excess return on a satellite. When open access launchers anticipate cooperative satellite owners removing debris, they begin launching again at a rate independent

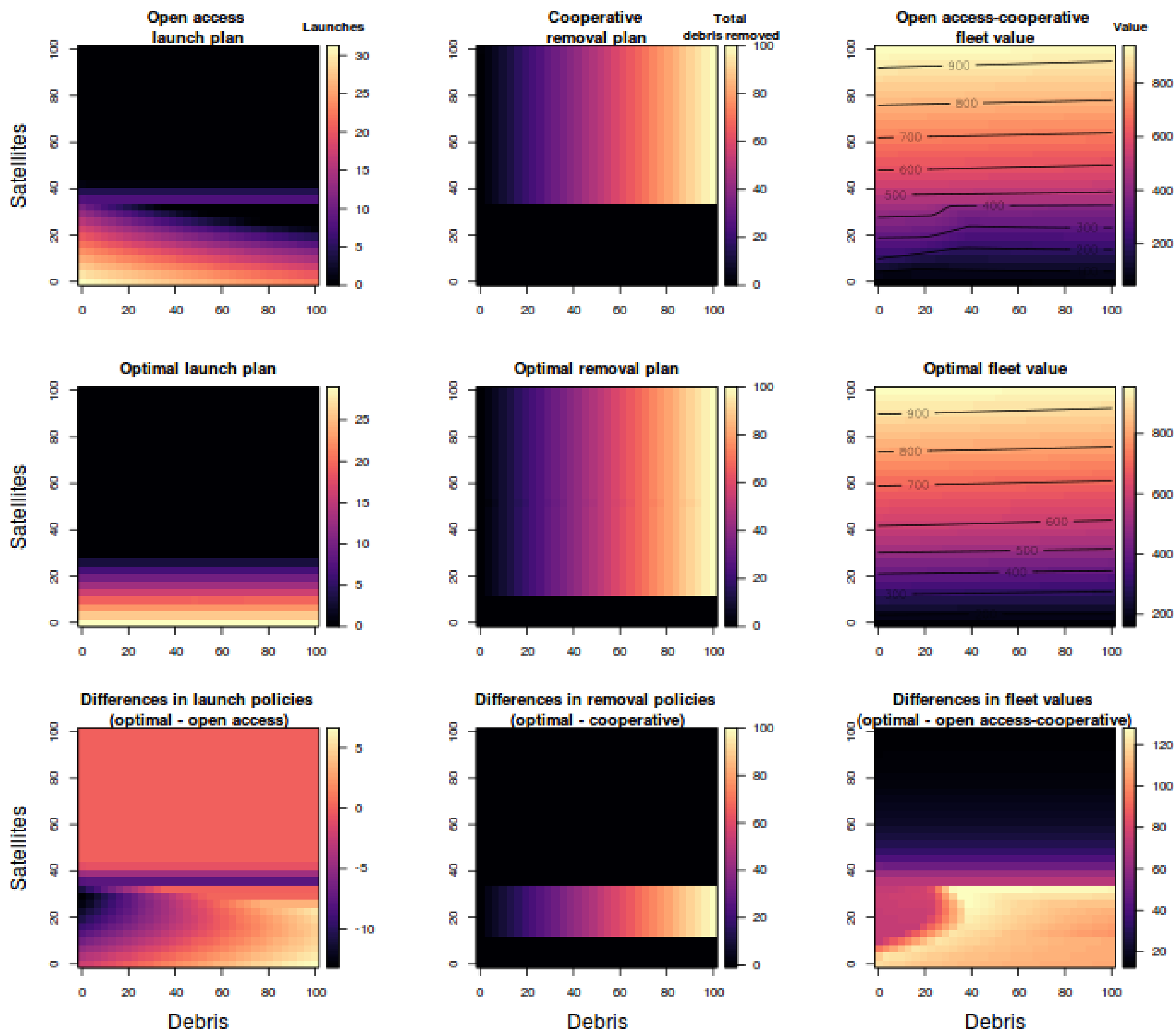

Figure 3.10: *Comparing optimal and open access-cooperative launch and removal plans.*
Upper row: The open access launch plan (left), cooperative removal plan (middle), and resulting fleet value (right). The jump in the launch plan just above 10 reflects open access launches taking advantage of debris removal beginning, as shown in the time paths in Figure 3.9.
Middle row: The optimal launch plan (left), optimal removal plan(middle), and resulting fleet value (right).
Bottom row: The gap between optimal plans/values and open access-cooperative plans/values. The gap between optimal and open access-cooperative fleet values is maximized when (a) the planner would begin removing debris but cooperative satellite owners have not, and (b) just before open access launchers begin to launch again (anticipating removal) and the planner has stopped.

of the amount of debris. This jump is shown in the time paths in Figure 3.9. The jump occurs because debris removal by incumbent satellite owners allows new firms to enter the orbit. Since the planner keeps the debris stock at a constant level as soon as the fleet value justifies it, they ignore debris while launching. The net effect is that with debris removal technologies available, open access firms may launch too many or too few satellites relative to the planner. Under open access, firms launch too many satellites when the expected collision rate is low because they don't internalize the marginal external cost of their satellites, and too few as the expected collision rate increases because they don't account for debris removal. When open access launchers anticipate debris removal, they once again launch too many satellites because incumbent satellite owners can't exclude the launchers from taking advantage of the newly-cleaned orbits. Eventually, the profits of owning a satellite net of the cost of debris removal is no longer sufficient to justify launching.

The cooperative debris removal plan and the planner's removal plan are both corner solutions once debris removal starts.[21] The planner, however, begins debris removal with fewer satellites than the cooperative firms. Intuitively, the planner starts removing debris once the fleet is valuable enough to justify removal, while cooperative satellite owners start removing debris once there are enough owners sharing the removal costs to justify removal.

The discontinuity in the open access launch plan, its dependence on the debris stock, and the later-than-optimal start to cooperative debris removal all reduce the value of the open access-cooperative fleet relative to the planner's. The value loss from open access launching and cooperative debris removal follows the launch plan deviation and is intensified along the removal plan deviation. The gap is maximized just before open access launchers, anticipating removal, begin to launch again. At that point, the planner would have stopped launching and have begun removing debris while cooperative satellite owners would still be waiting for more contributors.

[21] See Appendix section D.8 for more details on nonconvexities and corner solutions in debris removal. The functional form in equation 3.4 implies that the marginal benefit of debris removal is constant with respect to debris, so it is optimal to either remove all debris or none.

3.3 Conclusion

In this chapter I showed how principles of economics should guide our stewardship of orbital resources. I established the equivalence of price and quantity instruments for orbital management and showed why space traffic controls should target satellite ownership rather than satellite launches. I considered the impacts of using active debris removal technology, and showed why, to reduce equilibrium collision risk, satellite owners must pay for debris removal. While satellite-focused policies can achieve first-best orbit use, attempts to control orbital debris growth and collision risk through launch fees or debris removal subsidies under open access may be ineffective or backfire.

Along the way I derived practically-useful results about orbit use management under physical uncertainty with and without active debris removal. These include how to use stock and flow space traffic controls, the fact that debris removal can induce more launches no matter how it is financed, and the possibility that the open access launch rate may be increasing in the launch cost. I also examined the effects of indirect orbit control through spectrum regulation or mandatory satellite insurance. These policies approximate stock controls and are potential avenues by which regulators can induce first-best orbit use.

Knowing these details will help regulators manage orbit use effectively. However, questions will only grow as humans develop a larger presence in space. Commercial satellite operators are increasingly using many small satellites arrayed in constellations to deliver services. How should satellite constellations be regulated? International agreements will be necessary to regulate orbit use and minimize leakages, but different nations have different interests. What kinds of international orbit use management agreements are incentive-compatible? Militaries are among the most prominent orbit users, and have objectives which may conflict with commercial operators or each other. How can strategic orbit use by militaries be efficiently managed without compromising national and international security objectives? These are all important directions for future research.

Satellites are important to the modern world. We depend on satellite telecommunications to reach remote parts of the planet, enabling telemedicine and timely rescue efforts. We depend on GPS for navigation, and will rely on it more as automated transportation infrastructure develops. We depend on satellite imagery to determine the extent of natural disasters and optimize our responses to them. A satellite-based orbit use management regime can prevent the tragedy of the orbital commons, preserving existing space assets and enabling new capabilities in the future.

Chapter 4

External Cost in Space: Optimally Pricing Low-Earth Orbit Use

Written with Matt Burgess and Dan Kaffine

Open access to Earth's orbit has serious negative externalities, and may collapse the resource entirely for generations. We know from Chapter 3 that the optimal policy in the absence of debris removal technologies is an orbital stock control, e.g. a satellite tax. But questions of magnitude remain open. What is the right price of an orbit? If society implements the right price, how much will satellite safety and global social welfare from orbit use increase? In this chapter, we calculate the path of an optimal LEO satellite tax and the improvement in LEO satellite safety and discounted net social welfare from implementing optimal LEO management. The tax is levied once per year on each satellite in orbit, and can be viewed as a yearly rental fee for orbit use. Methodologically, our contribution includes a class of simulation models for orbit management which are conceptually related to integrated assessment models for climate change analysis and the congestion measurement literature (Walters, 1961; Mayeres et al., 1996; Kelly and Kolstad, 1999; Qingyu et al., 2007; Hughes and Kaffine, 2017).

We wish to highlight three main results. First, shifting to optimal orbit management in 2020 would increase the net present value (NPV) of the satellite industry by approximately \$1.75 trillion in 2020, and by over \$4 trillion in 2040. Second, these gains can be achieved by imposing an annual tax on orbiting satellites, starting at around \$37,500 per satellite per year in 2020 and escalating to around \$1 million per satellite per year in 2040. Third, the cost of inaction in 2040, measured as the amount of long-run economic benefit forgone in 2040 due to suboptimal management under BAU relative to optimal management beginning in 2015, escalates from $900 billion if optimal management begins in 2020 to $4.6 trillion if optimal management begins in 2035.

We generate the path of an optimal satellite tax in three steps. First, we calibrate functions describing the physics and economics of orbit use to match observed data on satellite and debris stock levels and aggregate satellite industry costs and returns prior to 2018. Then, using the calibrated values, we generate open access and optimal launch paths from 2006 to 2040. Finally, by comparing the open access path of collision risk to the optimal path of collision risk, we calculate the path of the optimal satellite tax which induces open access satellite owners to internalize the externality they impose on other orbit users.

We obtain physical functions relating launches, satellites, and debris stocks to collisions, new fragments, and satellite and debris growth from the engineering literature, and economic functions relating the decision to launch to collision risk, costs, and returns from the economics literature. To calibrate the physical functions, we estimate the unknown parameters from satellite stocks, debris stocks, and launches observed over 1957-2017. We constrain the parameters to comply with theoretical restrictions imposed by the engineering model. To calibrate the economic functions, we estimate the unknown parameters from satellite stocks, debris stocks, launches, aggregate satellite industry costs, and aggregate satellite industry returns over 2005-2015. To allow the estimation process to adjust for unobserved launch market frictions, we do not constrain these parameter estimates.

In addition to the limitations imposed by modeling spatially and temporally heterogeneous physical and economic processes at an aggregated level, there are three main analytical limitations pertaining to unobservables in the past and present and unknowables in the future: launch market frictions, constellations, and satellite placement. Our conclusions about the suboptimality of open access to orbit and the necessity of a globally-coordinated satellite tax (or policies equivalent to one) are robust to these limitations. In general, these limitations may to make our satellite tax estimates lower bounds on the true values required to induce optimal orbit use.

Our economic model is founded on the assumption that all agents who want to launch satellites are able to do so with no frictions. In practice, there are factors other than orbital property rights and willingness-to-pay which limit agents' access to orbit, such as limited availability of launch windows and rockets. These factors constrain humanity's ability to launch satellites. To ensure that our simulations do not violate this launch constraint in observed years, we calculate the launch constraint in each observed period as the cumulative maximum number of launches observed so far. The shadow value of the launch constraint is recovered in the economic parameter calibration process, but the individual factors are not identifiable from the data. We then fit a linear time trend to the observed launch constraint, and project it into the future. To the extent that the launch constraint will be relaxed faster than a linear trend would predict, our estimates are economically conservative, i.e. we assume fewer launches than may occur.

Our economic model is also founded on the simplifying assumption of "one satellite per firm". In practice, there are a number of firms which own constellations or fleets of satellites. However, unless a single firm owned all satellites in orbit, orbit users would not internalize the full scope of the externality they impose on others. To the extent that the observed data reflects agents internalizing those externalities due to ownership of multiple satellites, our economic parameter estimates would entangle those factors with the estimated launch constraint shadow value, biasing our estimates toward negative infinity or zero (shown in equation 4.16). Our projections of single-satellite-owning firms' responses to increases in satellite profitability would therefore also be attenuated toward zero, making our projections environmentally conservative, i.e. closer to an environmental "worst-case" analysis.

Lastly, our model abstracts entirely away from the question of satellite placement. That is, two orbital objects within a given volume shell can be placed in orbits such that at one extreme they are guaranteed to collide, or at the other extreme they will never collide. Our projections are based on collision rate estimates which are calculated using historical placement patterns. Thus, our projections assume that the systematic factors which resulted in current object placements will continue into the future.

While technology and constellation ownership are likely to lead to improvements in placement patterns our collision risk projections would be biased for both the open access and optimal launch paths. However, the magnitude of the gap between open access and optimal collision risk may actually be understated by this issue. To the extent that economic agents have the placement margin available to them it is induces another externality, similar in spirit but different in detail to the orbit use externality we describe in this article, wherein firms do not account for the full magnitude of orbital use efficiency losses due to their placement. A fleet planner who coordinated all satellites in orbit would account for such placement-related externalities. By taking advantage of any efficiencies in placement, the planner would be able to reduce collision rates below what open access satellite owners would have an incentive to consider. Thus, while the inclusion of a placement margin may reduce levels of collision risk, the differences in collision risk between open access and optimal use will likely increase, making our satellite tax estimates a lower bound on average[1].

4.1 Data

4.1.1 Satellite and debris counts, 1958—2015

We use data on satellites in orbit from the Space-Track dataset hosted by the Combined Space Operations Center (CSpOC) (Combined Space Operations Center, 2018) to construct the satellite stock and launch rate series. The Space-Track dataset provides details on active payloads in LEO and their decay dates. The data are described in Table 4.1.

We construct the numbers of active satellites in each year by calculating the number of objects launched a particular year, adding the number of satellites previously calculated in orbit, and then subtracting the number of satellites listed as having decayed in that year.[2]

[1]While some regimes in a spatially-differentiated orbit model may have lower tax values than the ones we calculate here, the average tax value across all regimes will likely be larger.

[2]This procedure is likely to produce an upward-biased estimate of the returns-generating satellite stock in any given year, since satellites which are no longer operational will not be removed from the estimated stock until they have deorbited. Thus, the satellite stock in this procedure includes some objects which are, economically speaking, “socially-useless debris”. We use this procedure despite the attribution issue for two reasons. First, we do not have data on when specific satellites were declared nonoperational by

Letting ℓ_t be the number of collisions observed and Z_t be the number of payloads listed as decayed, we construct the launch rate from the satellite stock series as

$$X_t = S_{t+1} - S_t + Z_t + \ell_t, \tag{4.1}$$

where S_t is the number of active payloads in year t and Z_t is the number of payloads listed as decayed in year t.

The debris and collision risk series' we use were provided by the European Space Agency. We use debris data from the DISCOS database (European Space Agency, 2018) and collision probability data used in (Letizia et al., 2018) (the variable p_c in that paper). We use only objects with a semi-major axis of 2000km or less in all our data series. We prefer to use the DISCOS fragment data rather than the Space-Track fragment data as it tracks fragments from the time they were created or detected, whereas the Space-Track data tracks fragments from the time their parent body was launched. The DISCOS attribution method is closer to how economic agents in our model receive information and make decisions.

4.1.2 Aggregate satellite industry returns and costs, 2006—2015

We use data on satellite industry revenues from Wienzierl (2018), and data on satellites in LEO (semi-major axis less than 2000km) from the Union of Concerned Scientists' list of active satellites (Union of Concerned Scientists, 2018). The economic data provide a breakdown of revenues across satellite manufacture, launch, insurance, and products and services. The satellite industry revenues data cover

their owners. Such a determination can be particularly tricky when a mission has ended, but the satellite still has fuel and could be repurposed for another mission. Second, to the extent that our estimates of the satellite stock are biased upward (toward positive infinity), our physical and economic parameters estimates will be biased downwards. The downward bias in economic parameters will deflate both the open access and socially optimal launch rates, while the downward bias in physical parameters will inflate both the open access and socially optimal launch rates, with the net effect being difficult to determine. However, the downward bias in our estimated collision risk coefficients and the upward bias in our estimated satellite stock will bias our estimated satellite tax downward, so that it is a lower bound.

Table 4.1: LEO satellite and debris summary statistics

Statistic	Mean	St. Dev.	Min	Max
Satellites	1,057.897	720.298	4	2,271
Launch successes	89.931	47.137	5	334
Payloads decayed	50.810	24.166	3	90
Launch failures	13.845	9.817	1	42
ASAT tests	0.190	0.545	0	3
Debris	3,375.586	2,681.324	0	9,662
Collision rate	2.598	1.800	0	6

2006-2015, while the active satellites data cover 1958-2017. The historical data are described in Table 4.2.

We calculate the per-period returns on owning a satellite (π_t) as the revenues generated from commercial space products and services, and the per-period costs of launching a satellite (F_t) as the sum of revenues from commercial infrastructure and support industries, ground stations and equipment, commercial satellite manufacturing, and commercial satellite launching. The ratio π_t/F_t then gives a time series of the rate of return on a single satellite, as the number of satellites cancels out of the numerator and denominator. Since the numbers provided in Wienzierl (2018) are for the satellite industry as a whole, the ratio still needs to be adjusted to represent satellites in LEO. We do not explicitly conduct this adjustment, but let the adjustment be calculated during the estimation of equation 4.7.[3]

Table 4.2: Satellite industry historical aggregate revenue and cost summary statistics, 2006–2015

Statistic	Mean	St. Dev.	Min	Max
Total industry revenues	101.774	20.352	70.440	126.551
Total industry costs	185.376	40.050	136.162	254.387

[3]Another way to perform this adjustment is by calculating the yearly share of satellites in LEO and multiplying the ratio π_t/F_t by the share in LEO. This approach is difficult to generalize to future years since it requires projections of satellites in other orbits. It is also not clear that the returns of satellites in LEO are truly proportional to the LEO share of the total number of active satellites in all orbits.

To generate launch rate and satellite tax projections out to 2040, we revenue and cost projections from (Jonas et al., 2017). These projections are described in Table 4.3. Figure 4.1 shows the historical and projected data.

Table 4.3: Satellite industry projected aggregate revenue and cost summary statistics, 2016–2040

Statistic	Mean	St. Dev.	Min	Max
Total industry revenues	217.898	76.448	120.786	366.555
Total industry costs	362.207	86.974	210.674	525.072

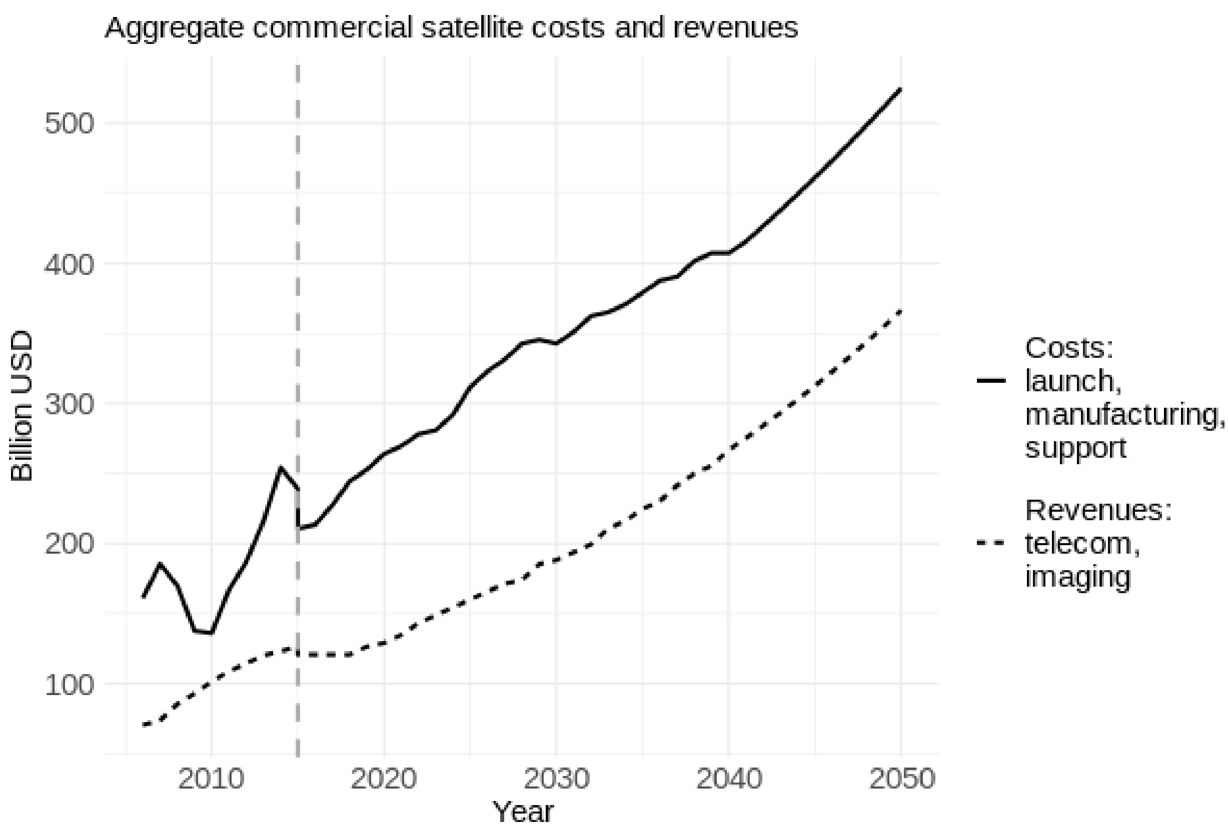

Figure 4.1: *Satellite industry revenues and costs.*

Historical and projected satellite industry revenues and costs. The vertical gray dashed line indicates where we begin to use projected data (2015).

4.2 Models

4.2.1 Orbital mechanics with limited lifespans, missile tests, and certainty

Our physical model uses accounting relationships in the aggregate stocks of satellites and debris for the laws of motion, and draws on (Letizia et al., 2017) for the functional forms of the new fragment creation and collision probability functions $G(S,D)$ and $L(S,D)$. The time period scale is set as one calendar year to match our data. S_t denotes the number of active satellites in an orbital shell in period t, D_t the number of debris objects in the shell in t, X_t the number of satellites launched in t, $L(S_t,D_t)$ the probability that an active satellite in the shell will be destroyed in a collision in t, μ is the fraction of satellites which do not deorbit in t, and m is the average amount of debris generated by deorbiting satellites. δ is the average proportion of debris objects which deorbit in t, and $G(S_t,D_t)$ is the number of new debris fragments generated due to all collisions between satellites and debris.[4] m is the number of debris pieces contributed by satellites launched. A_t is the number of anti-satellite missile tests conducted in t, and γ is the average number of fragments created by one test. We assume that satellites which deorbit do so without creating any additional debris.

The number of active satellites in orbit is modeled as the number of launches in the previous period plus the number of satellites which survived the previous period. The amount of debris in orbit is the amount from the previous period which did not decay, plus the number of new fragments created in collisions, plus the amount of debris in the shell created by new launches. Formally,

$$S_{t+1} = S_t(1-L(S_t,D_t))\mu + X_t \tag{4.2}$$

$$D_{t+1} = D_t(1-\delta) + G(S_t,D_t) + \gamma A_t + mX_t. \tag{4.3}$$

[4]For most of our sample, the number of observed collisions is zero. We use the probability of collisions in our models rather than the observed number for two reasons. First, it proxies for unobserved collisions, including non-catastrophic ones. Second, a model with stochastic collisions complicates the process of solving for the optimal time path by adding another state variable to the dynamic programming algorithm. As the number of objects in a single period increases, the fraction of satellites destroyed in collisions in that period converges to the probability of destruction, so this assumption provides a "mean field"-type approximation.

Letizia et al. (2017) use an analogy to kinetic gas theory to parameterize the probability of a collision as a negative exponential function, with the density of colliding objects one of the arguments of the exponential function. We therefore parameterize $L(S_t, D_t)$ as

$$L(S_t, D_t) = 1 - \exp(-\alpha_{SS} S_t - \alpha_{SD} D_t), \tag{4.4}$$

where α_{SS} and α_{SD} include the difference in velocities between the objects colliding, the total cross-sectional area of the collision, and scaling parameters which relate the number of objects to their density in the volume. We use these probability functional forms to parameterize $G(S_t, D_t)$ as

$$G(S_t, D_t) = \beta_{SS}(1 - \exp(-\alpha_{SS} S_t)) S_t + \beta_{SD}(1 - \exp(-\alpha_{SD} D_t)) S_t, \tag{4.5}$$

where the β_{jk} parameters are interpreted as "effective" numbers of fragments from collisions between objects of type j and k.[5] We refer to the α_{jk} and β_{jk} as "structural physics parameters".

We ignore the possibility of collisions between debris objects for two reasons. First, the data we have do not allow us to identify the effective number of fragments from such collisions, or the probability of such collisions, using our calibration approach. Second, our focus here is not on the probability of Kessler Syndrome, but on general launch patterns and their response to the extant stock of orbiting satellites and debris. Incorporating the possibility of Kessler Syndrome is an important piece of optimal orbit use analysis and policy design, and will likely require higher-fidelity physical modeling than the "aggregate calibration" approach we take here. This is an important area for future research.

Equations 4.3, 4.4, and 4.5 can be viewed as reduced-form statistical models which recreate the results of higher-fidelity physics models of debris growth and the collision probability. While higher-fidelity physics models may use the same functional forms, the key difference between our approach and the approach in such models is how we calibrate the models: rather than derive the appropriate parameter values from

[5]"Effective" numbers of fragments measure the number of new fragments weighted by the time they spend inside the volume of interest.

physical first principles, we estimate the values of those parameters which maximize the fit between model-predicted collision probabilities and debris stocks and the data.

4.2.2 Open access orbit use with time-varying aggregate returns and costs

The economic model of open access here is based on Chapter 2 to determine the satellite launch rate under open access, X_t, as a function of the collision probability, $L(S_{t+1}, D_{t+1})$, and the excess return on a satellite, $r_s - r$. In the simplest case, where all of the economic parameters are constant over time, the open access launch rate equates the collision probability with the excess return:

$$L(S_{t+1}, D_{t+1}) = \underbrace{r_s - r}_{\substack{\text{excess return} \\ \text{on a satellite}}}, \tag{4.6}$$

where r_s is the per-period rate of return on a single satellite (π/F, where π is the per-period return generated by a satellite and F is the cost of launching a satellite, inclusive of non-launch expenditures such as satellite manufacturing and ground stations) and r is the risk-free interest rate.[6]

Equation 4.6 can therefore also be used to calculate the implied IRR for satellite investments from observed data on collision risk and satellite returns. r is not observed in our data. When costs and returns are time-varying, equation 4.6 becomes

$$\begin{aligned} L(S_{t+1}, D_{t+1}) &= 1 + r_{s,t+1} - (1+r)\frac{F_t}{F_{t+1}} \\ \implies L(S_{t+1}, D_{t+1}) &= \underbrace{\left(r_{s,t+1} - r\frac{F_t}{F_{t+1}}\right)}_{\substack{\text{excess return} \\ \text{on a satellite}}} + \underbrace{\left(1 - \frac{F_t}{F_{t+1}}\right)}_{\substack{\text{capital gains} \\ \text{from open access} \\ \text{and satellite launch} \\ \text{cost variation}}} \end{aligned} \tag{4.7}$$

where $r_{s,t+1} = \pi_{t+1}/F_{t+1}$. With time-varying economic parameters, two sources of returns drive the collision risk. One is the excess return realized in $t+1$ from launching a satellite in t. The other is the

[6] More precisely, r is the opportunity cost of funds invested in launching a satellite, and may diverge from the risk-free rate if the satellite launcher's most-preferred alternate investment is not a risk-free security. This rate is sometimes referred to as the internal rate of return (IRR).

capital gain (or loss) due to open access and the change in satellite costs. Since open access drives the value of a satellite down to the total cost of launching and operating it, F_t becomes the cost of receiving F_{t+1} in present value the following period, and the returns are given as percentages of F_{t+1}. Since the discount rate is unobserved, we fix it to be constant over time to facilitate estimation.[7]

The maximum number of satellites which can be launched in a year are limited by a variety of factors, including weather, availability of rockets, and availability of launch sites. We estimate this "launch constraint" from the observed data for the historical period, and extrapolate it forward for the projection period. We describe this procedure in more detail in section G.5 of the Appendix.

4.2.3 Optimal orbit use with time-varying aggregate returns and costs

The fleet planner's problem is more complicated. The fleet planner launches satellites to maximize the value of the entire fleet into the (discounted) infinite future, subject to the laws of motion of satellite and debris stocks. Formally, letting $\beta = (1+r)^{-1}$ be the discount factor, the planner solves

$$W(S,D) = \max_{X\geq 0}\{\pi S - FX + \beta W(S',D')\} \tag{4.8}$$

$$S' = S(1-L(S,D))\mu + X$$

$$D' = D(1-\delta) + G(S,D) + \gamma A + mX.$$

We drop time subscripts and use primes on a variable's right to indicate future values, in keeping with the convention for infinite-horizon dynamic programming problems. The economic parameters π and F are allowed to be time-varying in our solution approach, though all other physical and economic parameters are constant over time.

Solving the planner's problem by taking the first-order condition and applying the envelope condition

[7] This equation was derived earlier in the Appendix of Chapter 2, in section A.3. In that setting the discount rate was not constant over time.

to recover the unknown functional derivatives yields the following relation between present and future values of debris:

$$W_D(S,D) = -SL_D(S,D)F + \beta[1-\delta+G_D(S,D)+mSL_D(S,D)]W_D(S',D'), \tag{4.9}$$

where $W_D(S,D)$, the marginal value of another unit of debris on the satellite fleet given the levels of S and D, is given by

$$W_D(S,D) = \left[\frac{'F}{\beta} - \pi - (1-L(S,D)-SL_S(S,D))F - \frac{G_S(S,D)-m(1-L(S,D)-SL_S(S,D))}{1-\delta+G_D(S,D)+mL_D(S,D)}L_D(S,D)SF\right] \cdot \left[\frac{G_S(S,D)-m(1-L(S,D)-SL_S(S,D))}{1-\delta+G_D(S,D)+mL_D(S,D)}+m\right]^{-1}, \tag{4.10}$$

subscripts indicate derivatives with respect to the subscripted variable, and $'F$ indicates the previous period's launch cost. A version of this equation is derived in the appendix.

4.3 Calibration and simulation

4.3.1 Physical parameters: deorbit, decay, collisions, and fragments

We calibrate the rate at which satellites deorbit, μ, by estimating the following analog to equation 4.1 by OLS:

$$S_{t+1} = S_t(1-L(S_t,D_t))\hat{\mu} + X_t. \tag{4.11}$$

This yields an estimated average operational lifespan of about 30 years, i.e. $\hat{\mu} = 0.967$. This is consistent with an average mission length of 5 years, followed by compliance with the 25-year deorbit guideline issued by the IADC (IADC, 2007). While only four in five LEO operators who launched between 2003 and 2014 are estimated to comply with the guideline, including this rate in our forward simulations is conservative in the sense that the estimated satellite tax becomes a lower bound relative to a model with imperfect compliance.

We calibrate equations 4.4 and 4.3 by estimating the following equations:

$$L(S_t, D_t) = 1 - \exp(-\hat{\alpha}_{SS} S_t - \hat{\alpha}_{SD} D_t) + \varepsilon_{Lt} \tag{4.12}$$

$$D_{t+1} = (1 \hat{-} \delta) D_t + \hat{\beta}_{SS}(1 - \exp(-\hat{\alpha}_{SS} S_t)) S_t + \hat{\beta}_{SD}(1 - \exp(-\hat{\alpha}_{SD} D_t)) S_t + \tag{4.13}$$

$$\hat{\gamma} A_t + \hat{m} X_t + \varepsilon_{Dt}, \tag{4.14}$$

where ε_{xt} are mean-zero error terms to minimize and a_{xi} are parameters to estimate. In theory, the α_{jk}, β_{jk}, and m are nonnegative, and δ is in $(0,1)$. We constrain the parameter estimates to comply with the theoretical restrictions.

We calibrate equations 4.12 and 4.13 in two stages. First, we estimate equation 4.12 by constrained NLS. Then, using the estimated values of $\hat{\alpha}_{SS}$ and $\hat{\alpha}_{SD}$ to generate $(1 - \exp(-\hat{\alpha}_{SS} S_t)) S_t$ and $(1 - \exp(-\hat{\alpha}_{SD} D_t)) S_t$, we estimate equation 4.13 by constrained ridge regression, estimating $(1 - \delta)$ directly.[8] We estimate both equations on the sample from 1957-2013. The fitted values are shown against the actual values with residuals in figure 4.2.

Tables 4.4 and 4.5 show the calibrated parameters for equations 4.12 and 4.13:

Collision probability parameters:	α_{SS}	α_{SD}
Parameter values:	1.29e-06	2.56e-08

Table 4.4: Parameter values from estimating equation 4.12.

Debris law of motion parameters	δ	m	γ	β_{SS}	β_{SD}
Parameter values:	0.49	4.84	144.13	292.72	5026.17

Table 4.5: Parameter values from estimating equation 4.13. All values are rounded to two decimal places. The penalty parameter λ was selected through cross-validation.

[8] We use ridge regression for the debris equation to improve the model fit, despite bias in the estimated parameters.

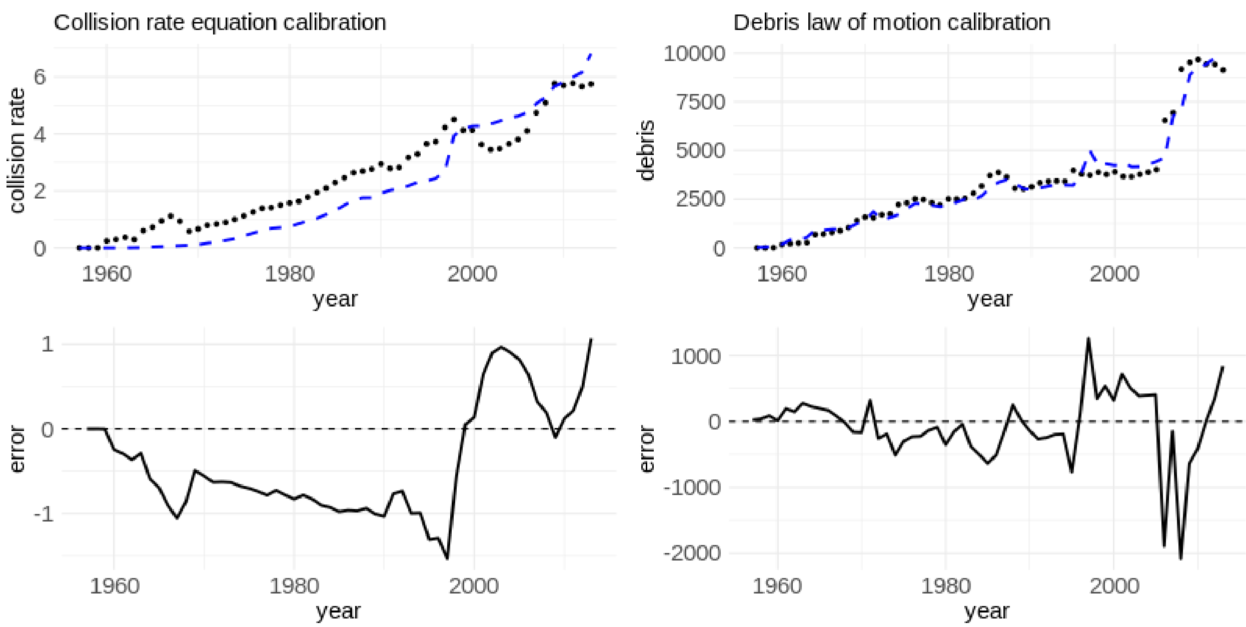

Figure 4.2: *Calibration fit.*
The upper panels show fitted values (blue dashed line) against actual values (black dots). The lower panels show the residuals.

These parameter values are physically plausible, with the values estimated for equation 4.13 being lower bounds.[9] For example, the value of m suggests that every satellite launched creates (at least) 4.84 pieces of debris on average, while the value of γ suggests that anti-satellite missile tests create (at least) 144.13 pieces of debris on average. While higher-fidelity physical models which derive these quantities from first principles will yield more accurate results, the estimated values appear to be a reasonable first-order approximation to the true values based on the model fits (shown in figure 4.2).

4.3.2 Economic parameters: returns, costs, and discounting

Since r is unobserved, we calibrate equation 4.7 by estimating

$$L(S_t, D_t) = a_{L1} + a_{L2} r_{st} + a_{L3} \frac{F_{t-1}}{F_t} + \varepsilon_{rt}, \tag{4.15}$$

using OLS on the sample of returns data from 2005-2015, omitting the first observation (for 2005) to construct F_{t-1}/F_t. ε_{rt} is a mean-zero error term, a_{L2} is a scale parameter, and a_{L3} measures the gross IRR, $1+r$. The fitted values are shown against the actual values with residuals in figure 4.3. Table 4.6 shows the calibrated parameters.

Economic calibration parameters:	a_{L1}	a_{L2}	a_{L3}
Parameter values:	0.004	0.009	-0.0004

Table 4.6: Parameter values from estimating equation 4.15. All values are rounded to the first non-zero digit.

If our data perfectly measured the costs and returns of satellite ownership, and our theoretical model held exactly, we would expect $a_{L1} = 1$, $a_{L2} = 1$, and $a_{L3} < -1$. Our estimates therefore suggest that our returns and cost series are measured with error or that our theoretical model is missing some important factors, such as constraints on the number of launches possible each period. Since constraints on the number of launches in each period are a type of flow control, we can adapt equation 3.26 from Chapter 3 to

[9] Ridge estimates are biased toward zero relative to OLS estimates. For a given penalty parameter $\lambda \geq 0$, the relationship between a ridge coefficient estimate $\hat{\beta}^{\text{ridge}}$ and the corresponding OLS estimate $\hat{\beta}^{\text{OLS}}$ is $\hat{\beta}^{\text{ridge}} = \hat{\beta}^{\text{OLS}}/(1+\lambda)$.

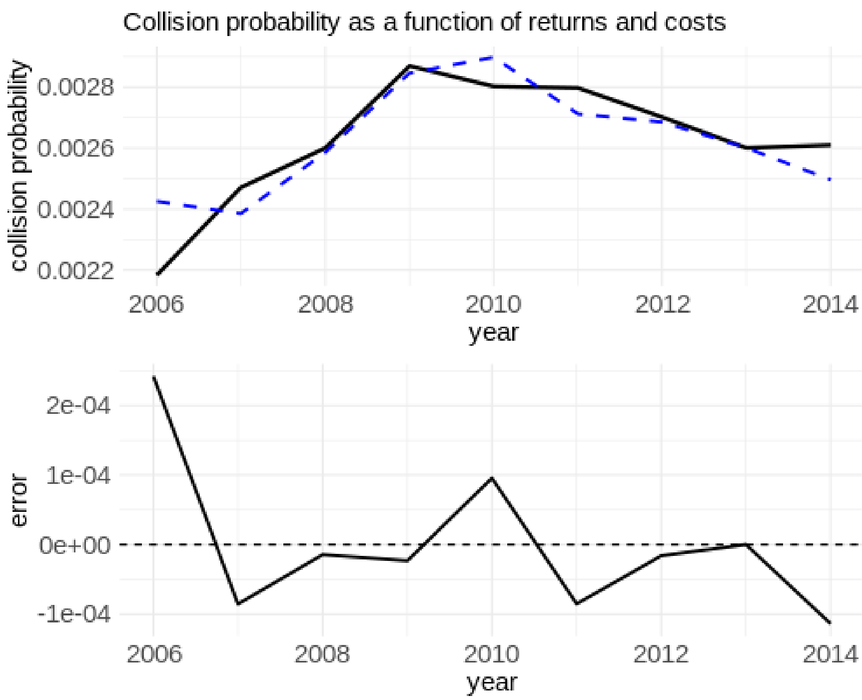

Figure 4.3: *Calibration fit.*
The upper panel shows fitted values (blue dashed line) against actual values (black solid line). The lower panel shows the residuals.

infer the aggregate shadow price of the launch constraint.

The open access model described so far assumes that any firm which wants to launch a satellite can do so. If launches are limited, as they are in practice, this assumption will be violated. The limits will prevent open access launching from equating the excess return on a satellite with the risk of its destruction. In this way, firms which are able to launch earn rents from having a satellite while the collision risk is below the excess return. The wedge between the collision risk and excess return will reflect the value of those rents.

In Chapter 3 we saw that flow controls which restrict the quantity of launches in each period are equivalent to flow controls which impose an additional positive price p_t on launching in t. Generalizing the flow-controlled equilibrium condition in equation 3.26, we obtain the open access equilibrium condition for launches when costs, returns, and discount rates are all time-varying:

$$L(S_{t+1}, D_{t+1}) = (1 - \frac{p_t}{F_{t+1} + p_{t+1}}) + \frac{\pi_{t+1}}{F_{t+1} + p_{t+1}} - (1 + r_t)\frac{F_t}{F_{t+1} + p_{t+1}}, \tag{4.16}$$

p_t can be interpreted in two ways. It can be viewed as the implied rent received by a firm which already owns a satellite in t due to launches in t being restricted. It can also be viewed as the implied launch tax paid by a firm which is allotted a launch slot in t. In either view, a binding launch constraint results in positive values of p_t and p_{t+1}, biasing the coefficients from regression 4.7. a_{L1} is biased toward negative infinity, while a_{L2} and a_{L3} are biased toward zero.

If the data were free of measurement error and an atomistic homogeneous open access model with a constant discount rate r was correct, we would have $a_{L3} = -(1 + r)$. Consistent with typical social discount rates in environmental economics, we set the discount rate to be 5% ($r = 0.05$, implying a discount factor of $\beta = (1 + r)^{-1} \approx 0.95$).

Regardless of the factors missing from the theoretical model, we use equation 4.15 to recursively

Year	Observed return	Observed cost	Implied cost
2006	70.44	161.02	194.26
2007	73.87	185.5	178.88
2008	85.5	170	154.86
2009	93.06	137.81	131.47
2010	101.51	136.16	140.48
2011	108.84	166.99	149.09
2012	114.55	186.88	165.69
2013	120.25	215.9	170.79
2014	123.18	254.39	170.68

Table 4.7: Launch costs implied by open access and observed revenues and costs. All values are given in billions of nominal US dollars per year. 2015 is omitted due to the recursive nature of equation 4.17.

calculate the sequence of launch costs implied by the combination of open access, observed launch rates, and observed satellite returns as

$$L(S_t, D_t) = a_{L1} + a_{L2} r_{st} + a_{L3} \frac{F_{t-1}}{F_t} + \varepsilon_{rt}$$
$$\implies \hat{F}_t = \frac{a_2 \pi_t + a_3 F_{t-1}}{L_t - a_1}. \tag{4.17}$$

Table 4.7 shows the observed satellite returns (π_t), observed launch costs (F_t), and implied launch costs ($\hat{F}_t$).

4.3.3 Algorithms for open access and optimal policy functions

We generate two sequences of policy functions: one function for each period under consideration, and one sequence for each management regime type. We compute each sequence through backwards induction: beginning at the final period in our projection horizon, and iteratively working backwards to the initial period. This procedure implies "perfect foresight" planning under each management regime, i.e. that all agents under any management regime are able to perfectly forecast the sequence of returns, costs, interest rates, launch rates, and other model objects. The perfect foresight assumption is clearly unrealistic, but our purpose is not to show how uncertainty in economic parameters propagates over time. Rather, our purpose is to show how an optimal satellite tax would vary over time and the time paths of orbital aggregate stocks under different management regimes. Such assumptions are used in integrated assessment models of climate change with similar rationales, e.g. the models studied in Kelly and Kolstad (1999); Nordhaus (2013); Wilkerson et al. (2015). Our work here is conceptually similar to integrated assessment modeling.

To compute the open access time path, we first generate a grid of satellite and debris levels, $(grid_S, grid_D)$. We generate this grid as an expanded Chebyshev grid to reduce numerical errors from interpolation, provide higher fidelity near boundaries, and economize on overall computation time. In contrast to a standard Chebyshev grid, an expanded Chebyshev grid allows for computation (rather than extrapolation) at the boundary points. The formula for the k^{th} expanded Chebyshev node on an interval $[a,b]$ with n points is

$$x_k = \frac{1}{2}(a+b) + \frac{1}{2}(b-a)\sec\left(\frac{\pi}{2n}\right)\cos\left(\frac{k}{n} - \frac{1}{2n}\right)$$

We set different values of a and b for S and D, creating a rectangular grid. The main issue in setting b is ensuring that the time paths we solve for (described in section 4.3.4) do not run into or beyond the boundary. To avoid this issue while minimizing the number of points in regions the time paths never reach, we set different a and b bounds for open access and the optimal plan, with the open access grid being strictly larger in both dimensions than the optimal plan grid.

In general, computing decentralized solutions under open access is simpler than computing the planner's solutions. This is because open access simplifies the continuation value to the cost of launching a satellite. We use `R` for all simulations, parallelizing where possible. To facilitate convergence of policy and value functions, we normalize the returns and costs parameters so that $\pi_1 = 1$ during computation, and rescale the parameters after the time paths have been generated.

We compute optimal value functions by value function iteration on a grid of the state variables S and D. We initialize the algorithm with a guess of the value and policy functions. Then, at each point on the grid, we solve the first-order condition for the planner's problem (equation 4.9). Since there may be multiple solutions, only one of which leads to a global maximum, we then evaluate the value function at each solution (including zero) and select the launch rate attached to the largest level of the value function. Algorithm 1 describes how we compute the optimal policy and value functions for a given grid and given value function guess $guess(S,D)$, while algorithm 2 describes how we compute the open access policy and value functions.

We construct our initial guess of the planner's value function as the terminal value of the fleet. In the penultimate period, we assume it is not optimal to launch any satellites ($X^*_{T-1} = 0$), making the final fleet size

$$S_T = S_{T-1}(1 - L(S_T, D_T)).$$

In the final period (T), the payoff of the fleet is πS_T. Our assumption that it is not optimal to launch any satellites in the penultimate period implies that the one-period returns of a satellite do not cover the cost of building and launching ($\beta \pi_T < F_{T-1}$), which we verify to hold in every period of our data. We use the implied series of F_t given the observed π_t and launch rate series in solving for open access and optimal policies.

4.3.4 Projected time paths

We use algorithms 1 and 2 to compute policy and value functions in each period, and run them sequentially from the final period to the first period to generate a series of policy and value functions for each period's set of economic parameters. Algorithm 3 describes this process.

It is important to note that when obtaining the sequence of policy functions we do not do backwards induction within each economic time period prior to the final period. Instead, we hold the continuation value ($W(S_{t+1}, D_{t+1})$) fixed and iterate on the policy functions, using previous iterations' policies as starting points. This ensures that the continuation value incorporates each period's returns and costs only once until the final period, while allowing for any numerical errors in initial policy solves to be corrected. This type of "policy iteration" typically takes 1-2 iterations to converge to within $1e$-3. Backwards induction on the

Algorithm 1: Value function iteration

Set

$$W_0(S,D) = guess(S,D),$$
$$X_0 \equiv 0$$

for all $(S,D) \in (grid_S, grid_D)$

Set $i = 1$ and $\delta = 100$ *(some large initial value)*.

while $\delta > \varepsilon$ **do**

At each grid point in $(grid_S, grid_D)$, use a numerical rootfinder to obtain

$$X_i : W_D(S,D) + SL_D(S,D)\hat{F} - \beta[1 - \delta + G_D(S,D) + mSL_D(S,D)]W_D(S',D') = 0,$$

where $W_D(S,D)$ is given by equation 4.10.

Evaluate $W_i(S,D) = \pi S - \hat{F}X_i + \beta W_{i-1}(S',D')$ at $X_i \cup \{0\}$, and select whichever value of $X_i \cup \{0\}$ maximizes $W_i(S,D)$. $W_{i-1}(S',D')$ is computed by linear interpolation.

$\delta \leftarrow ||W_i(S,D) - W_{i-1}(S,D)||_\infty$.

i $\leftarrow$ i+1

end

Algorithm 2: Open access launch plan computation

Use a numerical rootfinder to find the X^o_{t-1} which solves

$$L(S_t, D_t) = a_{L1} + a_{L2}\hat{r}_{st} + a_{L3}\frac{\hat{F}_{t-1}}{\hat{F}_t},$$

using the estimated laws of motion for S_t and D_t as functions of X_{t-1}, and the estimated function for $L(S_t, D_t)$.

Approximate $W_i^\infty(S,D) = \sum_{t=1}^\infty \beta^{t-1}(\pi S_t - \hat{F}X^o_t)$ as $W_i^T(S,D) = \sum_{t=1}^{T-1} \beta^{t-1}(\pi S_t - \hat{F}X^*_t)$ by backwards induction, using the estimated laws of motion for S_{t+1} and D_{t+1} and the estimated function for $L(S_t, D_t)$. We use $T = 500$.

Algorithm 3: Generating a perfect-foresight sequence of policy functions

Set economic parameters to final period values.

Run algorithm 1 (for an optimal path) or 2 (for an open access path).

for t in T-1:1 do

  Set $i = 1$ and $\delta = 100$ *(some large initial value)*. Set $X_{0t} = X_T$. **while** $\delta > \varepsilon$ **do**

    Using the value function from the previous step as $W(S_{t+1}, D_{t+1})$, calculate

$$\begin{aligned} X_{it}^{*} : & W_{Dt}(S_t, D_t) + S_t L_D(S_t, D_t)\hat{F}_t - \\ & \beta[1 - \delta + G_D(S_t, D_t) + mS_t L_D(S_t, D_t)] W_{Dt}(S_{t+1}, D_{t+1}) = 0, \\ & \text{(for an optimal path, where } W_{Dt}(S, D) \text{ is given by equation 4.10),} \end{aligned}$$

    or

$$X_{it}^{o} : L(S_{t+1}, D_{t+1}) = a_{\ell 1} + a_{\ell 2}\hat{r}_{st} + a_{\ell 3}\frac{\hat{F}_{t-1}}{\hat{F}_t} \quad \text{(for an open access path),}$$

    using the estimated laws of motion for S_{t+1} and D_{t+1}, the estimated function for $L(S_t, D_t)$, linearly interpolating to compute $W_{t+1}(S_{t+1}, D_{t+1})$.

    $\delta \leftarrow ||X_{it} - X_{i-1t}||_{\infty}$

  end

  If calculating an optimal path, set $W_t(S_t, D_t) = \pi_t S - \hat{F}_t X^{*} + W_t(S_{t+1}^{*}, D_{t+1}^{*})$. If calculating an open access path, set $W_t(S_t, D_t) = \pi_t S - \hat{F}_t X^{o} + W_t(S_{t+1}^{o}, D_{t+1}^{o})$.

end

value function in the final period treats that period's costs and returns as steady state values. This is why we change the notation for the fleet value function for algorithm 3, indexing by time to indicate that the launch cost and satellite per-period return are changing in each period.

Once we have a sequence of policy functions for each period's economic parameters, we generate time paths by picking a starting condition (S_0, D_0), computing the launch rate X_0 by thin-plate spline interpolation of the policy function, using the launch rate to compute the next-period state variables, and repeating the process until the terminal period.

Figure 4.4 shows the simulated open access and optimal paths of launches, satellites, debris, and collision risk over an "in-sample" period from 2005-2015. Figure 4.5 shows the same values over the in-sample period as well as projections from 2016-2040.

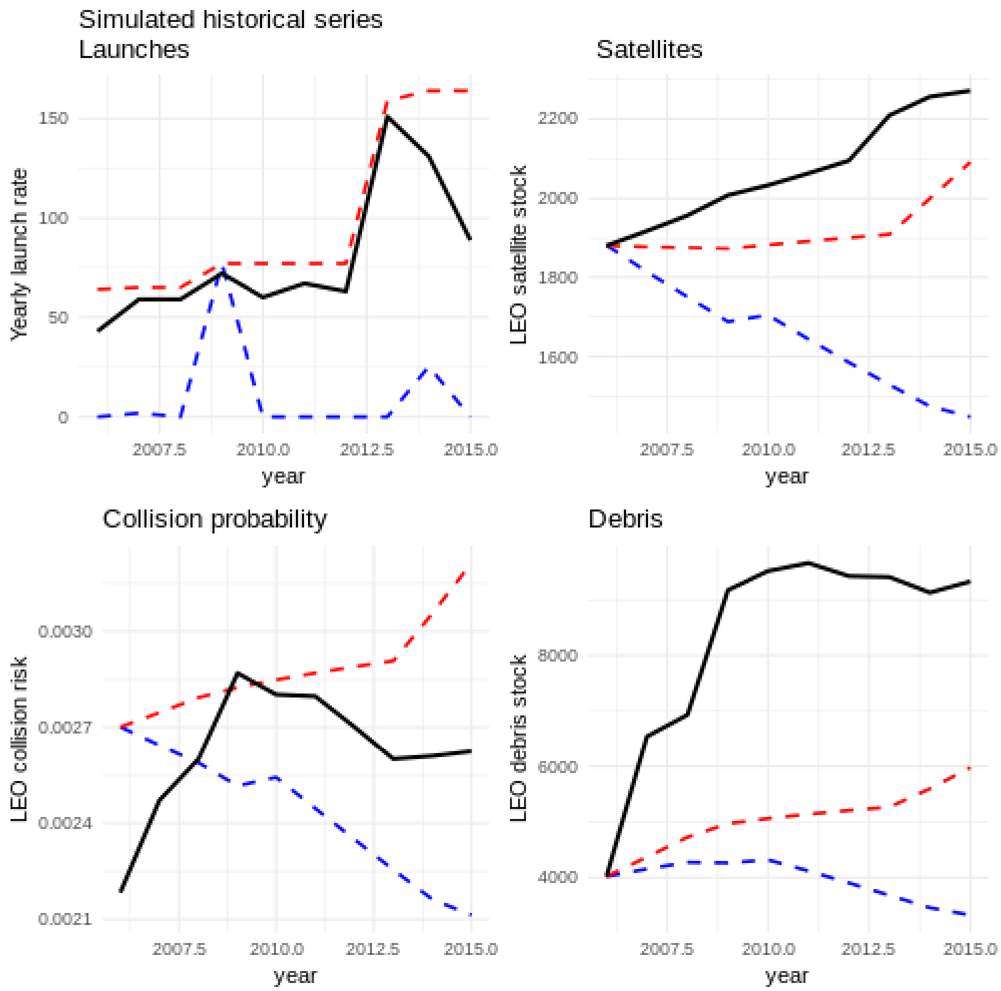

Figure 4.4: *Simulated historical time paths of orbital aggregates.*

The red lines show simulated open access paths. The blue lines show simulated optimal paths. The black lines show observed time paths.

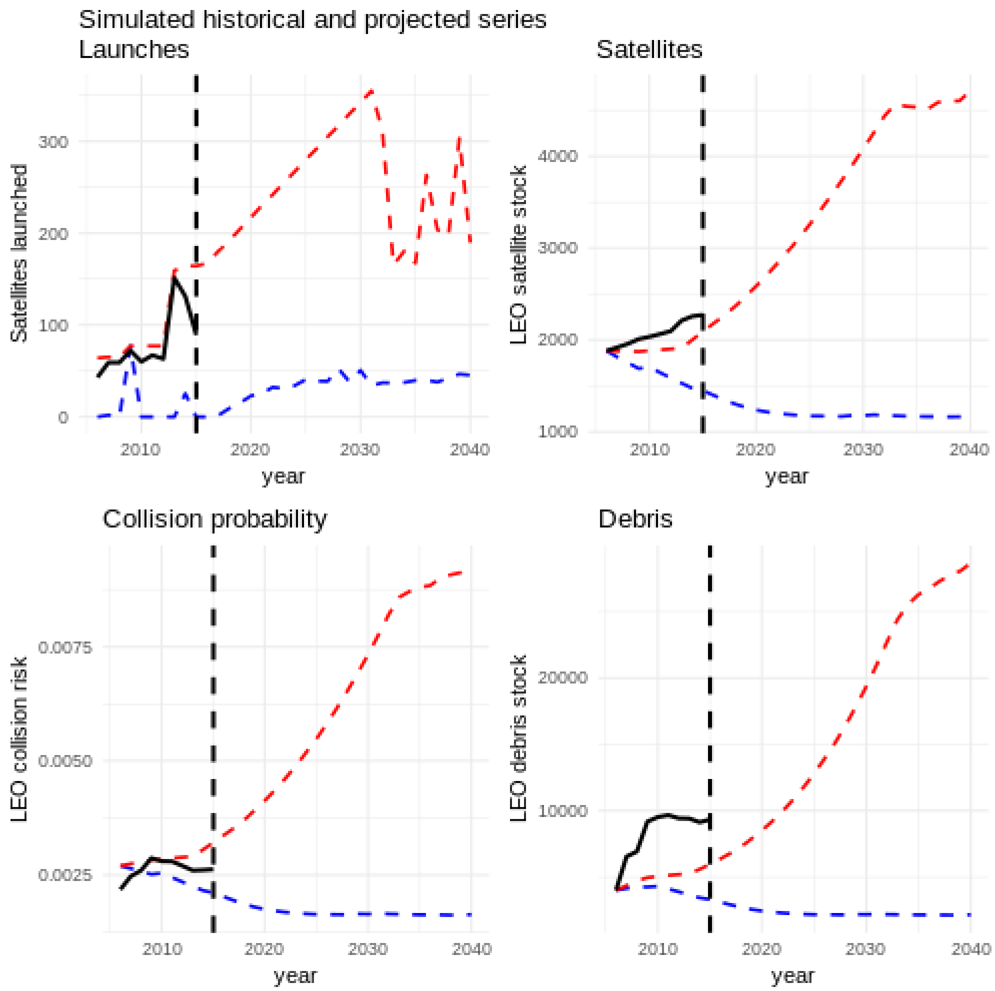

Figure 4.5: *Projected time paths of orbital aggregates.*

The red lines show simulated open access paths. The blue lines show simulated optimal paths. The black lines show observed time paths.

While the historical fit of the open access model is not exact, it is on the right order of magnitude and exhibits the correct trends. Our model is highly simplified and aggregated, and does not include a detailed model of orbital mechanics (necessary to correctly propagate the state variables over time) or a detailed model of launch industry dynamics (necessary to correctly model the idiosyncrasies of the launch path).[10]

[10]For example, on the physical side we abstract from different types of objects in orbit and time variation in their inclinations and relative velocities. On the economic side, we abstract from market power in the launch industry, the different types of commercial industries using satellites, and increasing returns to scale in satellite launch and ownership.

Our focus is on obtaining approximations which are the correct order of magnitude and exhibit the correct time trends. Future research which integrates detailed physical models of orbital mechanics with detailed models of satellite and launch industry dynamics will be required to provide precise values for satellite taxes and welfare loss.

4.4 Optimal satellite taxes

With the calibrated parameter values, we turn to computing satellite tax projections. We split this process into two stages. In the first stage, we compute the time paths of the satellite stock, debris stock, and launch rate, given the open access and fleet planner models of orbit use. These describe the projected evolution of the orbital aggregates. In the second stage, we use the computed time paths with the estimated collision probability function and launch cost path to calculate the optimal satellite tax. The tax is derived from the same open access and fleet planner models. It describes the amount by which a satellite owner would have to be taxed, beginning from the projection horizon's initial conditions, in order to align their incentives with the fleet planner's.[11]

First, we compute open access and optimal launch policy functions in each period using the calibrated parameter values. The policy functions prescribe the number of satellites launched under each type of management regime for given levels of satellite and debris stocks and economic parameters. We compute these policy functions recursively from the final period in our projection to the initial period. We then interpolate between solved points in each period's policy functions to generate launch rate predictions, beginning with the initial condition (observed satellite and debris stocks) in the first projection period and continuing recursively forward until the final projection period. Second, we use a simple formula to calculate the optimal satellite tax from the projected open access and optimal expected value losses due to satellite-destroying collisions. The expected value losses are easy to calculate from the observed or assumed time path of launch costs and the collision probabilities calculated during the launch rate prediction step.

[11]This can also be thought of as "How much of the profits from orbit use currently reflect resource rents which should not have been dissipated?"

We show the in-sample fit of our open access projections to establish that our approach can approximate the observed history, and then use predictions of space economy revenues and costs from (Jonas et al., 2017) to project out the open access and optimal launch rates given those predictions.

We consider two types of taxes: a “jump” tax, and a “stay” tax. The jump tax answers, “What is the right tax to make firms take the most-rapid approach path to a particular management path?”, while the stay tax answers, “What is the right tax to make firms start behaving as a planner faced with the current state would?” While the jump tax can be used to reach a regulator’s desired satellite and debris path as quickly as possible, the stay tax can be used to begin and maintain a program of optimal management from the current orbital state.

For either type, we calculate the time path of an optimal satellite tax from equation 4.18:

$$\tau_t = (L(S^o_{t+1}, D^o_{t+1}) - L(S^*_{t+1}, D^*_{t+1}))F_{t+1}, \tag{4.18}$$

where S^o_{t+1} and D^o_{t+1} are satellite and debris stocks in $t+1$ under open access management, and S^*_{t+1} and D^*_{t+1} are satellite and debris stocks in $t+1$ under optimal management. The optimal tax is positive whenever the planner would maintain a lower collision probability than firms under open access would. The planner, in turn, will maintain a lower collision probability if the lifetime social benefits of another satellite in orbit are less than that satellite’s expected future damages to other satellites in the fleet. By charging open access firms the marginal external cost of their satellite as a tax, their incentives are aligned with those of the planner despite the institutional differences. With their incentives aligned, their decisions to launch or not are shifted to optimize society’s intertemporal economic welfare from orbit use rather than their own individual profit.

Formally, equation 4.18 can be derived by comparing the open access equilibrium condition (equation 4.7) to the fleet planner’s optimality condition for launching (the first-order condition of system of equations

4.8). These conditions can be written to express the expected loss in satellite value (collision probability multiplied by replacement cost) in terms of economic and, in the case of the optimality condition, physical parameters. Those economic parameters include terms for the current excess return on a satellite in addition to the capital gain or loss from changes in the cost of a replacement satellite. By subtracting the optimal expected loss from the open access expected loss, we recover the additional physical and economic term the social planner accounts for — the marginal external cost of a satellite.

4.4.1 The jump tax: jumping to a management path

A jump tax will induce firms to immediately begin targeting a particular management path. For example, supposing a regulator wanted to induce firms to target the 2020 optimal management path (the path of satellites and debris had optimal management begun in 2020), their tax rule $\tau_{t+1}^{j,2020}$ would be

$$\tau_{t+1}^{j,2020} : L(\hat{S}_{t+1}, \hat{D}_{t+1} | \hat{S}_t, \hat{D}_t) = L(S_{t+1}^*, D_{t+1}^* | S_t^{*,2020}, D_t^{*,2020}), \tag{4.19}$$

where $\hat{S}$ and $\hat{D}$ indicate optimal values, and superscripted numbers indicate the year that optimal management began.

The jump tax can be used to target any desired management path. Intuitively, the value of the jump tax is exactly the yearly rent accruing to satellites on the targeted management path, which is dissipated by the business-as-usual path the firms are currently on. But while the jump tax induces firms to take the MRAP to a particular optimal management path, it is not necessarily optimal to do so. The stay tax addresses this concern. On the other hand, computing the jump tax at any point in time does not require any estimates of the launch constraint.

Figure 4.6 shows optimal jump tax paths to target management paths beginning in 2006, 2010, 2015, 2020, 2025, 2030, and 2035.

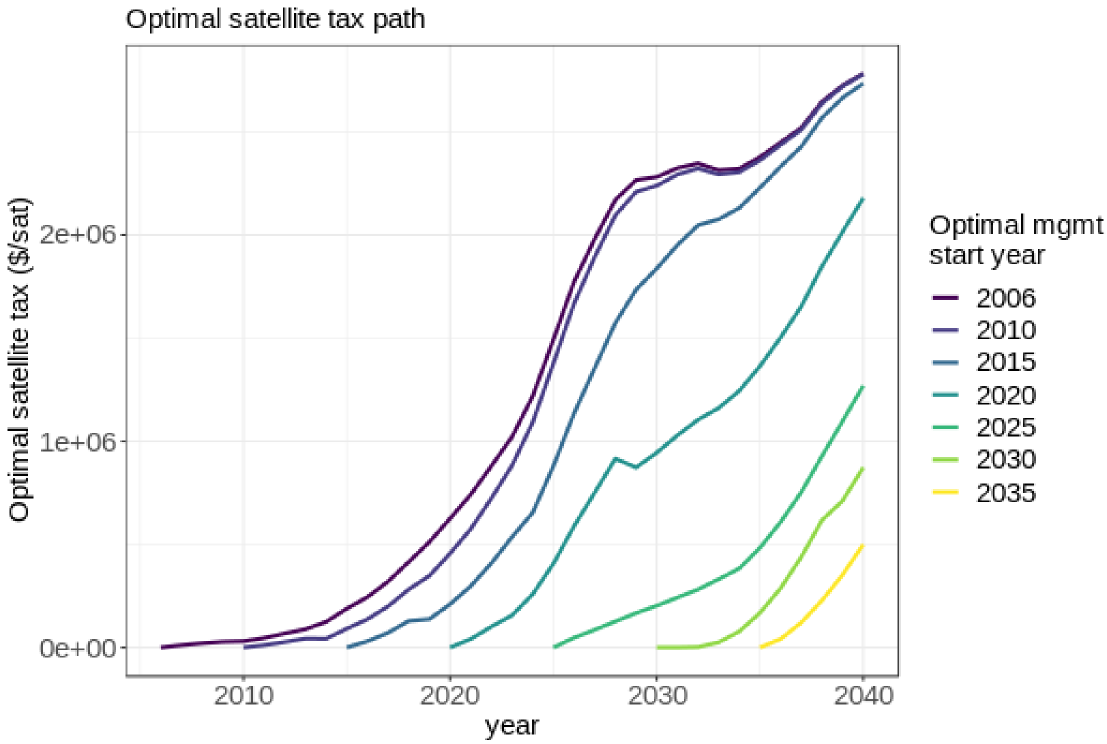

Figure 4.6: *Sample jump tax paths.*

4.4.2 The stay tax: staying on a management path

A stay tax will induce firms to act as a planner faced with the current orbital state would. Another way to think of the stay tax is that it prevents firms from a feasible one-period open access deviation off the 2020 optimal management path. For example, supposing a regulator wanted to induce firms to begin following the optimal management path in period t, their tax rule τ^s_{t+1} would be

$$\tau^s_{t+1} : L(\hat{S}_{t+1}, \hat{D}_{t+1} | S_t, D_t) = L(S^*_{t+1}, D^*_{t+1} | S_t, D_t) \quad (4.20)$$

The stay tax can be used from any initial state to begin and maintain optimal management. Intuitively, the value of the stay tax is exactly the yearly marginal dissipatable rent accruing to satellites on the targeted management path. While the stay tax is necessarily optimal from the given initial state, it is sensitive to the estimated launch constraint. If the launch constraint is incorrectly estimated, the stay tax may also be incorrectly estimated. The influence of the launch constraint can be seen in the rate at which the tax increases: whenever the launch constraint binds, it constrains the magnitude of the estimated optimal stay tax.

Figure 4.7 shows optimal stay tax paths beginning in 2006.

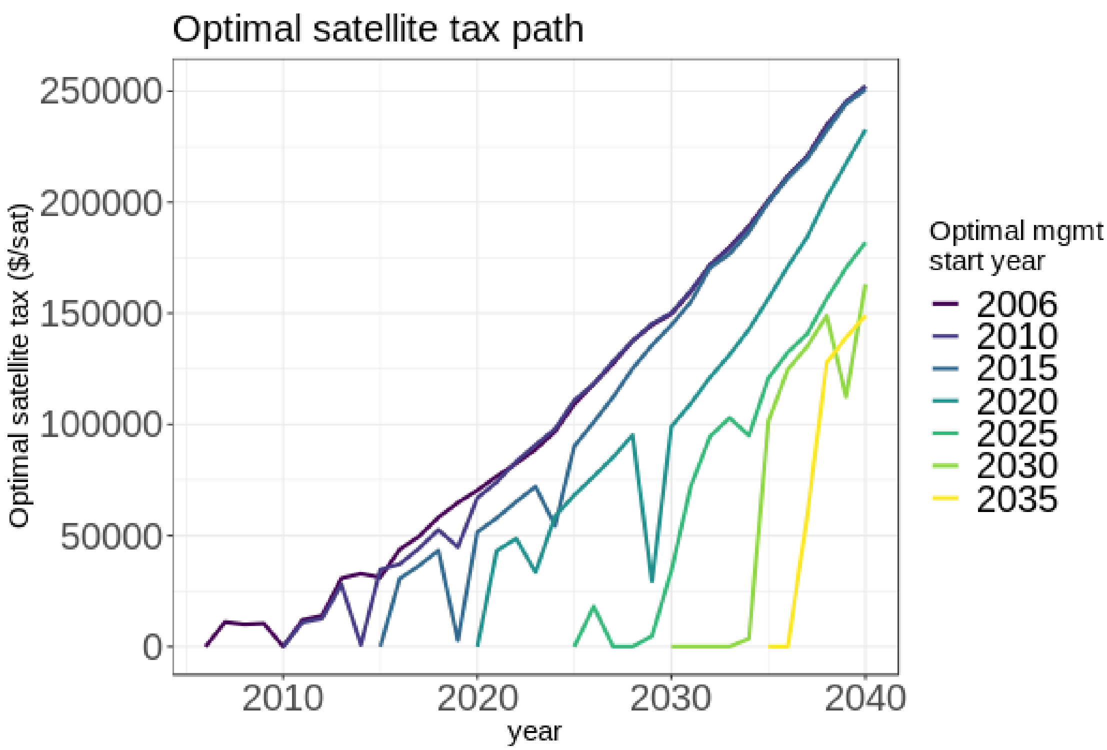

Figure 4.7: *Sample stay tax paths.*

4.4.3 The benefits of satellite taxes and the costs of inaction

Figure 4.8 shows the paths of an optimal satellite tax, the factor of satellite safety improvement, and the gain in net present value beginning at different points over the 2006—2040 sample period. We show these time paths for different optimal management start dates to illustrate the gains from earlier action. Figure 4.9 shows the associated paths of orbital aggregates with optimal management beginning in the same years.

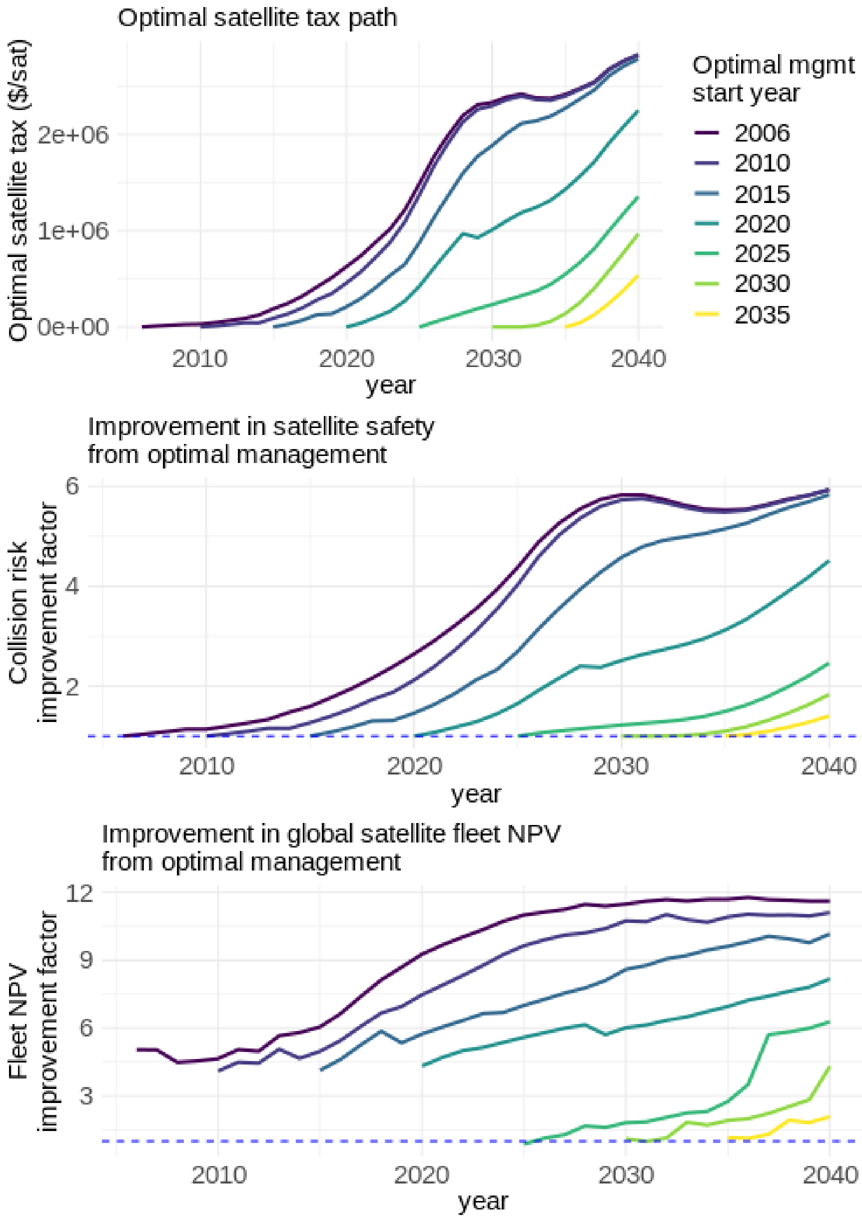

Figure 4.8: *Projected optimal orbit rental fee and gains from optimal management.*

Top panel: The optimal satellite jump tax over time. Values are in nominal US dollars.

Middle panel: The factor of collision risk reduction under optimal management, beginning in different years. A factor of 1 indicates no gain from optimal management, while 2 indicates a halving in risk.

Bottom panel: The factor of discounted net social welfare improvement under optimal management, beginning in different years. A factor of 1 indicates no gain from optimal management, while 2 indicates a doubling in welfare.

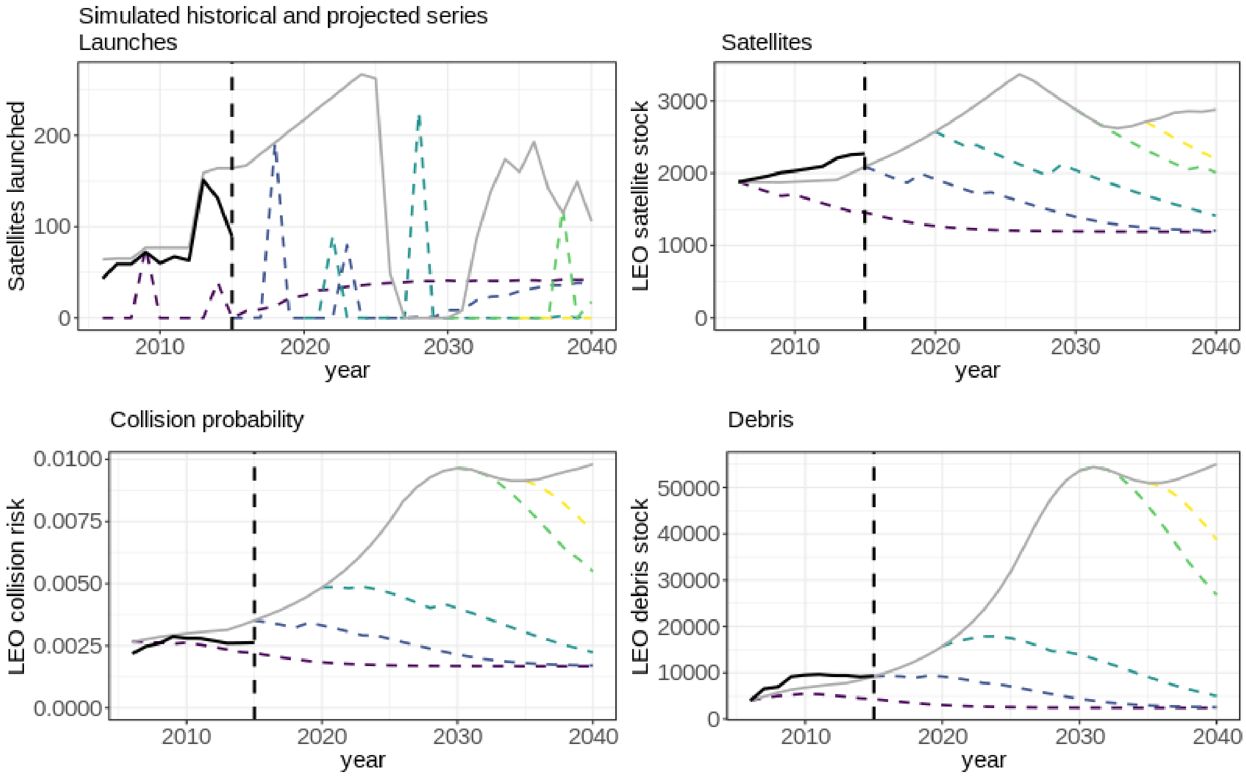

Figure 4.9: *Projected time paths of orbital aggregates with different optimal management start dates.*

The solid gray lines show the simulated open access paths. The dotted colored lines show simulated optimal paths, beginning from that year's open access state. The black lines show observed time paths.

An optimal satellite jump tax beginning in 2006 would have, by 2020, risen to a value of approximately \$250,000 USD, while one beginning in 2015 would have risen to approximately \$150,000 by 2020. The tax value rises over time to reflect both the projected growth in value of the space industry and the endogenous growth in dissipated resource rents from orbit use as the policy restores the orbits.

In an untaxed open access equilibrium, a tax of a single dollar would be sufficient to push the excess return on a satellite below the collision risk. With potential entrants dissuaded, the collision risk falls, requiring a larger tax to continue to dissuade potential entrants from dissipating the newly-restored resource rents. These resource rents can be seen in both the reduction in collision risk and the increase in fleet NPV.

While the optimal jump and stay taxes begin at the same value if they target the same optimal management path, note that the optimal stay tax rises much more slowly than the optimal jump tax. Both features are shown in Figure 4.10, where the hypothetical regulator wants to induce optimal management in 2020. This gap is driven by the nature of the dissipated/dissipatable rents the taxes are reflecting: the rents reflected in the optimal jump tax is roughly the cumulative value of the rents reflected in the optimal stay tax.

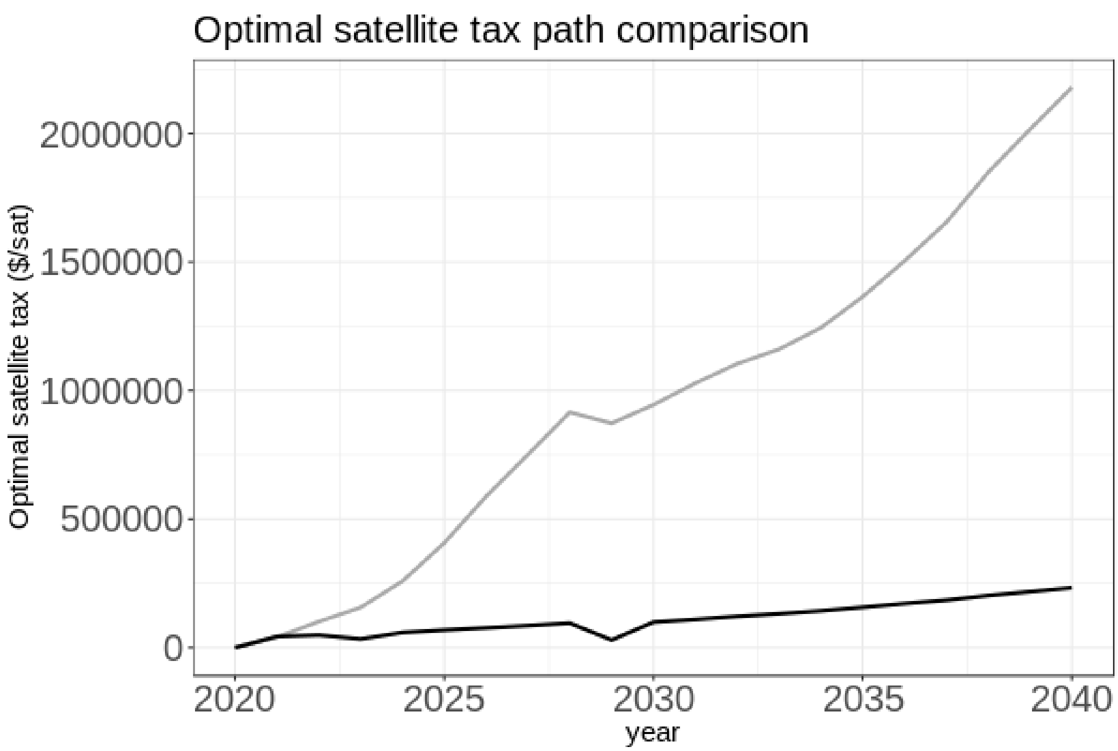

Figure 4.10: *Comparison of optimal jump and stay taxes beginning in 2020.*

The solid gray line shows the optimal jump tax — the tax which would induce firms to jump to the 2020 optimal management path. The solid black line shows the optimal stay tax — the tax which would induce firms to begin optimal management in 2020 and stay on that optimal path.

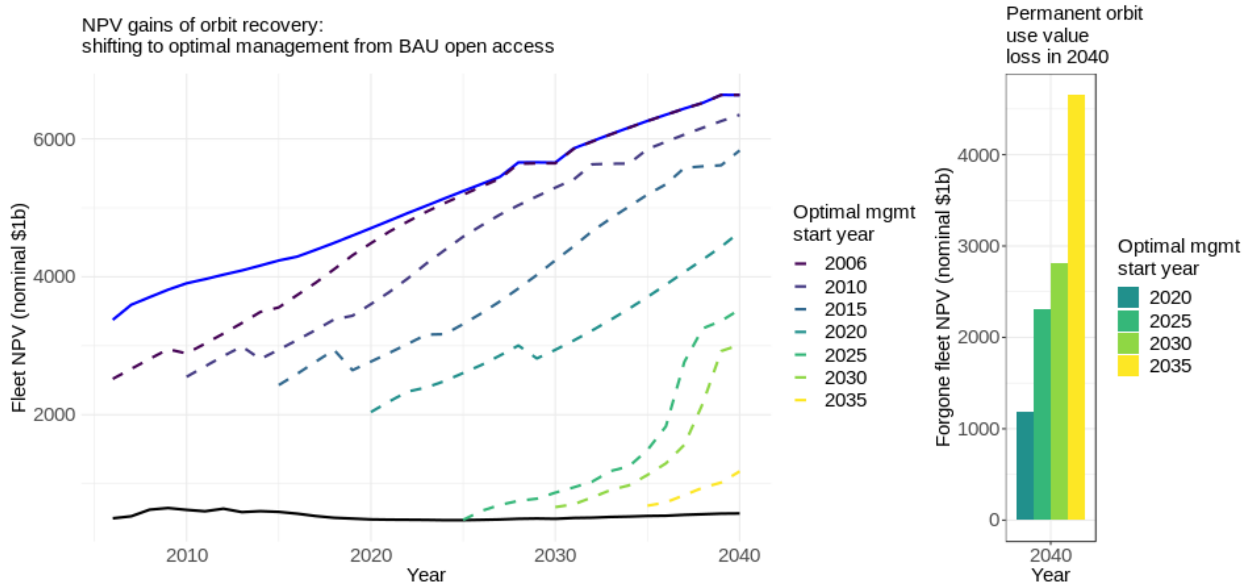

Figure 4.11: *Projected gains from optimal management and permanent orbit use value loss from delaying action.*

Left panel: NPV gains of orbit recovery, with optimal management beginning in different years. The solid blue line shows the path of the satellite fleet's NPV, had 2006 begun at the optimal steady state. The black line shows the business-as-usual open access fleet value path.

Right panel: Gap in permanent orbit use value from delaying optimal management, relative to starting optimal management in 2015.

These taxes can generate significant improvements in satellite safety and fleet NPV relative to the open access baseline — for a tax beginning in 2006, the safety improvement is on the order of 2.5x and the fleet NPV improvement is on the order of 4.5x, while a tax beginning in 2015 would have realized a safety improvement on the order of 2x and a fleet NPV improvement on the order of 4.25x. Figure 4.11 shows the global permanent orbit use value losses in 2040 from not adopting optimal management, relative to adopting optimal management in 2015.[12] By 2020, nearly $1 trillion in permanent orbit use value will have been lost

[12]"Permanent orbit use value" is the net present value, looking forward, of the satellite fleet. This value ignores losses or gains incurred prior to the reference year. In the comparisons here, the actual losses are larger as they will include value lost in the time between 2015 and the reference year. The name comes from analogy with "permanent income" in other economic settings.

relative to the 2015 optimal management baseline. By 2035, this will escalate to over \$4.5 trillion in lost permanent orbit use value.

4.5 Conclusion

There are novel dynamic externalities associated with orbit use, explored in Chapter 2. The most effective policy instrument to correct these externalities is a satellite tax, shown in Chapter 3. Knowing these facts, one may reasonably ask "how big should the tax be, and what will it yield in benefits?" In this chapter, we provide a method to compute the optimal satellite tax and welfare gains, and estimates of these quantities. A tax on orbiting satellites on the order of \$250,000 per satellite per year in 2020 is necessary to achieve outcomes in line with a world where optimal orbit management began in 2006. This tax would halve orbital collision risk in 2020, and quadruple the society's net present value from orbit use in 2020. The costs of inaction are substantial: relative to having begun optimal management in 2015, society stands to lose on the order of \$4.5 trillion if action is delayed until 2035.

Much work remains to be done in this area. We used aggregate data to obtain order-of-magnitude estimates for key policy and benefit values; studies using more detailed data will be able to measure optimal tax and social benefit values with greater precision, potentially even teasing out how the tax values should be scaled for different satellite types. More detailed calibrations, ideally using physics simulations and disaggregating to thinner shells, will provide both more precise estimates as well as the necessary spatially-differentiated tax schedule.

Though a satellite tax is the ideal corrective approach to the aforementioned extraterrestrial externalities, to quote Cicero, we live not in the glittering Republic of Plato but among the dregs of Romulus. A spatially-differentiated globally-harmonized satellite tax schedule is exceedingly unlikely to be adopted in the near future. Of considerable import, then, are questions of reality-constrained policy design and implementation. These constraints include national security uses of orbits and the associated geostrategic concerns, incentive compatibility for firms and national governments, dealing with legacy orbit users and their debris, and market structures in satellite launch and ownership. Armed with knowledge of the nature and benefits of optimal corrective policies, constrained policies can be evaluated in terms of their performance relative to unconstrained policies. Future research will be necessary to examine the nature and impact of real-world constraints on orbital management policy.

Bibliography

Adilov, Nodir, Peter J. Alexander, and Brendan M. Cunningham, “Earth Orbit Debris: An Economic Model,” *Environmental and Resource Economics*, January 2015, *60*, 81–98.

Ailor, William, James Womack, Glenn Peterson, and Norman Lao, “Effects of Space Debris on the Cost of Space Operations,” Technical Report IAC-10.A6.2.10 September 2010.

Akers, Agatha, “To Infinity and Beyond: Orbital Space Debris and How to Clean It Up,” *University of La Verne Law Review*, 2012, *33*.

Aldy, Joseph E, Todd D Gerarden, and Richard L Sweeney, “Investment versus Output Subsidies: Implications of Alternative Incentives for Wind Energy,” Working Paper 24378, National Bureau of Economic Research March 2018.

Ansdell, Megan, “Active Space Debris Removal: Needs, Implications, and Recommendations for Today’s Geopolitical Environment,” *Journal of Public & International Affairs*, 2010, *21*, 7–22.

Barrett, Scott, *The Theory of International Environmental Agreements*, Elsevier B.V.,

_ , “Climate treaties and approaching catastrophes,” *Journal of Environmental Economics and Management*, January 2013, *66*, 235–250.

Böhringer, Christoph, Knut Einar Rosendahl, and Halvor Briseid Storrøsten, “Robust policies to mitigate carbon leakage,” *Journal of Public Economics*, March 2017, *149*, 35–46.

Bradley, Andrew M. and Lawrence M. Wein, “Space debris: Assessing risk and responsibility,” *Advances in Space Research*, February 2009, *43*, 1372–1390.

Briggs, R.J., “Prices vs. quantities in a dynamic problem: Externalities from resource extraction,” *Resource and Energy Economics*, May 2011, *33*, 843–854.

Brodkin, Jon, “With latency as low as 25ms, SpaceX to launch broadband satellites in 2019,” *Ars Technica*, May 2017.

Carroll, J., “Bounties for orbital debris threat reduction,” in “in” NASA-DARPA International Conference on Orbital Debris Removal 2009.

Combined Space Operations Center, “Space-Track.org Satellite Catalog,” 2018. From Space-Track.org at `https://www.space-track.org/`.

Costello, Christopher, Steven D. Gaines, and John Lynham, "Can Catch Shares Prevent Fisheries Collapse?," *Science*, 2008, *321* (5896), 1678–1681.

Dietz, Thomas, Elinor Ostrom, and Paul C. Stern, "The Struggle to Govern the Commons," *Science*, 2003, *302* (5652), 1907–1912.

Dvorsky, George, "California Startup Accused of Launching Unauthorized Satellites Into Orbit: Report," *Gizmodo*, March 2018.

European Space Agency, "DISCOS database," 2018. From ESA at `https://discosweb.esoc.esa.int/web/guest/home`.

Fischer, Carolyn and Alan K. Fox, "Comparing policies to combat emissions leakage: Border carbon adjustments versus rebates," *Journal of Environmental Economics and Management*, March 2012, *64*, 199–216.

Fowlie, Meredith L., "Incomplete Environmental Regulation, Imperfect Competition, and Emissions Leakage," *American Economic Journal: Economic Policy*, 2009, *1*, 72–112.

Gordon, H. Scott, "The Economic Theory of a Common-Property Resource: The Fishery," *Journal of Political Economy*, 1954, *62.*

Gorove, Stephen, "Interpreting Article II of the Outer Space Treaty," *Fordham Law Review*, 1969, *37.*

Hardin, Garrett, "The Tragedy of the Commons," *Science*, 1968, *162* (3859), 1243–1248.

Haveman, Robert H, "Common Property, Congestion, and Environmental Pollution," *The Quarterly Journal of Economics*, May 1973, *87*, 278–287.

Hughes, Jonathan and Daniel Kaffine, "When is Increasing Consumption of Common Property Optimal? Sorting, Congestion and Entry in the Commons," *Journal of Environmental Economics and Management*, January 2017, *81.*

IADC, "IADC Space Debris Mitigation Guidelines," Technical Report IADC-02-01 September 2007.

Jehn, R., V. Agapov, and C. Hernandez, "The situation in the geostationary ring," *Advances in Space Research*, 2005, *35*, 1318–1327.

Johnson, Nicholas L, "Space traffic management concepts and practices," *Acta Astronautica*, 2004, *55* (3-9), 803–809.

Jonas, Adam, Armintas Sinkevicius, Simon Flannery, Benjamin Swinburne, Patrick Wellington, Terence Tsui, Rajeev Lalwani, James E Faucette, Brian Nowak, Ravi Shanker, Kai Pan, and Eva Zlotnicka, "Space: Investment Implications of the Final Frontier," Technical Report, Morgan Stanley November 2017.

Jones, Noelle, Jan Skora, Calvin Monson, Robert W Jones, Stuart Jack, and Kurt Eby, "Study on the Global Practices for Assigning Satellite Licences and Other Elements," Technical Report November 2010.

Kelly, David L and Charles D Kolstad, "Integrated assessment models for climate change control," *International yearbook of environmental and resource economics*, 1999, *2000*, 171–197.

Keohane, Nathaniel, Benjamin Van Roy, and Richard Zeckhauser, "Managing the quality of a resource with stock and flow controls," *Journal of Public Economics*, 2007, *91*, 541–569.

Kessler, Donald J and Burton G Cour-Palais, "Collision Frequency of Artificial Satellites: The Creation of a Debris Belt," *Journal of Geophysical Research*, June 1978, *83*, 2637–2646.

_ **and Phillip D. Anz-Meador**, "Critical Number of Spacecraft in Low-Earth Orbit: Using Satellite Fragmentation Data to Evaluate the Stability of the Orbital Debris Environment," *Proceedings of the Third European Conference on Space Debris*, March 2001.

_ **, Nicholas L Johnson, J-C Liou, and Mark Matney**, "The Kessler Syndrome: Implications to Future Space Operations," *33rd Annual AAS Guidance and Control Conference*, February 2010.

Ketz, Philipp, "Subvector inference when the true parameter vector may be near or at the boundary," *Journal of Econometrics*, 2018, *207* (2), 285 – 306.

Klima, Richard, Daan Bloembergen, Rahul Savani, Karl Tuyls, Daniel Hennes, and Dario Izzo, "Space Debris Removal: A Game Theoretic Analysis," *Games*, 2016, *7* (3).

Klinkrad, H. and N. L. Johnson, "Space Debris Environent Remediation Concepts," in "Fifth European Conference on Space Debris," Vol. 672 of ESA Special Publication March 2009, p. 50.

Letizia, F., S. Lemmens, and H. Krag, "Application of a debris index for global evaluation of mitigation strategies," in "in" 69th International Astronautical Congress October 2018.

Letizia, Francesca, Camilla Colombo, Hugh Lewis, and Holger Krag, "Extending the ECOB space debris index with fragmentation risk estimation," 2017.

Liou, J.C., "A statistical analysis of the future debris environment," *Acta Astronautica*, December 2006, *62*, 264–271.

_ **and N.L. Johnson**, "Instability of the present LEO satellite populations," *Advances in Space Research*, January 2008, *41*, 1046–1053.

_ **and** _ , "A sensitivity study of the effectiveness of active debris removal in LEO," *Acta Astronautica*, August 2009, *64*, 236–243.

_ **, D.T. Hall, P.H. Krisko, and J.N. Opiela**, "LEGEND - a three-dimensional LEO-to-GEO debris evolutionary model," *Advances in Space Research*, February 2004, *34*, 981–986.

Macauley, Molly K, "Allocation of Orbit and Spectrum Resources for Regional Communications: What's at Stake?," *Journal of Law and Economics*, October 1998, *41*, 737–764.

_ , "The economics of space debris: Estimating the costs and benefits of debris mitigation," *Acta Astronautica*, May 2015, *115*, 160–164.

_ **and Paul R Portney**, "Slicing the Geostationary Pie: Property Rights in Orbit," *AEI Journal on Regulation*, July/August 1984.

Mayeres, Inge, Sara Ochelen, and Stef Proost, "The marginal external costs of urban transport," *Transportation Research Part D: Transport and Environment*, 1996, *1* (2), 111–130.

Merges, Robert P. and Glenn H. Reynolds, "Rules of the Road for Space?: Satellite Collisions and the Inadequacy of Current Space Law," *The Environmental Law Reporter (ELR) News&Analysis*, 2010, *40.*

Muller, Catherine, Olga Rozanova, and Miguel Urdanoz, "Economic Valuation of Debris Removal," in "68th International Astronautical Congress" 2017.

NASA Orbital Debris Program Office, *Orbital Debris Quarterly News*, May 2017, *21.* From `https://www.orbitaldebris.jsc.nasa.gov/quarterly-news/pdfs/odqnv21i2.pdf`.

Newell, Richard G. and William A. Pizer, "Regulating stock externalities under uncertainty," *Journal of Environmental Economics and Management*, 2003, *45*, 416–432.

Newell, Richard, William Pizer, and Jiangfeng Zhang, "Managing permit markets to stabilize prices," *Environmental and Resource Economics*, 2005, *31* (2), 133–157.

Nordhaus, William, "Chapter 16 - Integrated Economic and Climate Modeling," in Peter B. Dixon and Dale W. Jorgenson, eds., *Handbook of Computable General Equilibrium Modeling SET, Vols. 1A and 1B*, Vol. 1 of Handbook of Computable General Equilibrium Modeling, Elsevier, 2013, pp. 1069 – 1131.

Ostrom, Elinor, "COPING WITH TRAGEDIES OF THE COMMONS," *Annual Review of Political Science*, 1999, *2* (1), 493–535.

Parish, Ross M and Keith Robert McLaren, "Relative cost-effectiveness of input and output subsidies," *Australian Journal of Agricultural Economics*, 1982, *26* (1), 1–13.

Pearson, J., J. Carroll, and E. Levin, "Active Debris Removal: EDDE, the ElectroDynamic Debris Eliminator," *Space Manufacturing*, October 2010, *14.*

Qingyu, LUO, JUAN Zhicai, SUN Baofeng, and JIA Hongfei, "Method research on measuring the external costs of urban traffic congestion," *Journal of Transportation Systems Engineering and Information Technology*, 2007, *7* (5), 9–12.

Rossi, A., L. Anselmo, A. Cordelli, P. Farinella, and C. Pardini, "Modelling the evolution of the space debris population," *Planetary Space Science*, May 1998, *46*, 1583–1596.

Salter, Alexander William, "Space Debris: A Law and Economics Analysis of the Orbital Commons," *Mercatus Working Paper*, September 2015.

_ **and Peter T Leeson**, "Celestial Anarchy," *Cato Journal*, January 2014, *34.*

Sandler, Todd and William Schulze, "The Economics of Outer Space," *Natural Resources Journal*, April 1981, *21*, 371–393.

Sandmo, Agnar, "Direct versus indirect Pigovian taxation," *European Economic Review*, 1978, *7* (4), 337–349.

Schaub, Hanspeter, Lee E.Z. Jasper, Paul V. Anderson, and Darren S. McKnight, "Cost and risk assessment for spacecraft operation decisions caused by the space debris environment," *Acta Astronautica*, April 2015, *113*, 66–79.

Scott, Anthony, "The Fishery: The Objectives of Sole Ownership," *Journal of Political Economy*, 1955, *62.*

Selk, Avi, “Thousands of tiny satellites are about to go into space and possibly ruin it forever,” *The Washington Post*, April 2017.

Senechal, Thierry, “Orbital Debris: Drafting, Negotiating, Implementing a Convention,” *Masters Thesis*, June 2007.

Shimabukuro, Alessandro, “No deal in space: A bargaining model analysis of U.S. resistance to space arms control,” *Space Policy*, April 2014, *30*, 13–22.

Stavins, Robert N, “The problem of the commons: still unsettled after 100 years,” *American Economic Review*, 2011, *101* (1), 81–108.

Svensson, Lars EO, “What is wrong with Taylor rules? Using judgment in monetary policy through targeting rules,” *Journal of Economic Literature*, 2003, *41* (2), 426–477.

Sweeney, James L., “Economic Theory of Depletable Resources: an Introduction,” in Allen V. Kneese and James L. Sweeney, eds., *Allen V. Kneese and James L. Sweeney, eds.*, Vol. 3 of Handbook of Natural Resource and Energy Economics, Elsevier, 1993, pp. 759 – 854.

TelAstra, Inc., “Communications Satellite Databases,” 2017. From TelAstra, Inc. at `http://telastra.com/` . The data shown here are from the 2017 edition, which is available through the CU Boulder Libraries.

Telser, L.G., “A Theory of Self-Enforcing Agreements,” *Journal of Business*, January 1980, *53*, 27–44.

Tkatchova, Stella, *Space Debris Mitigation*, Berlin, Heidelberg: Springer Berlin Heidelberg,

Union of Concerned Scientists, “UCS Satellite Database,” 2018. From UCS at `https://www.ucsusa.org/nuclear-weapons/space-weapons/satellite-database`.

Walters, Alan A, “The theory and measurement of private and social cost of highway congestion,” *Econometrica: Journal of the Econometric Society*, 1961, pp. 676–699.

Weeden, Brian C, “Overview of the Legal and Policy Challenges of Orbital Debris Removal,” Technical Report IAC-10.A6.2.10 September 2010.

_ **and Tiffany Chow**, “Taking a common-pool resources approach to space sustainability: A framework and potential policies,” *Space Policy*, July 2012, *28*, 166–172.

Weitzman, Martin L, “Prices vs. Quantities,” *Review of Economic Studies*, 1974, *41* (4), 477–491.

Wienzierl, Matthew, “Space, the Final Economic Frontier,” *Journal of Economic Perspectives*, Spring 2018, *32*, 173–192.

Wilkerson, Jordan, Benjamin Leibowicz, Delavane Diaz, and John Weyant, “Comparison of Integrated Assessment Models: Carbon Price Impacts on U.S. Energy,” *Energy Policy*, 01 2015, *76*, 18–31.

Appendix A

Extensions to the deterministic model

A.1 The effects of spectrum congestion

Satellite applications generally require transmissions to and from the Earth. These transmissions may be the satellite's main output or incidental to its operation. In both cases, satellite operators must secure spectrum use rights from the appropriate national authorities for their broadcast and receiving locations. How will spectrum management affect collisions and debris growth?

Spectrum congestion degrades signal quality, making the per-period output of a satellite decreasing in the number of satellites in orbit, i.e. $\pi = \pi(S), \pi'(S) < 0$. If satellites are launched only when they have appropriate spectrum rights and spectrum use is optimally managed, then firms will be forced to account for their marginal effects on spectrum congestion in their decision to launch or not. The equilibrium condition becomes

$$L(S_{t+1}, D_{t+1}) = r_s(S_{t+1}) + r_s'(S_{t+1}) - r, \tag{A.1}$$

where $r_s'(S_{t+1}) = \pi'(S_{t+1})/F < 0$ by assumption. The equilibrium set would no longer be a collision rate isoquant, although it would still be a surface in the state space. Even if it isn't managed optimally, spectrum congestion will reduce the equilibrium collision rate by reducing the rate of excess return on a satellite.[1]

[1] If spectrum use were also under open access then the marginal congestion effect ($r_s'(S_{t+1})$) would not be in the equilibrium condition, but the equilibrium set would still not be a collision rate isoquant.

Open access orbit use will still be inefficient. Although spectrum congestion can reduce the chance of Kessler Syndrome, efficient spectrum management will not incorporate the marginal external cost of collisions and debris growth, $\xi(S_{t+1}, D_{t+1})$. Collision and debris growth management policies could be implemented through spectrum pricing. The rest of the analysis in this paper still goes through when spectrum congestion is considered, though some proofs become more complicated since the equilibrium set is no longer a collision rate isoquant.

A.2 The effects of limited satellite lifespans

Satellites do not, in general, produce returns forever until destroyed in a collision. Over 1967-2015, planned satellite lifetimes ranged from 3 months to 20 years, with longer lifetimes being more representative of larger and more expensive GEO satellites.[2] How would finite lifetimes affect the collision rate and debris growth problems?

Suppose satellite lifetimes are finite and exogenously distributed with mean μ^{-1}. Satellites live at least one time period on average, so that $\mu^{-1} > 1$. The equilibrium condition becomes

$$L(S_{t+1}, D_{t+1}) = \frac{r_s - r - \mu}{1 - \mu}, \tag{A.2}$$

which is lower than the equilibrium collision rate when satellites are infinitely lived. Intuitively, the fact that the satellite will stop generating returns at some point reduces its expected present value, and with it the incentive to launch. All else equal, shorter lifetimes reduce the equilibrium collision rate. The rest of the analysis in the paper goes through with minor modifications.

Satellites built for GEO tend to be longer lived than satellites launched for LEO. If shorter lifetimes have reduced satellite costs, then the downward shift in the collision rate isoquant from the shorter lifetimes will be balanced against the upward shift caused by higher rates of return. The net effect may be higher or

[2]These numbers are taken from the Union of Concerned Scientists' publicly available data on satellites. The data are available at https://www.ucsusa.org/nuclear-weapons/space-weapons/satellite-database.

lower equilibrium rates of collisions.

The distinction between exogenous and endogenous lifetimes is relevant here. The above analysis hinges on satellite lifetimes being exogenously set. This is not the case in reality. Satellite lifetimes are determined by cost minimization concerns, technological constraints, expectations of component failures, and expectations of future technological change. Incorporating all these features to realistically model the choice of satellite life along with launch decisions and their effects on orbital stock dynamics is beyond the scope of this paper, though it is an interesting area for future research. The assumption that the end-of-life is random simplifies the analysis, but does not change any conclusions over imposing a pre-specified end date in this model.[3]

A.3 The effects of changes in satellite returns, costs, and discount rates

Though it simplifies the analysis, the rate of return on a satellite is not constant over time. How would changes over time in these economic parameters affect our conclusions? For simplicity, suppose that costs and returns vary exogenously and are known in advance.[4] The open access equilibrium condition is then

$$\pi_{t+1} = (1+r_t)F_t - (1-L(S_{t+1},D_{t+1}))F_{t+1} \tag{A.3}$$

Equation A.3 can be rewritten as

$$L(S_{t+1},D_{t+1}) = 1 + r_{s,t+1} - (1+r_t)\frac{F_t}{F_{t+1}} \tag{A.4}$$

$$\implies L(S_{t+1},D_{t+1}) = \underbrace{\left(r_{s,t+1} - r_t\frac{F_t}{F_{t+1}}\right)}_{\substack{\text{excess return} \\ \text{on a satellite}}} + \underbrace{\left(1 - \frac{F_t}{F_{t+1}}\right)}_{\substack{\text{capital gains} \\ \text{from open access} \\ \text{and satellite launch} \\ \text{cost variation}}} \tag{A.5}$$

---

[3]This simplification could matter in a model where firms own multiple satellites and have to plan replacements.

[4]Uncertainty over costs and returns doesn't change the qualitative results, though it introduces expectations over the changes. Endogeneity in the changes, for example due to investment in R&D or marketing, may have more significant consequences which are beyond the scope of this paper.

If $\pi_{t+1} > (1+r_t)F_t$, then the one period return on a satellite is greater than the gross return on the launch cost from the safe asset, and the collision rate will be 1. Ignoring that corner case, the equilibrium collision rate is decreasing in the current cost of launching a satellite, but increasing in the future cost of launching a satellite. All else equal, the collision rate in $t+1$ will be lower when F_{t+1} increases. This highlights the role of the launch cost under open access: if firms enter until zero profits in each period, future increases in the cost deter firms from entering in the future, increasing the value of satellites already in orbit by the amount of the cost increase.

When the costs and returns are time-varying, the equilibrium set is still a collision rate isoquant, though the isoquant selected may vary over time. These changes do not affect the physical dynamics or the Kessler threshold, though they may affect how close the selected equilibrium is to the threshold. If the parameters vary so that the ratio $\frac{\pi_{t+1}-(1+r_t)F_t}{F_{t+1}}$ is stationary, then the equilibrium set will stay on the same isoquant.

A.4 The effects of space weather

Sunspots have two effects: first, changes in radiation pressure force satellites in higher orbits to spend more fuel on stationkeeping; second, the Earth's atmosphere expands in response to the solar activity, increasing drag and debris decay in all but the highest orbits. The latter effect can be particularly significant for active satellites in LEO. Formally, suppose sunspots cause decay rates to vary periodically with mean $\bar{\delta}$, making the laws of motion

$$S_{t+1} = S_t(1 - L(S_t, D_t)) + X_t$$

$$D_{t+1} = D_t(1 - \delta_t) + G(S_t, D_t) + mX_t$$

If the variations are exogenous and publicly known, there are no changes to the firm or planner's Bellman equations. Figure A.1 shows example time paths with and without sunspot activity.

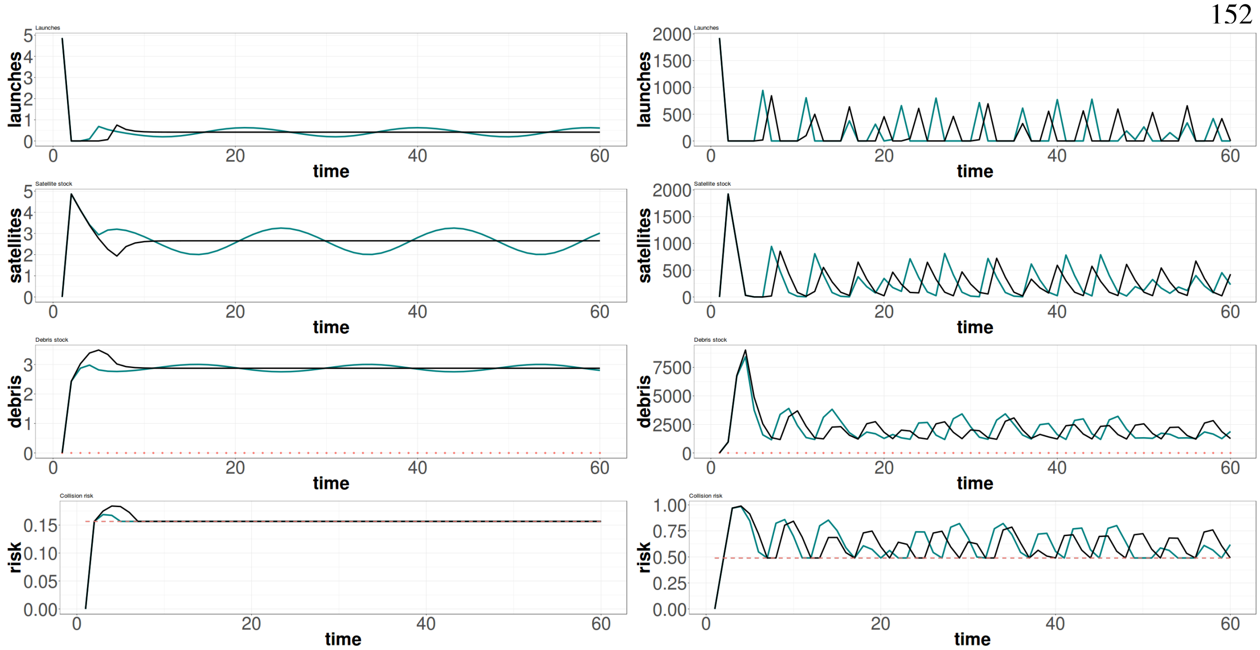

Figure A.1: Open access equilibrium paths with no sunspots (black line) and with subspots driven by $\delta_t = \bar{\delta} + \omega_1 \sin(\omega_2 t)$ (teal line).

The results in Proposition 5 describe how open access equilibria will respond to increases in the decay rate caused by space weather, holding costs constant. As the activity increases, firms will have incentives to take advantage of the increased renewal by launching more satellites. This may increase or decrease the net amount of debris. As the activity decreases, firms will reduce the launch rate, decreasing or increasing the net amount of debris. These effects will be offset by the extent to which sunspot activity increases the cost of operating a satellite.

Increases in the decay rate have the benefit of shifting the Kessler threshold up (Proposition 4). If the net increase in debris due to the launch response is not too large, then increased sunspot activity may stabilize orbit use. On the other hand, if the rebound effect for decay rates is strong enough, sunspot activity may push the equilibrium debris level into the basin of the Kessler region. Even when the rebound effect is not very strong, a decrease in the decay rate at the end of a period of extra sunspot activity may make the original open access steady state locally unstable given the increase in debris due to the rebound effect.

Appendix B

Full proofs for Chapter 2

Proof of Proposition 4:

Proof. Period t values are shown with no subscript, and period $t+1$ values are marked with a $'$, e.g. $S_t \equiv S, S_{t+1} \equiv S'$. Applying the Implicit Function Theorem to equation 3.51, we get

$$\frac{\partial X}{\partial S} = -\frac{\left[\frac{\partial L(S',D')}{\partial S'}(1 - L(S,D) - \frac{\partial L(S,D)}{\partial S}S) + \frac{\partial L(S',D')}{\partial D'}\frac{\partial G(S,D)}{\partial S}\right]}{\left[\frac{\partial L(S',D')}{\partial S'}\frac{\partial S'}{\partial X} + \frac{\partial L(S',D')}{\partial D'}\frac{\partial D'}{\partial X}\right]} < 0$$

$$\frac{\partial X}{\partial D} = -\frac{\frac{\partial L(S',D')}{\partial S'}(-S\frac{\partial L(S,D)}{\partial D}) + \frac{\partial L(S',D')}{\partial D'}(1 - \delta + \frac{\partial G(S,D)}{\partial D})}{\frac{\partial L(S',D')}{\partial S'}\frac{\partial S'}{\partial X} + \frac{\partial L(S',D')}{\partial D'}\frac{\partial D'}{\partial X}} \lesseqgtr 0$$

$\frac{\partial X}{\partial D} < 0$ if

$$(S,D): \quad \underbrace{\frac{\partial L(S',D')}{\partial S'}S}_{\substack{t+1 \text{ marginal collision probability from} \\ t+1 \text{ satellites} \cdot \text{present satellite stock}}} \quad < \quad \underbrace{\frac{\frac{\partial L(S',D')}{\partial D'}}{\frac{\partial L(S,D)}{\partial D}}}_{\substack{\text{growth rate of} \\ \text{marginal collision rate} \\ \text{due to debris}}} \quad \cdot \underbrace{\left(1 - \delta + \frac{\partial G(S,D)}{\partial D}\right)}_{\substack{\text{net growth in debris} \\ \text{due to debris}}} \tag{B.1}$$

and $\frac{\partial X}{\partial D} > 0$ otherwise.

□

Proof of Proposition 5:

Proof. Period t values are shown with no subscript, and period $t+1$ values are marked with a $'$, e.g. $S_t \equiv S, S_{t+1} \equiv S'$. All present-period variables are assumed to be at equilibrium values. Applying the Implicit Function Theorem to equation 3.51, we get

$$\frac{\partial X}{\partial r_s} = \frac{1}{L_S(S',D') + mL_D(S',D')} > 0$$
$$\frac{\partial D'}{\partial r_s} = \frac{1}{L_D(S',D')} > 0$$

$$\frac{\partial X}{\partial m} = -\frac{L_D(S',D')X}{L_S(S',D') + mL_D(S',D')} < 0$$
$$\frac{\partial D'}{\partial m} = X + m\frac{\partial X}{\partial m}$$
$$\frac{\partial D'}{\partial m} > 0 \implies 1 > \frac{mL_D(S',D')}{L_S(S',D') + mL_D(S',D')},$$

which is true as long as $L_S > 0$.

$$\frac{\partial X}{\partial \delta} = \frac{L_D(S',D')D}{L_S(S',D') + mL_D(S',D')} > 0$$
$$\frac{\partial D'}{\partial \delta} = -D + m\frac{\partial X}{\partial m}$$
$$\frac{\partial D'}{\partial \delta} < 0 \implies 1 > \frac{mL_D(S',D')}{L_S(S',D') + mL_D(S',D')},$$

which is also true as long as $L_S > 0$. □

Proof of Proposition 6:

Proof. The open access steady states are defined by equations 2.33, 2.34, and 2.35. Equation 2.33 implicitly determines the number of satellites as a function of the amount of debris, the excess return on a satellite, and the collision rate function,

$$L(S,D) = r_s - r \implies S = S(r_s - r, D). \tag{B.2}$$

Since we assumed $L(S,D)$ is increasing in each argument, $S(r_s - r, D)$ is decreasing in D. This implicit function allows us to reduce equations 2.33, 2.34, and 2.35 to a single equation in debris,

$$\mathscr{Y}(D) = -\delta D + G(S(r_s - r, D), D) + m(r_s - r)S(r_s - r, D) = 0. \tag{B.3}$$

$-\delta D$ and $m(r_s-r)S(r_s-r,D)$ are both monotonically decreasing in D. Since $G(S,D)$ is increasing in both arguments but $S(r_s-r,D)$ is decreasing in D, $G(S(r_s-r,D),D)$ may be increasing or decreasing in D over an arbitrary positive interval. In the limiting cases where $D\to 0$ or $D\to\infty$, $G(S(r_s-r,D),D)\to G(S(r_s-r,0),0)>0$ and $G(S(r_s-r,D),D)\to G(0,D)>0$. If $G(S,D)=L(S,D)$, $G(S(r_s-r,D),D)$ would be constant for all D because $S(r_s-r,D)$ is such that $L(S,D)=r_s-r$ for all S,D where this is possible. If $G(S,D)$ is locally more convex than $L(S,D)$, then $G(S(r_s-r,D),D)$ will be first decreasing and then increasing (over the local interval) as D increases, and vice versa if $G(S,D)$ is locally more concave than $L(S,D)$ over a local interval. So,

$$\delta D = G(S(r_s-r,D),D)+m(r_s-r)S(r_s-r,D)$$

may have zero, one, or more than one interior solutions. If $L(S,D)$ is globally strictly concave and $G(S,D)$ is globally weakly convex, there can be up to 2 solutions. $\square$

Proof of Proposition 7:

Proof. We use the reduction from Proposition 6 to simplify the proof. The open access steady states are solutions to equation B.3, and the sign of $\frac{\partial \mathscr{Y}}{\partial D}$ at the solutions allows us to classify the stability of the system. Applying the Implicit Function Theorem to equation 2.33 to calculate $S_D(r_s-r,D)$, then differentiating $\mathscr{Y}(D)$ in a neighborhood of an arbitrary solution D^*,

$$\frac{\partial \mathscr{Y}}{\partial D}(D^*) = -\delta - \frac{L_D(S^*,D^*)}{L_S(S^*,D^*)}(G_S(S^*,D^*)+m(r_s-r))+G_D(S^*,D^*), \tag{B.4}$$

where $S^*\equiv S(r_s-r,D^*)$. The first two terms of $\frac{\partial \mathscr{Y}}{\partial D}$ are negative and the last term is positive. Since $\frac{\partial \mathscr{Y}}{\partial D}(D^*)$ may be positive or negative, the generic solution considered may be stable or unstable depending on the physical and economic parameters.

Analysis of $\frac{\partial^2 \mathscr{Y}}{\partial D^2}(D^*)$ shows a necessary condition for $\frac{\partial \mathscr{Y}}{\partial D}(D^*)$ to be increasing in D, again using the

Implicit Function Theorem to calculate $S_{DD}(r_s - r, D)$:

$$\begin{aligned}\frac{\partial^2 \mathscr{Y}}{\partial D^2}(D^*) > 0 &\implies S_D G_S D + S_D^2 G_{SS} + S_{DD}(G_S - m(r_s - r)) > 0\\ &\implies \frac{L_D L_S (G_{SD} L_S - G_{SS} L_D)}{L_{DD} L_S^2 + L_{SD} L_D^2} > m(r_s - r) - G_S,\end{aligned}$$

which is not necessarily satisfied by every open access steady state without further restrictions on the physical and economic parameters. □

Proof of Proposition 8:

Proof. Consider an arbitrary open access steady state (S', D'). Our approach will be to first characterize the set of points which can reach the open access steady state in one period, then show that other points in the action region must overshoot at least one state variable in approaching the equilibrium isoquant.

The set of initial conditions which can reach that steady state are one iteration of the physical dynamics away from a line segment intersecting the steady state. Given an arbitrary point (S, D), (S', D') can be written as

$$\begin{bmatrix} S' \\ D' \end{bmatrix} = \begin{bmatrix} S \\ D \end{bmatrix} + \begin{bmatrix} -L(S,D)S \\ -\delta D + G(S,D) \end{bmatrix} + \begin{bmatrix} X \\ mX \end{bmatrix}. \tag{B.5}$$

The second term on the right hand side is the effect of the physical dynamics independent of launches. The third term on the right hand side is a line segment with magnitude $X(1+m^2)^{1/2}$ and slope m. $X > 0$ implies that the sum of the first two terms is in the action region. Under open access, X is determined so that $L(S', D') = r_s - r$. Since the magnitude of X only changes the length of the line segment from $[S(1 - L(S,D)), D(1-\delta) + G(S,D)]^T$ but not its slope, any initial condition for which firms launch satellites and reach an open access steady state in one step must be one iteration of the physical dynamics away from a line segment intersecting the steady state in question. The case where $X = 0$ and the system reaches an open access steady state is less interesting, since in this case one iteration of the physical dynamics places the system exactly at the steady state in question (which is by definition on a line segment with slope m

intersecting itself).

Now consider a different initial condition, (S_a, D_a), which still leads to the interior of the action region after one iteration of the physical dynamics, but does not end up on a line segment with slope m which intersects an open access steady state. Since it leads to the interior of the action region, $X > 0$ holds and will be such that $L(S'_a, D'_a) = r_s - r$. Now, $L(S,D)$ and $G(S,D)$ are both monotonic in each argument, and $L(S'_a, D'_a) = L(S', D') = r_s - r$ since $X > 0$ implies next period aggregates are on the equilibrium isoquant. If $S'_a = S$, the monotonicity of $L(S,D)$ implies that $D'_a = D$ and vice versa. If $S'_a < S'$ and $D'_a < D$ (or both inequalities were reversed), either (S', D') or (S'_a, D'_a) would not be an equilibrium. So, either $S'_a > S'$ or $D'_a > D'$ must be true. Combining these facts, if $S'_a > S'$ then $D'_a < D'$ must be true, and vice versa.

If open access launching from an initial condition overshot the steady state in both satellites and debris, it would no longer be on the equilibrium collision rate isoquant. So any initial condition which leads to the action region must do one of three things: it must lead directly to an open access steady state, or it must overshoot the steady state satellite level, or it must overshoot the steady state debris level. $\square$

Proof of Proposition 9:

Proof. We suppress function arguments to reduce notation; all functions are evaluated at an arbitrary open access steady state. Applying the Implicit Function Theorem to equation B.3, and applying it again to equation 2.33 to calculate $\frac{\partial S}{\partial r_s}$ where necessary, we get

$$\frac{\partial D}{\partial r_s} = \frac{\frac{G_S}{L_S} + m(S(r_s - r, D) + \frac{r_s - r}{L_S})}{\delta + \frac{L_D}{L_S}(G_S + m(r_s - r)) - G_D} \lesseqgtr 0,$$

$$\frac{\partial D}{\partial m} = \frac{(r_s - r)S(D)}{\delta + \frac{L_D}{L_S}(G_S + m(r_s - r)) - G_D} \lesseqgtr 0,$$

$$\frac{\partial D}{\partial \delta} = \frac{D}{-\delta - \frac{L_D}{L_S}(G_S + m(r_s - r)) + G_D} \lesseqgtr 0.$$

An open access steady state is locally stable if $-\delta - \frac{L_D}{L_S}(G_S + m(r_s - r)) + G_D < 0$. If a steady state is locally stable, then $\frac{\partial D}{\partial r_s} > 0$, $\frac{\partial D}{\partial m} > 0$, and $\frac{\partial D}{\partial \delta} < 0$. The inequalities are reversed when a steady state is locally unstable. $\square$

Appendix C

Extensions to the stochastic model

C.1 Spectrum use management and price effects

So far I have assumed that there is no spectrum congestion from satellites and that the entry of new firms does not affect the price of satellite services. In practice, radio frequency interference is one of the major concerns of space traffic control. However, policy to manage spectrum use is not a focus of this paper because it is generally handled well by existing institutions Johnson (2004). The effect of optimally managed spectrum congestion on the expected collision risk in a deterministic setting is shown in section A.1 of the Appendix. In this section, I adapt the result to the stochastic setting setting and show that permits or fees for spectrum use can approximate stock controls.

Spectrum congestion degrades the quality of the signals to and from satellites. This makes the per-period output from a satellite decreasing in the number of orbital spectrum users. For simplicity, suppose that all satellites in orbit use enough spectrum to have some congestion impact. The per-period return function is then $\pi = \pi(S), \pi'(S) < 0$, and the one-period rate of return on a satellite is $\pi(S)/F = r_s(S)$. Assuming spectrum is optimally managed, firms will account for their marginal impact on spectrum congestion when they launch their satellite. The open access equilibrium condition, equation 3.11, becomes

$$E_t[\ell_{t+1}] = r_s(S_{t+1}) - r + r_s'(S_{t+1}). \tag{C.1}$$

Satellite owners would internalize the final term, $r'(S_{t+1})$, through a permit or fee system. Though spectrum permits may be purchased before the satellite is launched, their continued use is contingent on the firm abiding by non-interference protocols and any other stipulations by the appropriate regulatory body. Similarly, an optimal fee for spectrum use would adjust to reflect the marginal spectrum congestion from another broadcasting satellite. In general, regulated spectrum use will adjust the equilibrium collision rate to be

$$E_t[\ell_{t+1}] = r_s(S_{t+1}) - r - q_{t+1}, \tag{C.2}$$

where q_{t+1} is the spectrum use fee or permit price. Note that equation C.2 is similar to equation 3.23,

$$\begin{aligned} \pi &= rF + E_t[\ell_{t+1}]F + p^s_{t+1} \\ \implies E_t[\ell_{t+1}] &= r_s - r - \frac{p^s_{t+1}}{F}. \end{aligned} \tag{C.3}$$

This suggests another avenue for controlling the equilibrium collision rate. By setting the price of spectrum use, q_{t+1}, equal to the sum of marginal spectrum and collision risk congestion costs, $r_s'(S_{t+1}) + E_t[\xi(S_{t+1}, D_{t+1})]$, a spectrum regulator can implement an optimal stock control. More generally, this would be an optimal space traffic control in the sense of Johnson (2004), as it would account for both radio frequency and physical interference.

The same argument applies to price reductions due to downward-sloping demand. Some satellite-using industries, such as telecommunications in many regions, compete with terrestrial alternatives. Bertrand competition between satellite-provided and terrestrially-provided services, with the terrestrial alternatives acting as the low-price alternative, would place an upper bound on π. In regions where satellite-provided services are more expensive, π would be constant with respect to orbital congestion and debris growth. In industries where satellite-provided services compete with no or high-cost terrestrial alternatives, such as satellite imaging, the price effects of new entry would reduce collision risk and increase the sensitivity of the launch rate to the number of firms in orbit. Price changes on their own cannot prevent Kessler Syndrome and will not ensure optimal orbit use, but they still approximate a stock control as described in equation C.2.

C.2 Mandatory satellite insurance

Can insurance markets correct the orbital congestion externality in the absence of active debris removal? Suppose that satellites were required to be fully insured against loss once they reached orbit and the satellite insurance sector was perfectly competitive. The insurance payment will act as a stock control, so the only question remaining is how the insurance industry will price the product. Denote the price of insurance in period t by p_t, and the profits of the insurance sector by I_t.

$$Q(S_t, D_t, \ell_t, p_t) = \pi - \underbrace{p_t}_{\substack{\text{Insurance}\\\text{premium}}} + (1-\ell_t)F + \underbrace{\ell_t F}_{\substack{\text{Insurance}\\\text{payout}}} = \pi - p_t + F \tag{C.4}$$

$$I(S_t, \ell_t) = \underbrace{p_t S_t}_{\substack{\text{Inflow of}\\\text{premium payments}}} - \underbrace{\ell_t S_t F}_{\substack{\text{Outflow of}\\\text{reimbursements}}}. \tag{C.5}$$

Competitive insurance pricing

With competitive insurance pricing, satellite insurance will be actuarially fair. Plugging this price into the open access equilibrium condition, we can solve for the loss rate under mandatory insurance:

$$p_t : I(S_t, \ell_t) = 0 \implies p_t = \ell_t F \tag{C.6}$$

$$\pi = rF + E_t[\ell_{t+1}]F + p_{t+1} \tag{C.7}$$

$$\implies E_t[\ell_{t+1}] = r_s - r. \tag{C.8}$$

Proposition 19. *(Competitive insurance won't change collision risk) The equilibrium collision risk given mandatory satellite insurance with actuarially fair pricing is the same as the equilibrium collision risk given uninsured open access.*

Proof. From equation C.8, the equilibrium expected loss rate with actuarially fair insurance is

$$E_t[\ell_{t+1}] = r_s - r.$$

From equation 3.51, the equilibrium expected loss rate with no insurance is

$$E_t[\ell_{t+1}] = r_s - r.$$

$\square$

Regulated insurance pricing As in the case of spectrum management policies, mandatory satellite insurance premiums approximate a stock control. This suggests another avenue by which a regulator could induce optimal orbit use without assigning property rights over orbits or levying an explicit satellite tax.

Suppose the regulator was able to give insurers a per-satellite penalty or subsidy of τ_t to ensure insurance would be priced at the marginal external cost ($p_{t+1} = E_t[\xi(S_{t+1}, D_{t+1})]$) while still allowing free entry into the insurance sector. When τ_t is positive the regulator would be issuing an underwriting subsidy, and when τ_t is negative the regulator would be issuing an underwriting penalty. The insurance sector's profit is then

$$I(S_t, \ell_t) = (E_{t-1}[\xi(S_t, D_t)] - \ell_t F + \tau_t)S_t \tag{C.9}$$

$$\tau_t : I(S_t, \ell_t) = 0 \implies \tau_t = \ell_t F - E_{t-1}[\xi(S_t, D_t)]. \tag{C.10}$$

Equation C.10 shows that the socially optimal mandatory insurance pricing can be achieved by an incentive which imposes the difference between the actuarial cost of satellite insurance and the marginal external cost on the insurer. The insurer then passes the marginal external cost on to the satellite owner. Depending on the magnitude of the risk and the marginal external cost, this may be a net subsidy or tax on the insurer.

C.3 Competitive debris removal pricing

The profits of the cleanup industry, which supplies active debris removal, are

$$I_t(R_t) = \underbrace{c_t R_t}_{\substack{\text{Cleaning}\\\text{revenues}}} - \underbrace{\gamma R_t^2}_{\substack{\text{Cleaning}\\\text{costs}}} . \tag{C.11}$$

If the cleanup industry is competitive and a positive amount of debris is removed, debris will be removed from orbit until industry profits are zero,

$$R_t^s : I_t(R_t^s) = 0 \tag{C.12}$$

$$\implies R_t^s(c_t - \gamma R_t^s) = 0 \tag{C.13}$$

$$\implies R_t^s = \frac{c_t}{\gamma}, \tag{C.14}$$

subject to the constraint that $R_t^s \leq D_t$. Tkatchova (2018) examines the potential for debris removal markets.

Combining equation 3.56 with equation C.14 and the market clearing condition $R_t^s = R_t$, the aggregate amount of debris removed from orbit will be

$$R_t = \frac{\partial E_t[\ell_t | S_t, D_t - R_t]}{\partial D_t} \frac{F}{\gamma} S_t. \tag{C.15}$$

Appendix D

Full proofs and technical details for Chapter 3

D.1 Mapping between the deterministic and stochastic models

The dynamics of the satellite and debris stocks make the process for ℓ_t dependent and heterogeneously distributed over time. For exposition and to illustrate specific points (most often transition dynamics), I simulate deterministic versions of the models described in this paper. Using deterministic models simplifies making "apples-to-apples" comparisons between different policies without collision rate draws complicating matters. Making the collision risk process independent and identically distributed would remove the economically interesting physics of the problem - open access becomes socially optimal because launch and removal choices have no impact on the collision risk. In the deterministic model, I use the expected collision risk without drawing from a distribution. In the stochastic model, I draw the collision risk from a binomial distribution with floor(S_t) many trials and mean equal to $\min\{L(S,D),1\}$. Formally,

$$\text{Deterministic model collision risk: } \ell_t = L(S,D), \tag{D.1}$$

$$\text{Stochastic model collision risk: } \ell_t \sim Bin(\text{floor}(S_t), L(S,D)). \tag{D.2}$$

Figure D.1 compares sequences of launch rates, satellite and debris stocks, and expected collision risks under open access in the deterministic and stochastic models. The deterministic model behaves as expected - it is the mean of the stochastic model. The expected collision risk in each model is identical, which is unsurprising given that open access and the planner both target the expected collision risk.

Although the time paths are similar, there are some interesting properties of the stochastic values and policies (such as Lemma 5) which are not captured by the deterministic ones.

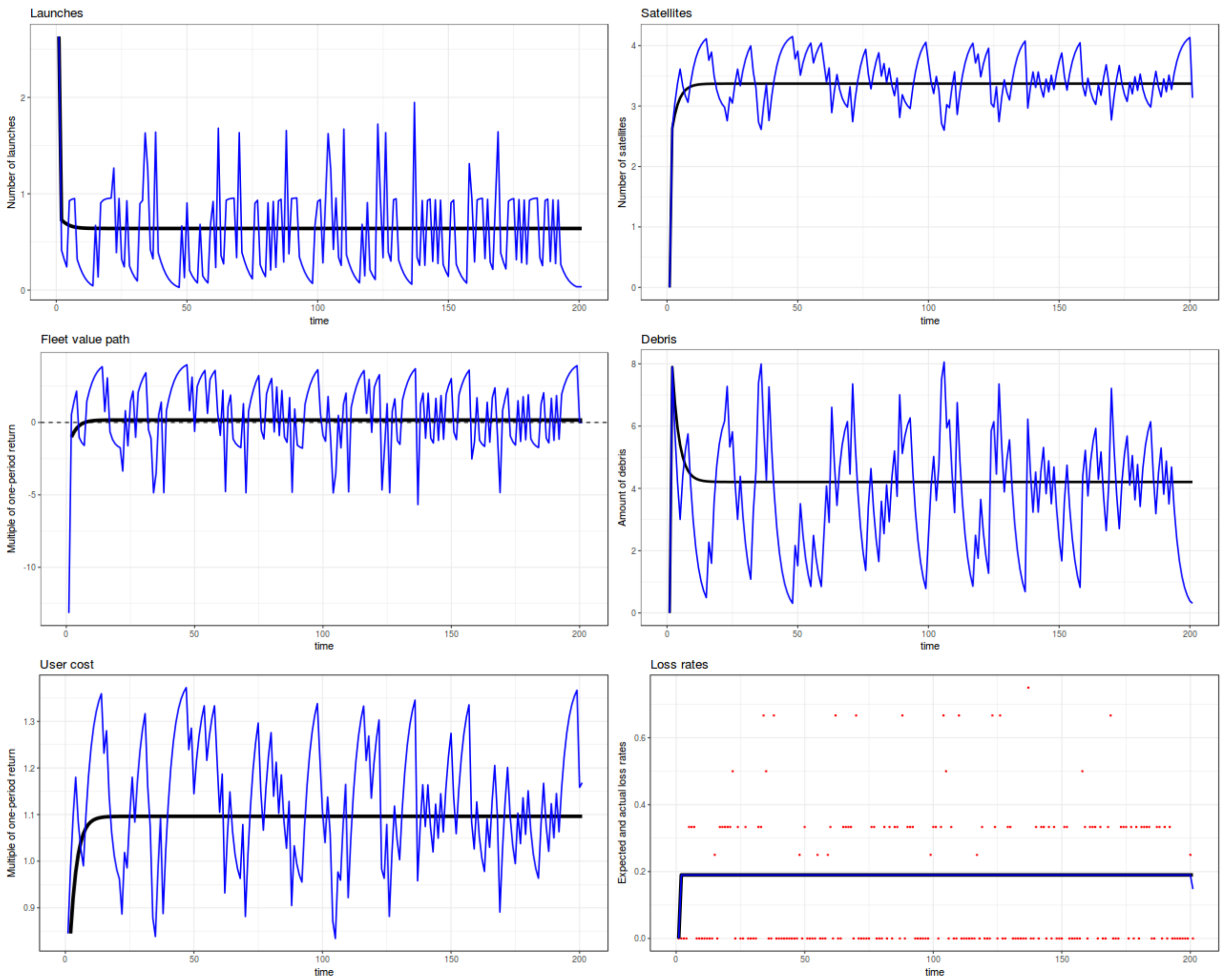

Figure D.1: Comparison of stochastic and deterministic open access time paths

Time paths under the stochastic (blue line) and deterministic (black line) models.
The red dots in the "collision rates" panel are the draws of ℓ_t.

D.2 Technical assumptions and lemmas

Assumption 6. *Let S be a vector of state variables, and I_k be a vector of same size as S with 1 in the k^{th} position and 0 in all other positions. $\phi(\ell|S)$ is a conditional density which satisfies the following properties:*

(1) The derivative of $\phi(\ell|S)$ with respect to the k^{th} argument of S,

$$\frac{\partial \phi(\ell|S)}{\partial S_k} = \lim_{h \to 0} \frac{\phi(\ell|S + I_k h) - \phi(\ell|S)}{h} \equiv \phi_S(\ell|S),$$

exists and is bounded $\forall \ell$, with $\phi_S(\ell|S) \neq 0$ for some ℓ, $\forall S$.

(2) Let S^1 and S^2 be two vectors which are identical except for the k^{th} entry, where $I_k S^1 < I_k S^2$. $\forall S^1, S^2$, and $\forall A \in [0,1]$, define $\bar{\ell}_1 \equiv \bar{\ell}(A, S^1), \bar{\ell}_2 \equiv \bar{\ell}(A, S^2)$ such that

$$\int_0^{\bar{\ell}_1} \phi_S(\ell|S^1) d\ell = \int_0^{\bar{\ell}_2} \phi_S(\ell|S^2) d\ell = A.$$

Then $\phi(\ell|S)$ satisfies a Lipschitz condition

$$\left|\left| \int_{\bar{\ell}_1}^{\bar{\ell}_2} \phi(\ell|S^2) d\ell \right|\right| < ||S^2 - S^1||,$$

The first condition places a lower bound on the change in collision probability from new satellite placements and ensures some smoothness for the changes in the density across the support.

The second condition places an upper bound on changes to the density $\phi(\ell|S, D)$ over the space of (S, D). The idea is that the additional area under the new density required to achieve a target area under the old density is bounded by the change in S or D. Physically, this requires that new satellites or debris will not be placed in orbits that will cause a drastic change in the collision probability. Rather, the change in collision probability from new launches should be bounded and proportional to the number of new satellites placed in orbit. Together, the physical implication of these two conditions is that new satellites or debris will cause some changes to the collision probability, but that those changes will be bounded across the possible outcomes. This is economically reasonable for satellites - a violation of this implies that firms are deliberately placing their satellites in risky orbits. This may be less reasonable for debris, since the orbits of debris objects resulting from collisions are uncontrolled and difficult to predict. These conditions facilitate the proofs of the lemmas below, but are not crucial to the main results of the paper.

Note that the proofs of the lemmas below often assume uniformly bounded functions. While no such property is proven for the value functions studied, realistic parameter choices should guarantee the existence of uniform bounds on the value functions.

Lemma 2. *(Measurable functions under changes in distribution) Let ℓ be a random variable with a conditional density $\phi(\ell|S)$ defined on the compact interval $[a,b]$ and with range $[r(a),r(b)]$. Let $f(\cdot)$: $[r(a),r(b)] \to [f(a),f(b)]$ be a measurable function of ℓ. Then*

$$\int_a^b f(\ell)\frac{\partial\phi(\ell|S)}{\partial S}d\ell = \frac{\partial E[f(\ell)|S]}{\partial S}$$

Proof.

$$\begin{aligned}
\int_a^b f(\ell)\frac{\partial\phi(\ell|S)}{\partial S}d\ell &= \int_a^b f(\ell)\lim_{h\to 0}\frac{\phi(\ell|S+h)-\phi(\ell|S)}{h}d\ell \\
&= \lim_{h\to 0}\frac{1}{h}\left(\int_a^b f(\ell)\phi(\ell|S+h)d\ell - \int_a^b f(\ell)\phi(\ell|S)d\ell\right) \\
&= \lim_{h\to 0}\frac{1}{h}\left(E[f(\ell)|S+h]-E[f(\ell)|S]\right) \\
&= \frac{\partial E[f(\ell)|S]}{\partial S}.
\end{aligned}$$

□

Lemma 3. $\frac{\partial E[f(x)|S]}{\partial S} = 0$ *$\forall S$ and $\forall f(x)$ which do not depend on ℓ, the argument of $\phi(\ell|S)$.*

Proof. From Assumption 6 and Lemma 2,

$$\begin{aligned}
\frac{\partial E[f(x)|S]}{\partial S} &= \int_0^1 f(x)\left[\lim_{h\to 0}\frac{\phi(\ell|S+h)-\phi(\ell|S)}{h}\right]d\ell \\
&= f(x)\lim_{h\to 0}\frac{1}{h}\left[\int_0^1 \phi(\ell|S+h)d\ell - \int_0^1 \phi(\ell|S)d\ell\right] \\
&= f(x)\lim_{h\to 0}\frac{1}{h}[1-1] = 0.
\end{aligned}$$

□

Lemma 4. *If $f(\ell)$ is a nonnegative and uniformly bounded function, then under assumption 3*

(1) $\frac{\partial E[f(\ell)|S]}{\partial S} = 0$ *if* $\frac{\partial f(\ell)}{\partial \ell} = 0 \;\; \forall \ell$

(2) $\frac{\partial E[f(\ell)|S]}{\partial S} < 0$ *if* $\frac{\partial f(\ell)}{\partial \ell} < 0 \;\; \forall \ell$

(3) $\frac{\partial E[f(\ell)|S]}{\partial S} > 0$ *if* $\frac{\partial f(\ell)}{\partial \ell} > 0 \;\; \forall \ell$

Proof. For simplicity, the proof is written for a scalar-valued S. Extending the argument to vector-valued S is possible but not particularly informative.

The first statement, $\frac{\partial E[f(\ell)|S]}{\partial S} = 0$ if $\frac{\partial f(\ell)}{\partial \ell} = 0$, follows directly from Lemma 3 and the assumption that $f(\ell)$ is constant $\forall \ell$.

To show that $\frac{\partial E[f(\ell)|S]}{\partial S} < 0$ if $\frac{\partial f(\ell)}{\partial \ell} < 0 \;\; \forall \ell$, without any loss of generality let $S^2 > S^1$. Pick $\bar{\ell}_1, \bar{\ell}_2$: $\int_0^{\bar{\ell}_1} \phi_S(\ell|S^1)d\ell = \int_0^{\bar{\ell}_2} \phi_S(\ell|S^2)d\ell = A \in (0,1)$. Note that when $A = 0, \bar{\ell}_1 = \bar{\ell}_2 = 0$, and when $A = 1, \bar{\ell}_1 = \bar{\ell}_2 = 1, \;\; \forall S^1, S^2$. Assumption 3 implies that $\forall A \in (0,1)$, $\bar{\ell}_2 > \bar{\ell}_1$. Since $\frac{\partial f(\ell)}{\partial \ell} < 0 \;\; \forall \ell$,

$$
\begin{aligned}
& \int_0^{\bar{\ell}_1} f(\ell)\phi_S(\ell|S^1)d\ell > \int_0^{\bar{\ell}_2} f(\ell)\phi_S(\ell|S^2)d\ell \\
\Longrightarrow & \int_0^{\bar{\ell}_1} f(\ell)\phi_S(\ell|S^1)d\ell - \int_0^{\bar{\ell}_2} f(\ell)\phi_S(\ell|S^2)d\ell > 0 \\
\Longrightarrow & \int_0^{\bar{\ell}_2} f(\ell)\phi_S(\ell|S^2)d\ell - \int_0^{\bar{\ell}_1} f(\ell)\phi_S(\ell|S^1)d\ell < 0 \\
\Longrightarrow & \lim_{S^2 \to S^1} \frac{1}{S^2 - S^1}\left[\int_0^{\bar{\ell}_2} f(\ell)\phi_S(\ell|S^2)d\ell - \int_0^{\bar{\ell}_1} f(\ell)\phi_S(\ell|S^1)d\ell\right] \\
& = \lim_{S^2 \to S^1} \frac{1}{S^2 - S^1}\left[\int_0^{\bar{\ell}_1} f(\ell)\{\phi_S(\ell|S^2)d\ell - \phi_S(\ell|S^1)\}d\ell + \int_{\bar{\ell}_1}^{\bar{\ell}_2} f(\ell)\phi(\ell|S^2)d\ell\right] \\
& = \int_0^{\bar{\ell}_1} f(\ell) \lim_{S^2 \to S^1}\left\{\frac{\phi_S(\ell|S^2)d\ell - \phi_S(\ell|S^1)}{S^2 - S^1}\right\}d\ell + \lim_{S^2 \to S^1}\int_{\bar{\ell}_1}^{\bar{\ell}_2} f(\ell)\frac{\phi(\ell|S^2)}{S^2 - S^1}d\ell \\
& = \int_0^{\bar{\ell}_1} f(\ell)\phi_S(\ell|S^1)d\ell + \lim_{S^2 \to S^1}\int_{\bar{\ell}_1}^{\bar{\ell}_2} f(\ell)\frac{\phi(\ell|S^2)}{S^2 - S^1}d\ell < 0
\end{aligned}
$$

By assumption 6 and $f(\ell) \geq 0$, $S^2 > S^1$, $\phi(\ell|S^2) \geq 0$,

$$
\int_{\bar{\ell}_1}^{\bar{\ell}_2} f(\ell)\frac{\phi(\ell|S^2)}{S^2 - S^1}d\ell \geq 0.
$$

Now, taking the limit as A goes to 1, we get

$$\lim_{A\to 1}\int_0^{\bar{\ell}_1} f(\ell)\phi_S(\ell|S^1)d\ell + \lim_{S^2\to S^1}\lim_{A\to 1}\int_{\bar{\ell}_1}^{\bar{\ell}_2} f(\ell)\frac{\phi(\ell|S^2)}{S^2-S^1}d\ell < 0,$$

where

$$\lim_{A\to 1}\int_{\bar{\ell}_1}^{\bar{\ell}_2} f(\ell)\frac{\phi(\ell|S^2)}{S^2-S^1}d\ell = O(\bar{\ell}_2-\bar{\ell}_1)$$

is a nonnegative remainder term is bounded by $\bar{\ell}_2-\bar{\ell}_1$. This leaves us with

$$\lim_{A\to 1}\left[\int_0^{\bar{\ell}_1} f(\ell)\phi_S(\ell|S^1)d\ell\right] + \lim_{S^2\to S^1} O(\bar{\ell}_2-\bar{\ell}_1) < 0.$$

When $A=1, \bar{\ell}_2=\bar{\ell}_1=1$ and the remainder is exactly 0 $\forall S^1, S^2$. Since S^1 was chosen arbitrarily, we can therefore say that

$$\int_0^1 f(\ell)\phi_S(\ell|S)d\ell \equiv \frac{\partial E[f(\ell)|S]}{\partial S} < 0 \text{ if } \frac{\partial f(\ell)}{\partial \ell} < 0 \;\; \forall \ell.$$

Repeating the argument above with $f(\ell)$ strictly increasing instead of decreasing yields the third statement, $\frac{\partial E[f(\ell)|S]}{\partial S} > 0$ if $\frac{\partial f(\ell)}{\partial \ell} > 0 \;\; \forall \ell$. □

D.3 Proofs

Lemma 1 (Launch response to stock and flow controls): The open access launch rate is

- **decreasing in the future price of a stock control;**
- **decreasing in the current price and increasing in the future price of a flow control.**

Proof. Stock controls: From equation 3.23, we can write

$$\mathscr{I} = \pi - rF - E_t[\ell_{t+1}]F - p_{t+1} = 0. \tag{D.3}$$

Applying the Implicit Function Theorem, we get that

$$\frac{\partial X_t}{\partial p_{t+1}} = -\frac{\partial \mathcal{I}/\partial p_{t+1}}{\partial \mathcal{I}/\partial X_t} \tag{D.4}$$

$$= -\frac{-1}{-\frac{\partial E[\ell_{t+1}]}{\partial X_t}} \tag{D.5}$$

$$= -\frac{1}{\frac{\partial E_t[\ell_{t+1}]}{\partial S_{t+1}}\frac{\partial S_{t+1}}{\partial X_t} + \frac{\partial E_t[\ell_{t+1}]}{\partial D_{t+1}}\frac{\partial D_{t+1}}{\partial X_t}} \tag{D.6}$$

$$= -\frac{\partial E_t[\ell_{t+1}]}{\partial S_{t+1}} + m\frac{\partial E_t[\ell_{t+1}]}{\partial D_{t+1}} < 0. \tag{D.7}$$

Flow controls: From equation 3.26, we can write

$$\mathcal{G} = \pi - rF - E_t[\ell_{t+1}]F - (1+r)p_t + E_t[(1-\ell_{t+1})p_{t+1}] = 0. \tag{D.8}$$

Applying the Implicit Function Theorem, we get that

$$\frac{\partial X_t}{\partial p_t} = -\frac{\partial \mathcal{G}/\partial p_t}{\partial \mathcal{G}/\partial X_t} \tag{D.9}$$

$$= -\frac{-(1+r)}{-\frac{\partial E_t[\ell_{t+1}]}{\partial X_t}F - \frac{\partial E_t[\ell_{t+1}]}{\partial X_t}p_{t+1}} \tag{D.10}$$

$$= -\frac{1+r}{[\frac{\partial E_t[\ell_{t+1}]}{\partial S_{t+1}}\frac{\partial S_{t+1}}{\partial X_t} + \frac{\partial E_t[\ell_{t+1}]}{\partial D_{t+1}}\frac{\partial D_{t+1}}{\partial X_t}](F + p_{t+1})} \tag{D.11}$$

$$= -\frac{1+r}{[\frac{\partial E_t[\ell_{t+1}]}{\partial S_{t+1}} + m\frac{\partial E_t[\ell_{t+1}]}{\partial D_{t+1}}](F + p_{t+1})} < 0. \tag{D.12}$$

Similarly, we can obtain

$$\frac{\partial X_t}{\partial p_{t+1}} = -\frac{\partial \mathcal{G}/\partial p_{t+1}}{\partial \mathcal{G}/\partial X_t} \tag{D.13}$$

$$= -\frac{1 - E_t[\ell_{t+1}]}{-\frac{\partial E_t[\ell_{t+1}]}{\partial X_t}F - \frac{\partial E_t[\ell_{t+1}]}{\partial X_t}p_{t+1}} \tag{D.14}$$

$$= \frac{1 - E_t[\ell_{t+1}]}{[\frac{\partial E_t[\ell_{t+1}]}{\partial S_{t+1}}\frac{\partial S_{t+1}}{\partial X_t} + \frac{\partial E_t[\ell_{t+1}]}{\partial D_{t+1}}\frac{\partial D_{t+1}}{\partial X_t}](F + p_{t+1})} \tag{D.15}$$

$$= \frac{1 - E_t[\ell_{t+1}]}{[\frac{\partial E_t[\ell_{t+1}]}{\partial S_{t+1}} + m\frac{\partial E_t[\ell_{t+1}]}{\partial D_{t+1}}](F + p_{t+1})} > 0. \tag{D.16}$$

□

Proposition 12 (Smoothness at boundaries): Stock controls can be initiated without letting the launch rate exceed the open access launch rate. Flow controls cannot be initiated without forcing the launch rate to exceed the open access launch rate.

Proof. In both cases, I suppose that there is open access before the control is initiated.

Initiating a stock control: Suppose a stock control is scheduled to take effect at t, that is, satellite owners in t begin paying p_t. In $t-1$, firms would launch with this fact in mind:

$$X_{t-1} : \pi = rF + E_{t-1}[\ell_t]F + p_t. \tag{D.17}$$

Let the open access launch rate in $t-1$ with no stock control in t be $\hat{X}_{t-1} : \pi = rF + E_{t-1}[\ell_t]F$. Lemma 1 implies that for all $p_t > 0$, $\hat{X}_{t-1} > X_{t-1}$.

Initiating a flow control: Suppose a flow control is scheduled to be implemented at t, that is, satellite launchers in t begin paying p_t to launch. In $t-1$, firms would launch with this fact in mind:

$$X_{t-1} : \pi = rF + E_{t-1}[\ell_t]F - (1 - E_{t-1}[\ell_t])p_t. \tag{D.18}$$

Let the open access launch rate in $t-1$ with no flow control implemented in t be $\hat{X}_{t-1} : \pi = rF + E_{t-1}[\ell_t]F$. Lemma 1 implies that for all $p_t > 0$, $\hat{X}_{t-1} < X_{t-1}$.

□

Proposition 13 (Controlling the rate of deorbit): Stock controls with positive prices can make satellite owners deorbit their satellites and induce net deorbits. Flow controls with positive prices cannot make satellite owners deorbit their satellites or induce net deorbits.

Proof. A satellite owner facing a stock control in period t will deorbit if

$$p_t > \pi + (1 - E_t[\ell_{t+1}])F - V^d. \tag{D.19}$$

The regulator can induce firms to deorbit in t by raising p_t high enough in $t-1$. A potential launcher in $t-1$ will not launch if

$$p_t > \pi + (1 - E_{t-1}[\ell_t])F. \tag{D.20}$$

By raising p_t high enough, the regulator can both discourage further launches and induce existing satellite owners to deorbit their satellites.

A satellite owner facing a flow control in period t will deorbit if

$$p_t : \pi + (1-\ell_t)\beta E_t[Q_{t+1}] < V^d, \text{ where } X_t : \beta E_t[Q_{t+1}] = F + p_t \tag{D.21}$$

$$\implies (1-\ell_t)(F+p_t) < V^d - \pi, \tag{D.22}$$

which cannot be satisfied by positive p_t, given $V^d < \pi$. A potential launcher in t will not launch if

$$p_{t+1} : \ \pi + (1 - E_t[\ell_{t+1}])p_{t+1} < rF + E_t[\ell_{t+1}]F + (1+r)p_t. \tag{D.23}$$

If $\pi < rF + E_t[\ell_{t+1}]F + (1+r)p_t$, equation D.23 will not be satisfied for any positive p_{t+1}. Although equation D.23 can be satisfied if p_{t+1} is sufficiently negative, this would require the regulator to commit to a path of ever-decreasing negative prices as long as they wished to prevent launches (as described earlier and in Lemma 6). Regardless, the regulator cannot induce net deorbits (no new arrivals and some deorbits) in $t+1$ with a positive p_{t+1}. □

Proposition 17 (ADR can reduce collision risk):

Proof. I show the result first for the introduction of debris removal services, then for the ongoing use of debris removal services.

The introduction of ADR: Suppose an active debris removal service will become available at date t. To clarify whether removal is an option or not, I explicitly include the conditioning variables in the loss rate, that is, $E_t[\ell_{t+1}]$ is written as $E_t[\ell_{t+1}|S_{t+1}, D_{t+1} - R_{t+1}]$ when removal is an option.

Under open access, firms without satellites at $t-2$ will launch until

$$E_{t-2}[\ell_{t-1}|S_{t-1}, D_{t-1}] = r_s - r. \tag{D.24}$$

At $t-1$, launchers will expect to be able to remove debris once their satellites reach orbit. They will launch until

$$E_{t-1}[\ell_t|S_t, D_t - R_t] = r_s - r - \frac{c_t}{F} R_{it}. \tag{D.25}$$

Comparing $E_{t-2}[\ell_{t-1}|S_{t-1},D_{t-1}]$ and $E_{t-1}[\ell_t|S_t,D_t - R_t]$ yields the necessary and sufficient conditions:

$$E_{t-2}[\ell_{t-1}|S_{t-1},D_{t-1}] - E_{t-1}[\ell_t|S_t,D_t - R_t] > 0 \iff c_t R_{it} > 0. \tag{D.26}$$

Ongoing use of ADR: Under open access, the equilibrium collision risk in $t+1$ after debris removal in t is

$$E_t[\ell_{t+1}] = r_s - r - \frac{c_{t+1}}{F} R_{it+1}.$$

Similarly, the equilibrium collision risk in t after debris removal in $t-1$ is

$$E_{t-1}[\ell_t] = r_s - r - \frac{c_t}{F} R_{it}.$$

Subtracting one equilibrium risk from the other yields the necessary and sufficient condition for ongoing debris removal to continue to reduce the collision risk:

$$\begin{aligned} E_{t-1}[\ell_t] - E_t[\ell_{t+1}] > 0 &\iff r_s - r - \frac{c_t}{F} R_{it} - (r_s - r - \frac{c_{t+1}}{F} R_{it+1}) > 0 \\ &\iff c_{t+1} R_{it+1} > c_t R_{it}. \end{aligned}$$

□

D.4 Cost and congestion shifts in cooperative removal demands from new satellites

For brevity, I write $E_t[\ell_{t+1}]$ as $L(S_t, D_t - S_t R_{it})$ in this subsection and use S and D subscripts to indicate the respective partial derivatives. Since these results are intratemporal in nature, I also drop time subscripts.

To see the cost and congestion shifts in the cooperative private demand for removal, suppose there were two types of satellites, Kinds (k) and Unkinds (u). Kinds make sure their satellites can never collide with others and purchase debris removal, while Unkinds allow their satellites a non-zero chance of colliding with another satellite and never purchase debris removal. The private marginal benefit of debris removal for a Kind named i is

$$MB_i = L_D(u, D - kR_i)kF. \tag{D.27}$$

The congestion shift is the effect of another Unkind entering, while the cost shift is the effect of another Kind entering. Formally,

$$\text{Cost shift: } \frac{\partial MB_i}{\partial k} = -L_{DD}(u, D-kR_i)kR_iF + L_D(u, D-kR_i)F \tag{D.28}$$

$$\text{Congestion shift: } \frac{\partial MB_i}{\partial u} = L_{DS}(u, D-kR_i)kF \tag{D.29}$$

Adding them and normalizing by a function of the collisions rate's in debris,

$$\frac{\frac{\partial MB_i}{\partial k} + \frac{\partial MB_i}{\partial u}}{L_{DD}k^2F} = \frac{(-L_{DD}kR_iF + L_DF) + (L_{DS}kF + L_DF)}{L_{DD}k^2F} \tag{D.30}$$

$$= \frac{L_D}{L_{DD}k^2} + \frac{L_{DS} - L_{DD}R_i}{L_{DD}k} = \frac{\partial R_i}{\partial S}. \tag{D.31}$$

The congestion shift may be positive or negative. It is the effect of increasing the number of satellites on the marginal collision risk from a unit of debris. If the collision rate were decoupled from the satellite stock, the congestion shift would disappear. If a marginal Unkind would increase the effect of a marginal unit of debris, the congestion shift will be positive. Reducing the amount of debris would greatly reduce the threat posed by the marginal satellite. If a marginal Unkind would decrease the effect of a marginal unit of debris, the congestion shift will be negative. This could be the case if the Unkind was well-shielded from debris but a threat to other satellites. Reducing the amount of debris would not change the risk of the marginal Unkind by much then.

The cost shift is the effect of increasing the number of customers in the market for debris removal on the marginal collision threat from a unit of debris. There are two pieces to this. First, the debris removed by each Kind reduces the collision risk for all owners. As long as the collision rate is increasing in debris, reducing debris is always a good thing for everyone. This will tend to make the cost shift positive. Second, the debris removed by each Kind changes the marginal benefit of the next Kind's removal. Since the collision rate must be locally convex in debris at an interior solution, this effect will tend to make the cost shift negative. If the collision rate is sufficiently locally convex, this effect can make the cost shift negative in total. Generic satellites are both Kinds and Unkinds.

D.5 Open access launch response to collisions

Lemma 5. *(Open access launch response to collision risk draws) The open access launch rate in t can be non-monotonic in the realized collision risk in t.*

Proof. From equation 3.11,

$$\mathscr{F} = r_s - r - E_t[\ell_{t+1}] = 0. \tag{D.32}$$

Applying the Implicit Function Theorem to $\mathscr{F}$,

$$\frac{\partial X_t}{\partial \ell_t} = -\frac{\partial \mathscr{F}/\partial \ell_t}{\partial \mathscr{F}/\partial X_t} \tag{D.33}$$

$$= \frac{\frac{\partial E_t[\ell_{t+1}]}{\partial S_{t+1}} S_t + \frac{\partial E_t[\ell_{t+1}]}{\partial D_{t+1}} \frac{\partial G(S_t, D_t, \ell_t)}{\partial \ell_t}}{\frac{\partial E_t[\ell_{t+1}]}{\partial S_{t+1}} + m \frac{\partial E_t[\ell_{t+1}]}{\partial D_{t+1}}} \lesseqgtr 0 \tag{D.34}$$

□

Figure D.2 illustrates three cases of Lemma 5: one where the open access launch rate is first decreasing and then increasing in the satellite-destroying collision risk draw, another where it is uniformly decreasing in the collision risk draw, and a third where it is uniformly increasing in the collision risk draw. There are two competing effects of collisions driving this behavior: collisions generate debris, but collisions also remove other satellites from orbit.

In the first case, the effect of additional debris dominates the launch decision when the collision risk draw is low, and the effect of fewer satellites dominates when the collision risk draw is high. In the second case shown in Figure D.2, the new-debris effect dominates for all draws. In the third case shown in Figure D.2, the fewer-satellites effect dominates for all draws. The third case is the least-realistic, as it implies that the number of fragments from a satellite destruction is tiny compared to the number of fragments created by a debris-debris collision. Debris modeling studies such as Liou (2006); Letizia et al. (2017) find the opposite: satellite destructions contribute much more debris to the orbital environment than collisions between debris fragments. The first and second cases are plausible under realistic relative orders of magnitude between the number of fragments created by satellite-satellite, satellite-debris, and debris-debris collisions.

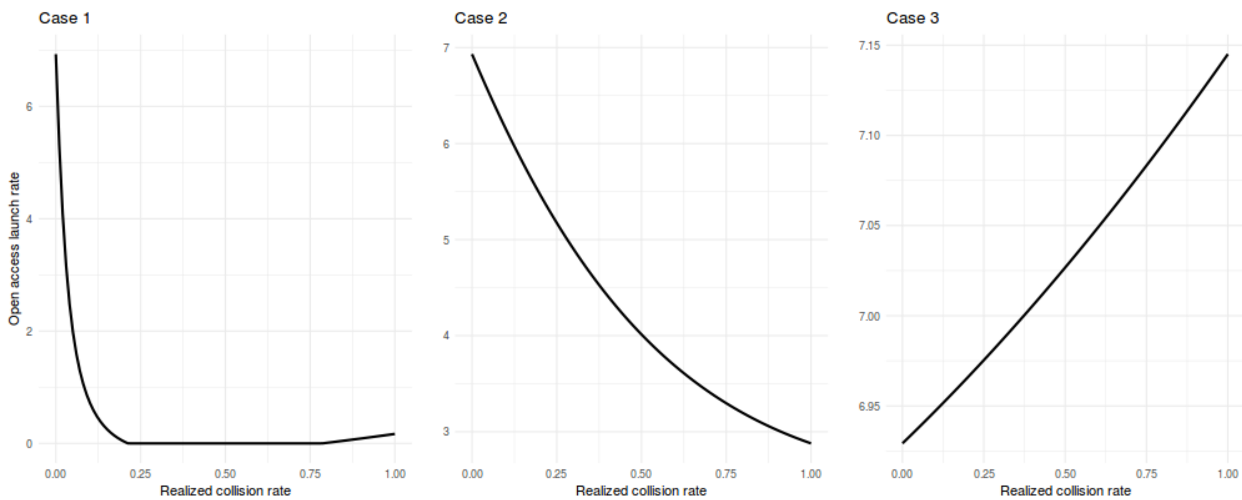

Figure D.2: *Three ways the open access launch rate may respond to collision probability draws.*
Left panel: The open access launch rate is decreasing then increasing in the collision risk draw. The effect of additional debris dominates the launch decision when the collision risk draw is low, and the effect of fewer satellites dominates when the collision risk draw is high.
Middle panel: The open access launch rate is uniformly decreasing in the collision risk draw. The new-debris effect dominates for all draws.
Right panel: The open access launch rate is uniformly increasing in the collision risk draw. The fewer-satellites effect dominates for all draws.
The left panel and middle panel cases are more plausible than the right panel case under realistic relative orders of magnitude between the number of fragments created by satellite-satellite, satellite-debris, and debris-debris collisions.

D.6 Two details for flow control implementations

Lemma 6. *(A feature of flow controls with no end date) If the discount rate is positive, the price of a time-consistent infinite-horizon flow control is an exploding process.*

Proof. From equation 3.26, we can write the price in period $t+1$ as a function of the price in period t as

$$p_{t+1} = \frac{1+r}{1-E_t[\ell_{t+1}]} p_t - \frac{\pi - rF - E_t[\ell_{t+1}]F}{1-E_t[\ell_{t+1}]}. \tag{D.35}$$

Differentiating with respect to p_t,

$$\frac{\partial p_{t+1}}{\partial p_t} = \frac{1+r}{1-E_t[\ell_{t+1}]} > 1 \ \ \forall r > 0. \tag{D.36}$$

This holds along a generic equilibrium path as shown above, or along an optimal path. Along an optimal path, the numerator of the second term on the right-hand side would be replaced with the marginal external cost of a satellite in period $t+1$. Since the first term on the right-hand side is unchanged, the flow control process still has a unit or greater-than-unit root. $\square$

Lemma 6 is likely to be practically relevant to designing a flow control policy, even though in theory such a control could be made optimal.[1]

Lemma 7. *(A limitation of flow controls) If the expected future loss rate is one, there is no value of the future flow control price which can affect the current launch rate.*

Proof. Rewriting equation 3.26 with $E_t[\ell_{t+1}] = 1$, the launch rate will satisfy

$$X_t : \pi = rF + E_t[\ell_{t+1}]F + (1+r)p_t. \tag{D.37}$$

Since equation D.37 no longer contains p_{t+1}, the regulator cannot use it to affect X_t. $\square$

I present this result for completeness, though it is likely a moot point. The destruction of all active satellites in an orbit would likely trigger Kessler Syndrome there. The regulator and the space industry

[1] Though it seems unlikely to me, perhaps there exists or could exist a regulatory body with the credibility to commit to an exploding price path for the foreseeable future; they would find this result irrelevant.

would then have bigger problems than whether or not a flow control in line with the pre-committed path exists. This may be relevant for low orbits where debris decays sufficiently quickly. If such destruction occurred in those orbits intentionally, for example missile tests or conflict, the regulator and space industry would again have bigger problems than deviation from prior commitments.

D.7 Incentives to cooperate in debris removal

Suppose all satellite owners agree to cooperate and individually purchase R_{it}^* units of debris removal. Owner i considers deviating and reducing her removal demands by $\varepsilon \in (0, R_{it}^*]$. Her payoff from not deviating by ε is

$$\underbrace{\pi - c_t R_{it}^* + (1 - \tilde{E}_t[\ell_t | S_t, D_t - R_t^*])F}_{\text{Expected value of cooperating}} - [\underbrace{\pi - c_t(R_{it}^* - \varepsilon) + (1 - \tilde{E}_t[\ell_t | S_t, D_t - (R_t^* - \varepsilon)])F}_{\text{Expected value of deviating by } \varepsilon}]$$
$$= -\varepsilon c_t + [\tilde{E}_t[\ell_t | S_t, D_t + \varepsilon - R_t^*] - \tilde{E}_t[\ell_t | S_t, D_t - R_t^*]]F.$$

Cooperation is a strictly dominant Nash equilibrium if and only if

$$\tilde{E}_t[\ell_t | S_t, D_t + \varepsilon - R_t^*] - \tilde{E}_t[\ell_t | S_t, D_t - R_t^*] > \varepsilon \frac{c_t}{F} \quad \forall \varepsilon \in (0, R_{it}^*]. \tag{D.38}$$

Proposition 20 establishes an intuitive necessary and sufficient condition for cooperation to strictly dominate small deviations. If the change in the expected loss rate before removal is greater than the ratio of the cost of removal in t to the cost of launching a satellite in t, then a tiny deviation will cost a satellite owner more expected value through a lower survival rate than it will yield in a removal expenditure savings.

Proposition 20. *(Local stability of cooperation) Cooperation with any non-zero debris removal plan strictly dominates small deviations if the change in the equilibrium collision risk from another unit of debris is greater than the ratio of the removal price to the launch cost,*

$$\left.\frac{\partial \tilde{E}[\ell_t | S_t, \bar{D} - R_t]}{\partial \bar{D}}\right|_{\bar{D} = D_t} > \frac{c_t}{F}.$$

Proof. Cooperation with a non-zero debris removal plan is robust to all deviations ε for which

$$\tilde{E}_t[\ell_t | S_t, D_t + \varepsilon - R_t^*] - \tilde{E}_t[\ell_t | S_t, D_t - R_t^*] > \varepsilon \frac{c_t}{F}.$$

Cooperation with a non-zero debris removal plan strictly dominates small deviations if

$$\lim_{\varepsilon \to 0} \frac{\tilde{E}_t[\ell_t|S_t, D_t + \varepsilon - R_t] - \tilde{E}_t[\ell_t|S_t, D_t - R_t]}{\varepsilon} > \frac{c_t}{F} \quad \forall R_t > 0$$
$$\implies \left. \frac{\partial \tilde{E}[\ell_t|S_t, \bar{D} - R_t]}{\partial \bar{D}} \right|_{\bar{D} = D_t} > \frac{c_t}{F} \quad \forall R_t > 0.$$

□

The following proposition establishes that the debris removal solution described in equation 3.56 is in fact the cooperative private debris removal solution.

Proposition 21. *(A cooperative private removal plan) The debris removal solution described by equation 3.56 is maximizes the value of the currently-orbiting satellite fleet, given open access in that period.*

Proof. Given open access launch rates, the value of a satellite already in orbit is

$$Q_i(S, D) = \pi - cR_i + (1 - \tilde{E}[\ell])F.$$

Given open access launch rates, the value of all satellites already in orbit is

$$W(S, D) = \int_0^S Q_i di$$
$$= \pi S - cR + (1 - \tilde{E}[\ell])FS.$$

Equation 3.56 is the first-order condition for the firm's problem,

$$Q_i(S, D) = \max_{0 \leq R_i \leq D/S} \{\pi - cR_i + (1 - \tilde{E}[\ell])F\}.$$

A constrained planner who maximizes the value of the currently-orbiting satellite fleet, taking open access to orbit as given, solves

$$W(S, D) = \max_{0 \leq R \leq D} \{\pi S - cR + (1 - \tilde{E}[\ell])SF\}$$
$$= \max_{\{0 \leq R_i \leq D/S\}_i} S\{\pi - cR_i + (1 - \tilde{E}[\ell])F\}$$
$$= \max_{\{0 \leq R_i \leq D/S\}_i} SQ_i(S, D).$$

The constrained planner's objective function is an individual satellite owner's objective scaled by the current size of the fleet, which the constrained planner takes as given. The individual removal solution given by equation 3.56 therefore characterizes a cooperative debris removal solution, where each firm behaves as an open-access-constrained social planner would command. □

D.8 Nonconvexities and corner solutions

For brevity, I write $E_t[\ell_{t+1}]$ as $L(S_t, D_t - S_t R_{it})$ in this subsection and use S and D subscripts to indicate the respective partial derivatives. Since these results are intratemporal in nature, I also drop time subscripts.

Upper bounds on damages, nonconvex stock decay rates, and complementarities between stocks in damage production each imply nonconvexities in the marginal benefits of abatement. The static private marginal benefits from abatement reflect two of these features:

(1) at most all satellites can be destroyed in collisions, implying the upper bound on the rate of satellite-destroying collisions;

(2) the marginal effect of debris on the number of satellite-destroying collisions depends on the number of satellites in orbit, which makes collisional complementarity or substitutability between satellites and debris possible.

When the satellite and debris couplings in the collision rate depend on each other, that is, $L_{SD} \neq 0$, changes in the satellite stock can change the returns to scale for debris removal. The dynamic benefits of debris abatement also include the effect of fragment growth from collisions between debris. This effect implies that the net marginal rate of debris decay ($\delta - G_D(S, D-R)$) can be negative.

The marginal benefit of removal is the private value of reducing the probability of a satellite-destroying collision. Debris removal has diminishing marginal benefits if and only if the collision rate is strictly convex in debris. The upper bound on $L(S,D)$ implies that debris removal will have increasing marginal benefits when the risk of a collision gets high enough. Figure D.3 shows two examples of this, one with a negative exponential collision rate (globally concave) and another with a sigmoid collision rate (first convex and later concave).

For any positive initial level of debris and satellites (S,D), removal must be nonnegative and no more than all of the debris can be removed. When all satellite owners are identical, the maximum that any one can remove is D/S. This closes the feasible set. Any intermediate amount can also be removed, making the feasible set convex.

The nonconvexity of marginal removal benefits complicates analysis of the optimal amount of removal. There are two cases: the collision rate is globally concave, or the collision rate is convex over some nonnegative interval.

(1) If the collision rate is globally concave, there can be no interior solution to the satellite owner's removal problem. Global concavity implies increasing marginal benefits of debris removal, so satellite owners will choose either to remove all debris or none of it.

(2) If the collision rate is convex over some nonnegative interval, an interior solution is possible but not guaranteed. For example, suppose the collision rate is convex initially and concave near the end, as in the sigmoid case in Figure D.3. Either the right-most intersection of marginal benefits and marginal costs is optimal (where equation 3.56 and inequality 3.57 hold) or else zero removal is optimal.

Determining which corner is optimal when the collision rate is globally concave case is straightforward. If the profits of full removal are greater than the profits of zero removal, full removal is optimal; if not, zero removal is optimal.

With local convexities, the problem is more complicated. One approach is as follows. First, select all solutions to the removal first-order condition (equation 3.56) which satisfy the second-order condition (inequality 3.57), and include them in a set with zero removal and full removal. This is the set of candidate

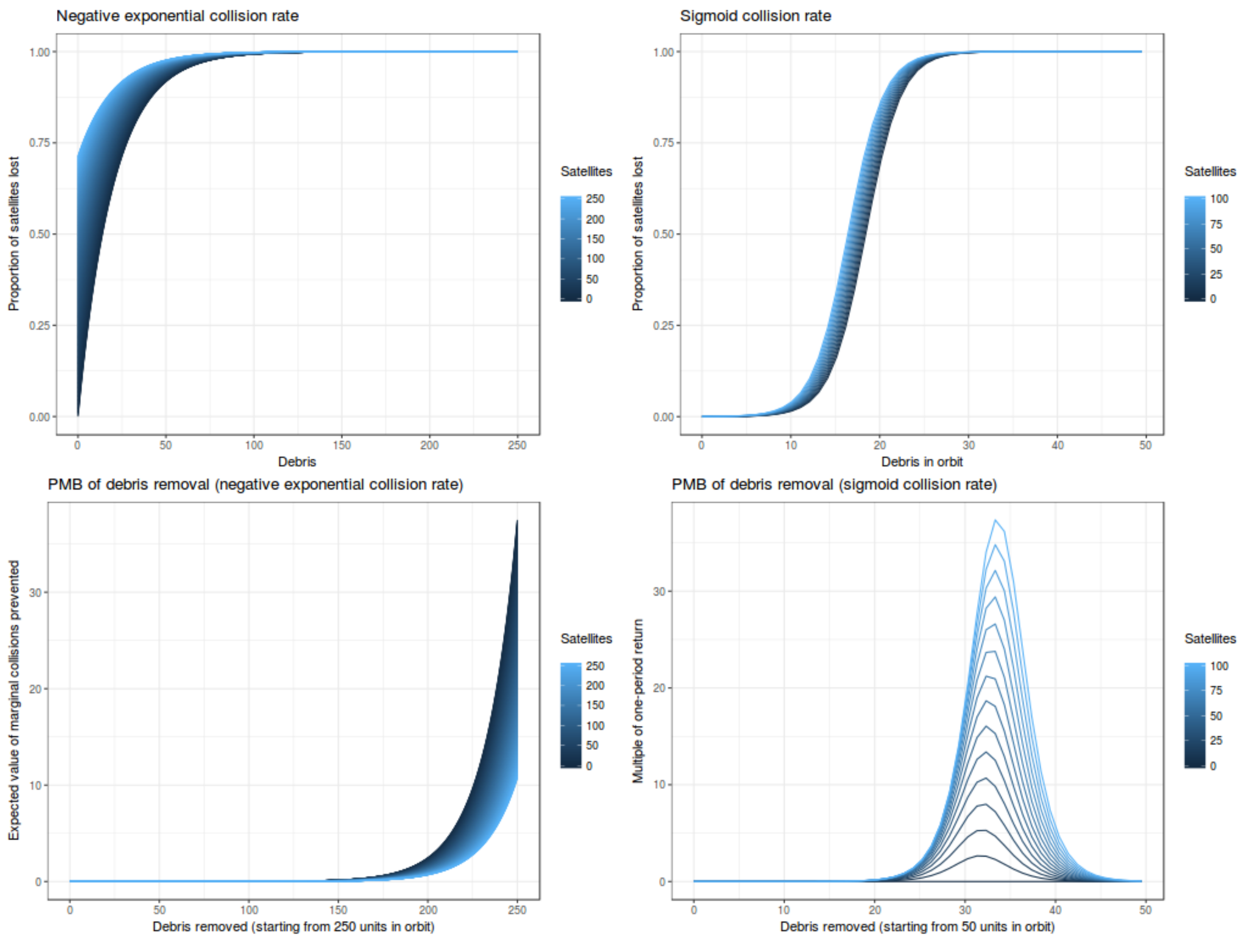

Figure D.3: *Two collision rate functions and the private marginal benefit of debris removal.*
Upper row: Collision risks given different levels of debris removal.
Lower row: Private marginal benefits of debris removal.
Left column: Negative exponential collision rate (globally concave).
Right column: Sigmoid collision rate (convex then concave).
Darker colors correspond to fewer satellites. More satellites may reduce or increase the marginal benefits of debris removal, depending on whether satellites and debris are complements or substitutes in collision production.
Not shown: More initial debris in orbit shifts the removal benefit curves to the right. This makes the optimal removal amount increase until a jump to zero removal.

solutions. Calculate the profits of each candidate solution, and select the one with the highest profits. This procedure is computationally tractable over a closed and convex support as long as the collision rate function is reasonably well-behaved. Figure D.4 illustrates how nonconvexity of the collision rate affects profits and the optimal level of removal.

D.9 Comparative statics of cooperative debris removal and open access launching

I show three results about the demand for debris removal in this section.

First, there is a unique cooperatively-optimal post-removal level of debris for any given level of the satellite stock. This is a consequence of the linear cost (to satellite owners) of debris removal and the monotonicity of the expected collision risk in debris. Due to the linearity, cooperative satellite owners will pursue a most-rapid approach path to the optimal post-removal level of debris in every period. Were the cost nonlinear, the most-rapid approach path would no longer be optimal but the optimal level of debris would remain unique due to monotonicity.

Second, if satellites and debris are "strong enough" complements in producing collision risk, increasing the number of satellite owners in orbit will reduce the optimal post-removal level of debris. This spillover effect in debris removal suggests that a "dynamic virtuous cycle" of active debris removal may be possible: removal in one period can spur entry in the next, which in turn spurs more removal in the following period. Although the functional forms I use rule this effect out, those forms are simplified from a statistical mechanics approximation of orbital interactions. A higher-fidelity model may allow this possibility. A static analog of this effect can be seen in Figure 3.9.

Third, the open access launch rate may be *increasing* in the launch cost. Though this result seems counterintuitive, it is a natural consequence of three features of open access orbit use:

(1) open access drives the value of a satellite down to the launch cost;

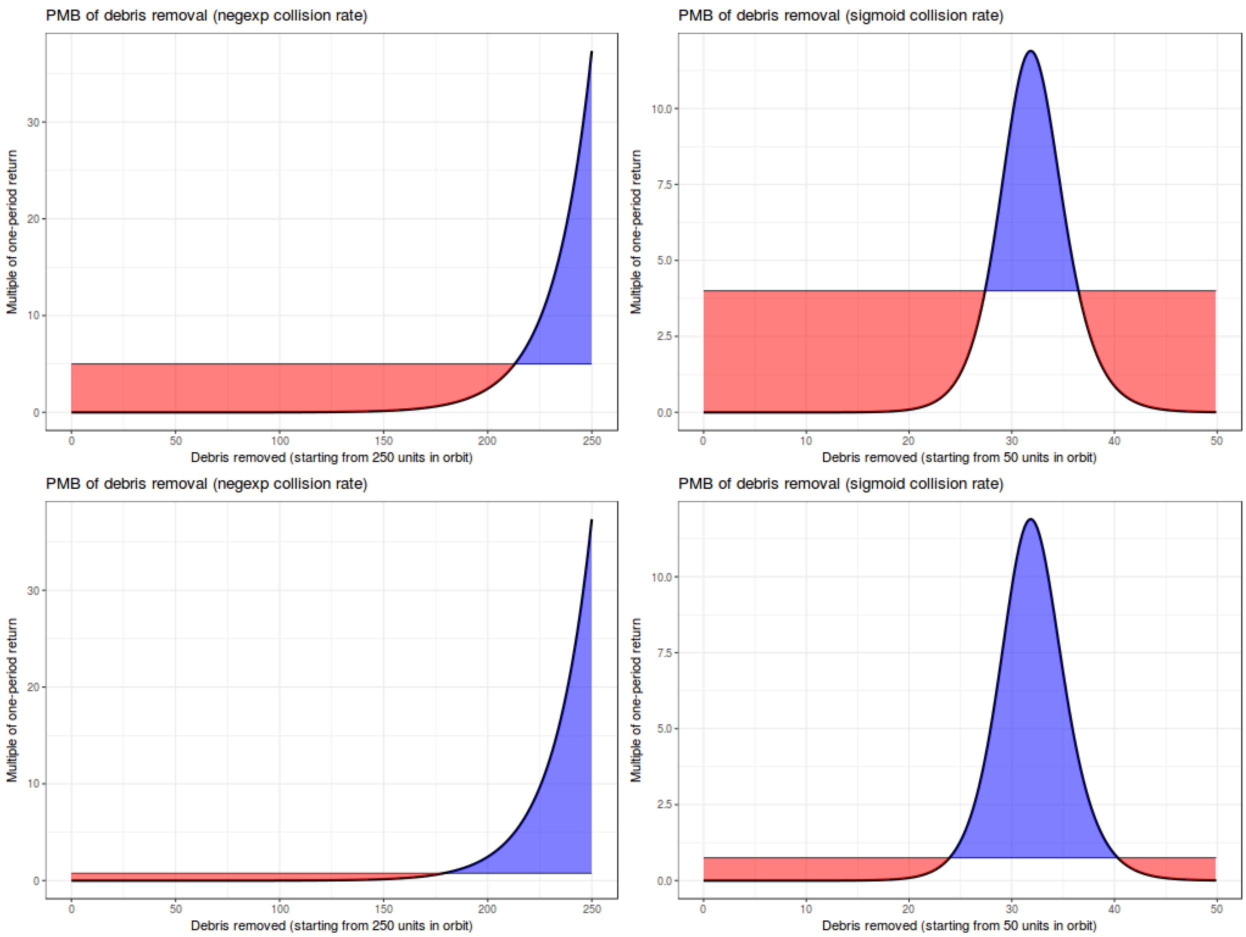

Figure D.4: *Nonconvexity and privately optimal removal.*
Upper row: High cost scenario where zero removal is cooperatively optimal.
Lower row: Low cost scenario where some removal is cooperatively optimal.
Left column: Negative exponential collision rate (globally concave), where the optimal removal demand is always on a corner.
Right column: Sigmoid collision rate (convex then concave), where the optimal removal demand may be in the interior.
The thin horizontal line is the marginal cost of removal. The thicker curve is the marginal benefit of removal. Red regions are losses, blue regions are profits. In the upper row, zero removal is optimal. In the lower left panel, full removal is optimal. In the lower right panel, removal of about 40 units is optimal. Because the collision risk is bounded in $[0,1]$, it cannot be strictly convex globally over S and D.

(2) the amount of removal is increasing in the launch cost;

(3) new entry can reduce the individual expenditure required from cooperative firms to achieve the optimal post-removal level of debris.

The cooperative cost-savings from new entry exceeding the effect of new entry on collision risk is necessary and sufficient for the open access launch rate to be increasing in the launch cost.

Together, these results suggest that the use of debris removal can result in interesting and counterintuitive dynamics in orbit use. Though these results are relevant to understanding the effects of debris removal technologies on orbit use, I omit their proofs from this section. Interested readers may find the proofs in the Appendix, section B.

Cooperative private debris removal:

Lemma 8. *(Law of cooperative private debris removal demand) The cooperative private debris removal demand is*

(1) weakly decreasing in the price of removing a unit of debris, and

(2) weakly increasing in the cost of launching a satellite.

Proof. I consider corner solutions first, then interior solutions. I characterize how interior solutions change in response to a change in the removal price, then show that increases in the price can only induce the firm to reduce their removal demands even at corners. I refer to the non-optimized value of a satellite as $Q_i(R_i)$.

The full removal corner: The first part of the proposition is trivially true at the full removal corner, since the amount of debris removal purchased cannot increase at this corner. So it must either stay the same, or decrease, in response to an increase in the price of removal. For the second part, suppose a firm initially finds full removal optimal. Reducing the amount of debris by a positive amount ε in response to a change in launch cost removed is optimal if and only if, at the new launch cost,

$$
\begin{aligned}
Q_i(D/S-\varepsilon)-Q_i(D/S) > 0 \;\; \forall \varepsilon \in \left(0, \frac{D}{S}\right] \\
\implies \pi + F - c\frac{D}{S} + c\varepsilon - \tilde{E}[\ell|S, D - S(D/S-\varepsilon)]F - \pi - F + c\frac{D}{S} + \tilde{E}[\ell|S,0]F > 0 \\
\implies c\varepsilon - (\tilde{E}[\ell|S,\varepsilon] - \tilde{E}[\ell|S,0])F > 0 \\
\implies \frac{\tilde{E}[\ell|S,\varepsilon] - \tilde{E}[\ell|S,0]}{\varepsilon} < \frac{c}{F} \;\; \forall \varepsilon \in \left(0, \frac{D}{S}\right].
\end{aligned}
$$

If full removal was optimal to begin with, then an increase in the launch cost cannot make it optimal to switch strategies. The above inequality also shows how an increase in the cost of removal can induce a firm to reduce the amount of removal purchased.

The zero removal corner: Consider the profits from increasing the amount of removal from zero to ε in response to a change in the launch cost or removal price. The change is privately optimal if and only if, at the new cost or price,

$$
\begin{aligned}
Q_i(\varepsilon) - Q_i(0) > 0 \;\; \forall \varepsilon \in \left(0, \frac{D}{S}\right] \\
\implies \pi + F - c\varepsilon - \tilde{E}[\ell|S, D - S\varepsilon]F - \pi - F + \tilde{E}[\ell|S,D]F > 0 \\
\implies -c\varepsilon - [\tilde{E}[\ell|S, D - S\varepsilon] - \tilde{E}[\ell|S,D]]F > 0 \\
\frac{\tilde{E}[\ell|S, D - S\varepsilon] - \tilde{E}[\ell|S,D]}{\varepsilon} > \frac{c}{F} \;\; \forall \varepsilon \in \left(0, \frac{D}{S}\right].
\end{aligned}
$$

If zero removal was optimal to begin with, then an increase in the price of removal cannot make it optimal to switch strategies. An increase in the cost of launching a satellite, however, may induce a firm to begin removing debris.

For interior solutions: From equation 3.56,

$$
R_i t : \mathscr{H} = c - \frac{\partial \tilde{E}[\ell]}{\partial D} SF = 0. \tag{D.39}
$$

Applying the Implicit Function Theorem to $\mathscr{H}$,

$$\begin{aligned}\frac{\partial R_i}{\partial c} &= -\frac{\partial \mathscr{H}/\partial c}{\partial \mathscr{H}/\partial R_i} \\ &= -\frac{1}{\frac{\partial^2 \tilde{E}[\ell]}{\partial D^2} S^2 F} < 0.\end{aligned}$$

Strict negativity follows from the second order condition (inequality 3.57). If there are multiple solutions and the removal price increase causes firms to jump from interior one solution to another, they must jump to a solution with less removal.

Similarly, from applying the Implicit Function Theorem to $\mathscr{H}$,

$$\begin{aligned}\frac{\partial R_i}{\partial F} &= -\frac{\partial \mathscr{H}/\partial D}{\partial \mathscr{H}/\partial R_i} \\ &= \frac{\frac{\partial \tilde{E}[\ell]}{\partial D} S}{\frac{\partial^2 \tilde{E}[\ell]}{\partial D^2} S^2 F} > 0.\end{aligned}$$

Strict positivity follows from the second order condition (inequality 3.57). If there are multiple solutions and the launch cost increase causes firms to jump from interior one solution to another, they must jump to a solution with more removal. □

The intuition for this result is simple. Satellite owners pay for debris removal. When the price of removal rises, the demand for removal falls. Under open access the continuation value of a satellite is the cost of launching. So, the demand for debris removal increases when satellites become more valuable. Figure D.5 illustrates Lemma 8.

What about changes in the satellite and debris stocks? Increases in the debris stock increase the cost of achieving any given level of reductions, but may also increase the marginal benefit of removal if the collision rate is locally concave. Increases in the satellite stock may increase the marginal benefit of removal if the collision rate is locally jointly concave, but also increase the number of firms in the market for removal and give existing firms an incentive to reduce their expenditures. I examine this question in Propositions 22 and 23.

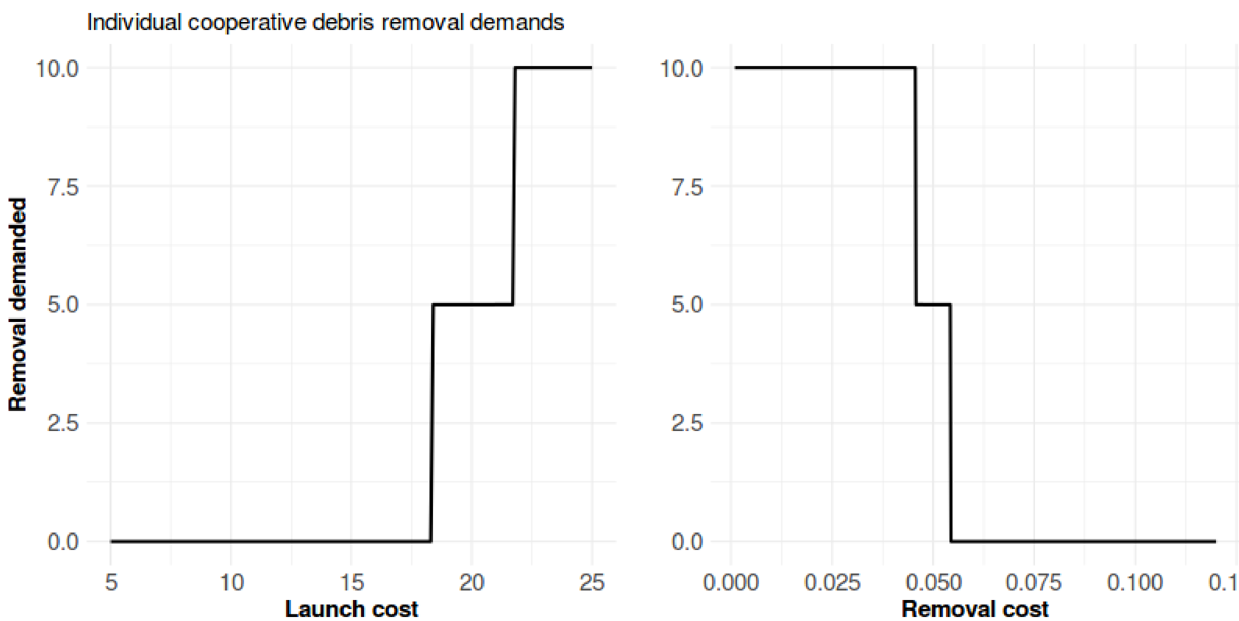

Figure D.5: *The effects of changes in satellite launch and debris removal costs on individual cooperative debris removal demands.*
Increases in the cost of launching a satellite increase the open-access value of satellites in orbit, increasing the amount of debris removal demanded. As expected, increases in the cost of debris removal decrease the amount demanded. Costs are stated in multiples of the one-period return generated by a satellite in orbit.

Proposition 22. *(Cooperative demand for debris removal and the state of the orbit) The optimal post-removal debris level is independent of the pre-removal debris level, but depends on the number of satellites in orbit.*

Proof. Proof. I focus on interior solutions, but the result follows for corner solutions as well due to the monotonicity of $\tilde{E}_t[\ell_t]$ in S_t and D_t.

From equation 3.56,

$$R_{it} : \mathscr{H} = c_t - \frac{\partial \tilde{E}_t[\ell_t|S_t, D_t - R_t]}{\partial D_t} S_t F = 0.$$

Applying the Implicit Function Theorem to $\mathscr{H}$,

$$\begin{aligned} \frac{\partial R_{it}}{\partial D_t} &= -\frac{\partial \mathscr{H} / \partial D_t}{\partial \mathscr{H} / \partial R_{it}} \\ &= \frac{1}{S_t} > 0. \end{aligned}$$

The total quantity of debris removed is

$$R_t = \int_0^{S_t} R_{it} di = S_t R_{it}.$$

Suppose that a positive amount of removal is optimal before and after the change in debris. Differentiating R_t with respect to D_t and using the earlier results for individual removal demands,

$$\frac{\partial R_t}{\partial D_t} = S_t \frac{\partial R_{it}}{\partial D_t} = 1.$$

The monotonicity of $\tilde{E}_t[\ell_t]$ in S_t and D_t also implies that, for any S_t, there is a unique $D_t - R_t$ such that

$$c_t = \frac{\partial \tilde{E}_t[\ell_t|S_t, D_t - R_t]}{\partial D_t} S_t F.$$

□

□

Proposition 22 shows that when it is optimal to remove debris, firms will increase their removal efforts in response to increases in debris. The uniqueness of the optimal post-removal debris stock requires aggregate removal demanded to match changes in the debris stock. This is analogous to the uniqueness of optimal escapement policies in fisheries management. Figure D.6 illustrates this behavior.

Proposition 23. *Additional satellite owners decrease the cooperative individual debris removal demand unless satellites and debris are "strong enough" complements in collision risk production.*

Proof. I show the result for interior solutions. A similar condition also holds for corner solutions.

The total quantity of debris removed is

$$R_t = \int_0^{S_t} R_{it} di = S_t R_{it}.$$

Differentiating R_t with respect to S_t,

$$\frac{\partial R_t}{\partial S_t} = R_{it} + S_t \frac{\partial R_{it}}{\partial S_t}.$$

R_{it} and S_t are both nonnegative by definition. It follows that

$$\frac{\partial R_t}{\partial S_t} > 0 \iff \frac{R_{it}}{S_t} > -\frac{\partial R_{it}}{\partial S_t}.$$

This is always true when individual removal demands increase in response to additional satellite owners ($\frac{\partial R_{it}}{\partial S_t} > 0$). The following steps establish the complementarity condition for interior solutions.

From equation 3.56,

$$R_{it} : \mathscr{H} = c_t - \frac{\partial \tilde{E}_t[\ell_t | S_t, D_t - R_t]}{\partial D_t} S_t F = 0.$$

Applying the Implicit Function Theorem to $\mathscr{H}$,

$$\begin{aligned} \frac{\partial R_{it}}{\partial S_t} &= -\frac{\partial \mathscr{H} / \partial S_t}{\partial \mathscr{H} / \partial R_{it}} \\ &= \frac{\frac{\partial \tilde{E}_t[\ell_t]}{\partial D_t}}{\frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t^2} S_t^2} + \frac{\frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t \partial S_t} - \frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t^2} R_{it}}{\frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t^2} S_t} \lesseqgtr 0. \end{aligned}$$

So, increases in the amount of debris must increase the privately optimal amount of removal at all interior solutions, while increases in the number of satellites will have ambiguous effects. The privately

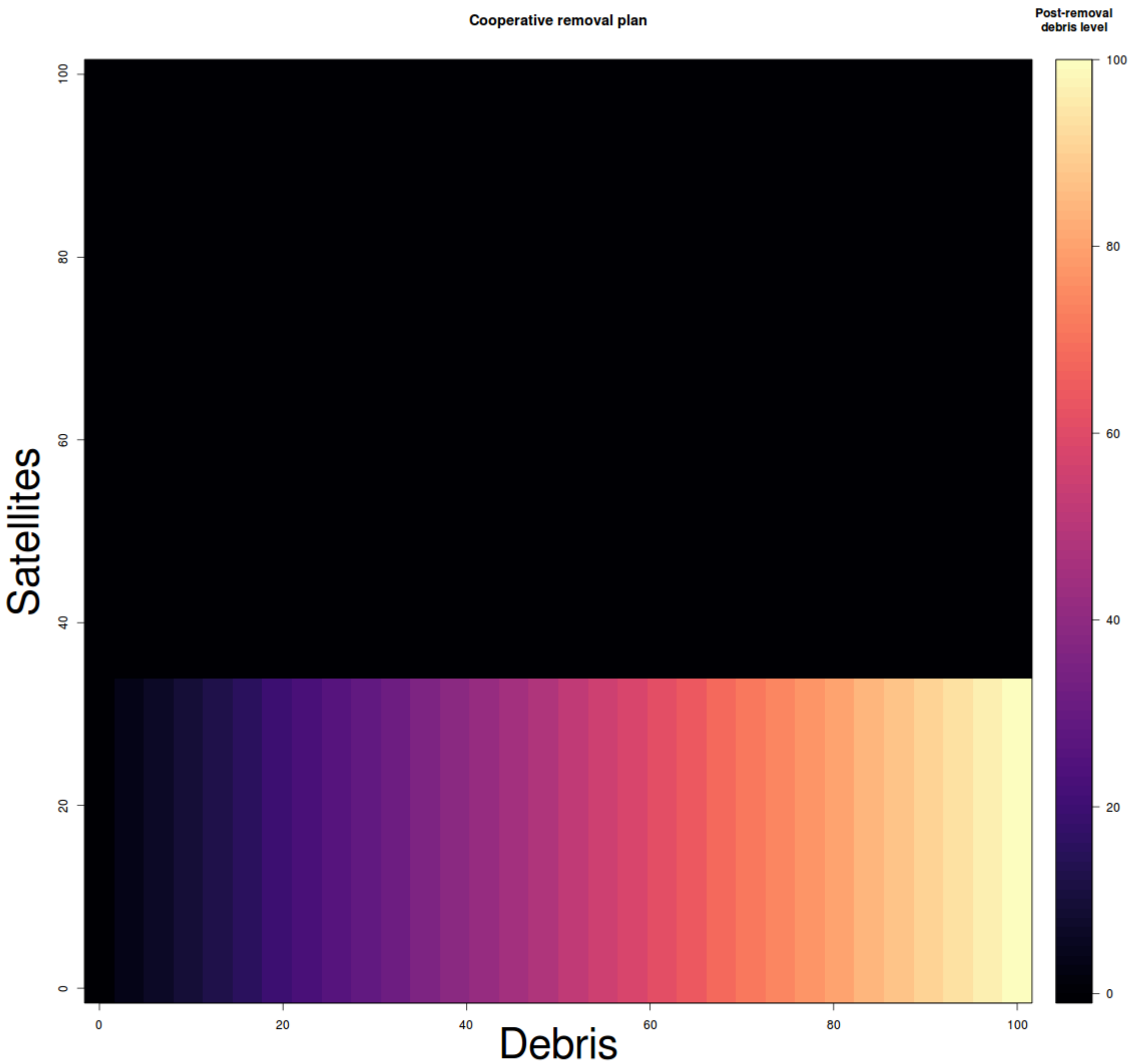

Figure D.6: *The effects of changes in the number of firms and debris in orbit on the post-removal level of debris.*
The color scale represents the amount of debris left in orbit after removal. The cooperatively optimal post-removal level of debris does not depend on the amount of debris initially in orbit, but on the number of firms who are available to share the cost of removal. Once there are enough firms to begin removal the post-removal debris level is constant (full removal).

optimal demand for removal will be increasing in the number of satellites if and only if

$$\frac{\partial R_{it}}{\partial S_t} > 0 \iff \frac{\frac{\partial \tilde{E}_t[\ell_t]}{\partial D_t}}{\frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t^2} S^2} + \frac{\frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t \partial S_t} - \frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t^2} R_i}{\frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t^2} S} > 0$$

$$\iff \frac{\frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t \partial S_t}}{\frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t^2} S} > -\frac{\frac{\partial \tilde{E}_t[\ell_t]}{\partial D_t}}{\frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t^2} S^2} + \frac{R_i}{S}$$

$$\iff \frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t \partial S_t} < -\frac{\frac{\partial \tilde{E}_t[\ell_t]}{\partial D_t}}{S} + \frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t^2} R_i.$$

The right hand side of the final line is strictly negative at an interior optimum from the second-order condition, inequality 3.57. So $\frac{\partial^2 \tilde{E}_t[\ell_t]}{\partial D_t \partial S_t}$ must be "sufficiently" negative for the presence of new satellites to increase privately optimal removal. Economically, this means that satellites and debris are "strong enough" substitutes.

□

Proposition 23 shows that firms may increase or decrease their removal demands in response to more firms entering the orbit. At the full removal corner, they will always reduce their demands. This is a cooperative cost-sharing effect: it remains optimal to remove all debris, but the contribution required of each firm decreases when more firms enter. At an interior solution, their response to more satellites depends on two effects: a congestion effect and a cooperation effect. Their net effect depends on the collision rate's convexity in debris and satellites, particularly whether satellites and debris are "strong enough" complements in producing collision risk. If

- the cooperation and congestion effects are collectively positive, then the presence of more satellites increases the marginal benefit of debris removal (positive spillover effects);
- the cooperation and congestion effects are collectively negative, then the presence of more satellites decreases the marginal benefit of debris removal (negative spillover effects).

These shifts are shown algebraically in the Appendix, section D.4. Conditional on a positive amount of removal being optimal, increases in debris are matched by increases in removal. Increases in the number of satellites can result in aggregate removal decreases if the individual demand reduction is large enough.

Given the functional forms I assume for simulations, the shifts are always collectively negative. I remain agnostic about the “correct” functional form to assume.

Any increase in debris is matched by a commensurate increase in aggregate removal. Since satellites are identical, owners collectively agree on the optimal level of debris. The total quantity of debris removed can not be decreasing in the number of satellites unless individual owners reduce their removal demands in response to more satellites in orbit. Proposition 23 shows this is a necessary but not sufficient condition.

Proposition 22 also offers some insight into the effects of launches on privately optimal removal. With no launch debris, the effect of a marginal launch on privately optimal removal is only the effect of a new satellite on privately optimal removal. With launch debris, the effect of a launch is a combination of the effect of a new satellite and the effect of new debris.

Open access launching: The option - or, in the cooperative case studied here, the obligation - to remove space debris alters the incentives of open access satellite launchers. While they will still launch until expected profits are zero, the expected collision risk is no longer the only object which equilibriates their launching behavior. In addition to expected collision risk, the expenditure they expect to incur removing debris as satellite owners will also adjust to equilibriate the launch rate. Though they will be price takers when their satellites reach orbit, they can anticipate the number of satellite owners who will contribute to debris removal. This acts in the opposite direction as the expected collision risk: while more satellites in orbit increases risk, more firms with satellites in orbit decreases each individual firm’s debris removal expense. Since debris removal also reduces collision risk, the net effect of introducing debris removal financed by satellite owners may be more launches than would otherwise occur. Indeed, this is precisely what occurs in the cases simulated here.

In addition to this perhaps-counterintuitive effect, it is plausible that an increase in the cost of launching a satellite could increase the launch rate. This is not as pathological a case as it may seem at first. Since open access drives the value of a satellite down to the launch cost, and the cooperatively-optimal amount of debris removal which satellite owners will pay for is increasing in the launch cost, and increase in the launch cost under open access could increase the value of owning a satellite by more than it increases the cost of launching it, at least locally near an existing equilibrium. This is not a violation of the law of demand for satellite ownership; rather, it is a violation of the "all else equal" clause. Assumption 7 describes a necessary and sufficient condition to rule this case out.

Assumption 7. *(New launches reduce the expected profits of satellite ownership) The change in individual removal expenses from a marginal satellite launch is smaller in magnitude than the sum of the change in expected future collision costs from a marginal satellite launch and the change in individual removal expenses from a marginal piece of launch debris. Formally,*

$$\left|\frac{\partial E_t[\ell_{t+1}]}{\partial S_{t+1}}F + m\left(\frac{\partial E_t[\ell_{t+1}]}{\partial D_{t+1}}F + \frac{\partial R_{it+1}}{\partial D_{t+1}}c_{t+1}\right)\right| > \left|\frac{\partial R_{it+1}}{\partial S_{t+1}}c_{t+1}\right|.$$

If this assumption is violated, then launches increase the profitability of owning a satellite through the debris removal expenditure channel described above. It is also possible that increases in the cost to satellite owners of removing a unit of debris could increase the launch rate Assumption 8 describes an additional condition necessary for increases in the price of debris removal to reduce the launch rate. An increase in the price of debris removal will reduce the cooperatively-optimal amount of debris removal satellite owners purchase, potentially reducing the total debris removal expenditure and increasing the profits of owning a satellite. As in the case of launch rates being increasing in launch costs, this is not a violation of the law of demand for satellite ownership; it is a violation of the "all else equal" clause.

Assumption 8. *(Removal expenditure is increasing in the removal cost) The cooperative private debris removal expenditure is increasing in the price of removing a unit of debris. Formally,*

$$\frac{\partial}{\partial c_{t+1}}(R_{it+1}c_{t+1}) = R_{it+1} + \frac{\partial R_{it+1}}{\partial c_{t+1}}c_{t+1} > 0.$$

Assumption 8 states that the amount of debris removed (R_{it+1}) is larger than the reduction in removal due to a price increase ($\frac{\partial R_{it+1}}{\partial D_{t+1}}c_{t+1}$, which is weakly negative from Proposition 8). This is likely to hold whenever the change in individual removal demands from a change in removal cost is small, for example,if removal demand is in the interior before and after the change. It is unlikely to hold if the opposite is true, for example,if the change in removal cost causes individual removal demands to jump from the full removal corner to the zero removal corner at a time when there are few satellites and many debris fragments. Though future cooperative private debris removal demands are an anticipated cost to current satellite launchers, those same launchers may find their willingness to launch increasing in the cost of removal if it reduces the burden of cooperating and purchasing removal.

Proposition 24. *(Private demand for satellite ownership) The open access launch rate is*

(1) strictly decreasing in the cost of launching a satellite if and only if new launches reduce the expected profits of satellite ownership; and

(2) strictly decreasing in the price of removing a unit of debris in $t+1$ *only if new launches reduce the expected profits of satellite ownership AND the cooperative private expenditure on debris removal is increasing in the cost of removal.*

Proof. From equation 3.54,

$$X_t : \mathscr{F} = \pi - rF - E_t[\ell_{t+1}]F - R_{it+1}c_{t+1} = 0.$$

Applying the Implicit Function Theorem to $\mathscr{F}$ and assuming R_{it+1} is chosen optimally (as described in Proposition 8),

$$\begin{aligned}\frac{\partial X_t}{\partial F} &= -\frac{\partial \mathscr{F}/\partial F}{\partial \mathscr{F} \partial X_t} \\ &= -\frac{r + E_t[\ell_{t+1}] + \frac{\partial R_{it+1}}{\partial F}c_{t+1}}{\frac{\partial E_t[\ell_{t+1}]}{\partial S_{t+1}}F + m\left(\frac{\partial E_t[\ell_{t+1}]}{\partial D_{t+1}}F + \frac{\partial R_{it+1}}{\partial D_{t+1}}c_{t+1}\right) + \frac{\partial R_{it+1}}{\partial S_{t+1}}c_{t+1}},\end{aligned}$$

which is negative for all parameter values when Assumption 7 holds.

Similar manipulations yield

$$\begin{aligned}\frac{\partial X_t}{\partial c_{t+1}} &= -\frac{\partial \mathscr{F}/\partial c_{t+1}}{\partial \mathscr{F} \partial X_t} \\ &= -\frac{R_{it+1} + \frac{\partial R_{it+1}}{\partial c_{t+1}}c_{t+1}}{\frac{\partial E_t[\ell_{t+1}]}{\partial S_{t+1}}F + m\left(\frac{\partial E_t[\ell_{t+1}]}{\partial D_{t+1}}F + \frac{\partial R_{it+1}}{\partial D_{t+1}}c_{t+1}\right) + \frac{\partial R_{it+1}}{\partial S_{t+1}}c_{t+1}}.\end{aligned}$$

$\frac{\partial X_t}{\partial c_{t+1}}$ is negative only if Assumptions 7 and 8 hold. $\square$

Given Assumption 7, Assumption 8 is necessary and sufficient for $\frac{\partial X_t}{\partial c_{t+1}}$ to be negative. A simultaneous violation of Assumptions 7 and 8 would indicate that the private marginal cost of orbit use was decreasing in the costs of access and debris removal - a counterintuitive situation, but not an *a priori* impossible one.

Appendix E

Legal and institutional features of orbit use policy

E.1 International laws regarding space traffic control

Orbits are inherently global resources, and space law is fragmented across nations and documents. Space law spans domestic policies, international treaties, bilateral agreements, and guidelines. Not all agreements are signed by all spacefaring nations, and many are non-binding. Most of the agreements are vague and suffer from enforcement problems. Four of the most relevant international agreements relating to orbit management are the 1967 Outer Space Treaty, the 1972 Liability Convention, the 1975 Registration Convention, and the 2007 COPUOS Guidelines.[1]

1967 Outer Space Treaty The Outer Space Treaty[2] established the legal framework for peaceful uses of outer space. Article 2 of the Treaty designates outer and orbital space as common pool resources, to be used "for the benefit of all" humankind. The only explicit restrictions are on military uses and claims of national sovereignty; the state of resource use is left ambiguous. The Treaty does not mention debris, only stating that nations should avoid causing (undefined) "harmful contamination" of outer space.

1972 Liability Convention The Liability Convention[3] established the framework for tort law of space activities. However, the Convention focused more on damage to terrestrial objects from re-entry than on damages to orbital objects which occur in space. "Damage" in this Convention is defined only in relation to realized outcomes for people and property, rather than potential outcomes caused by the

[1] A more detailed analysis of these laws can be found in Akers (2012).

[2] The "Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies."

[3] The "Convention on International Liability for Damage Caused by Space Objects."

environment. Additionally, the Convention places liability for such damages on the launching state rather than the launching entity. This has motivated nations like the US to require satellite owners insure their satellites, with the federal government indemnifying losses beyond a certain amount. The EU has different insurance requirements, with a similar motivation. There is no liability attached to producing debris in orbit, only to attributable damages. Liability extends to damage to people or property caused by re-entry. Such attribution is difficult in space, where damages may be caused by difficult-to-detect fragments of unknown origin.

1975 Registration Convention The Registration Convention[4] requires nations to register space objects launched from or by that nation with the UN Secretary-General. The responsibility for ensuring compliance lies with the launching state, with the UN being responsible for integrating all the registrations and publishing a publicly available international registry of objects in orbit. The Convention only requires basic orbital information to be provided: orbital parameters to ascertain the object's initial path, and the general function. It does not require more detailed information, such as orbit changes or satellite positions, or even continuous updates. The Convention does not offer a deadline by which a launched object must be registered, or specify a penalty or enforcement mechanism for noncompliance.

2007 COPUOS Guidelines The COPUOS Guidelines[5] are seven nonbinding guidelines for mitigating artificial space debris. This was the first international treaty to recognize the problem of orbital debris, though it is unenforceable and only focused on technical mitigation practices rather than economic control measures. Senechal (2007) discusses some features that an enforceable international space debris convention should possess; the COPUOS guidelines are a step in this direction, but they do not contain the kinds of clear definitions and enforceable provisions required.

In general, the legal doctrine of *nemo dat quod non habet* ("no one gives what he doesn't have") means that the Outer Space Treaty prevents states from issuing rights over orbital paths. Since they lack sovereignty in space, states do not have rights to give[6]. States retain their authority over launches within

---

[4]The "Convention on Registration of Objects Launched into Outer Space."

[5]The "Committee on the Peaceful Uses of Outer Space (COPUOS) 2007 Nonbinding Guidelines for Space Debris Mitigation."

[6]Salter and Leeson (2014) argue that this poses no difficulty to efficient decentralized orbit use management, as privately enforced

their borders, and satellites operated by firms within their borders.

E.2 US institutions regarding space traffic control

International law places responsibility for objects launched to space on the nations from which the objects are launched and the nations in which the launching entities are registered. This means that understanding the legal status of space objects also requires some background on national laws. A majority of currently-operational satellites were launched from the United States, Russia, and China. In this section, I focus on institutions in the United States.

In the US, the Federal Aviation Administration's Office of Commercial Space Launch handles issues related to launches and reentry, including issuing launch permits. The Department of Commerce currently regulates remote sensing satellite systems, with a focus on controlling the resolution and coverage of images that are sold. The DOC is set to take control of regulating on-orbit activities by US entities over the next two years. The FCC regulates satellite activities by US telecom entities through radio spectrum controls. Traffic management operations so far have therefore been limited to controlling launches or operations by entities primarily based within a country's national borders - for example, the FCC can use its control of spectrum rights to deny service to poorly-behaving telecom operators who want to provide service to North America, but have no leverage over providers who are interested in serving China. The patchwork of laws around the world has already led some firms to evade regulation in their home area by launching from another (for example, Dvorsky (2018)). Space Policy Directive-3, a presidential memorandum issued in June of 2018, directs federal agencies to cooperate in developing a framework for a national space traffic management policy centered on a new object tracking infrastructure to better predict the aggregate collision rate and coordinate collision avoidance maneuvers.

property rights have arisen in other common resource settings on Earth. Weeden and Chow (2012) offer suggestions guided by Ostrom's principles of commons management for developing decentralized protocols for orbit use. An interesting question, not pursued in detail here, is the degree to which physical dynamics allow cooperative mechanisms to operate. Section D.7 in the Appendix examines the stability of cooperation with debris removal plans.

E.3 The militarization of space

Military use of space accounts for over 10% of known active satellites in orbit Union of Concerned Scientists (2018). Although I focus on commercial orbit use with no reference to military use, some understanding of military incentives is necessary to place military use in the appropriate context. Sweeney (1993) describes some of generic logic for the effects of national security concerns on nationally optimal depletable resource extraction, whereby the socially optimal extraction rate may exceed the privately optimal extraction rate. Similarly, national security concerns may drive a national fleet planner to launch more than their national space sector would.[7]

Anti-satellite missile uses can trigger Kessler Syndrome. The space traffic control policies described here are not designed to control anti-satellite missile use. International agreements around space traffic control in the future are likely to be shaped in large part by the military stances of major space-faring nations, most notably the United States, Russia, and China. The current lack of binding international space traffic agreements is partly due to a lack of science and consensus around how such controls should be designed and why, but also partly due to military-related incentives facing the governments of space-faring nations which would be party to such agreements.

Article 4 of the Outer Space Treaty declares that state parties "undertake not to place in orbit around the Earth any objects carrying nuclear weapons or any other kinds of weapons of mass destruction, install such weapons on celestial bodies, or station such weapons in outer space in any manner." It also forbids establishing military installations, conducting weapons tests, or any other non-peaceful activities on the Moon and other celestial bodies. Despite these provisions, the Outer Space Treaty does not explicitly prohibit using near-Earth space for reconnaissance, terrestrial warfare coordination, or even

[7] The depletable resource case may be more apt than the renewable resource one for some orbital regimes, particularly higher-altitude ones such as GEO or high-Earth orbits. Natural rates of orbital decay in these regimes can be on the order of millennia - long enough that Kessler Syndrome without removal technologies can render them effectively unusable (the orbital volume "completely extracted") for economic purposes.

outright conflict so long as "weapons of mass destruction" are not used in orbit. The ongoing militarization of space has therefore involved these uses, with the US government being the largest such user of orbital space. The US government has not yet supported international treaty efforts to limit the militarization of space. Shimabukuro (2014) offers an explanation for the lack of more international regulation on space militarization in light of rising tensions between the US and China. The core of that explanation is that the US benefits from high ambiguity over acceptable uses of outer space given the US's high level of military dependence on space systems. This allows the US to induce its adversaries to spend ever-increasing amounts on developing comparable space capabilities, potentially collapsing their warfighting capabilities without a battle.

Appendix F

Full proofs for Chapter 4

F.1 Derivation of the optimal launch rate

In this section we derive the equations characterizing the planner's launch rule, equations 4.9 and 4.10. Period t values are shown with no subscript, and period $t+1$ values are marked with a $'$, e.g. $S_t \equiv S, S_{t+1} \equiv S'$. For notational simplicity, we keep the economic parameters are constant over time. The fleet planner's problem is

$$W(S,D) = \max_{X \geq 0} \{\pi S - FX + \beta W(S',D')\} \tag{F.1}$$

$$\text{s.t.} \quad S' = S(1 - L(S,D)) + X \tag{F.2}$$

$$D' = D(1-\delta) + G(S,D) + \gamma A + mX. \tag{F.3}$$

The fleet planner's launch plan will satisfy

$$X^* : \beta[W_S(S',D') + mW_D(S',D')] = F, \tag{F.4}$$

that is, the planner will launch until the marginal value to the fleet of a new satellite plus the marginal value to the fleet of its launch debris is equal to the launch cost.

Assuming an optimal policy function $X^* = H(S,D)$ exists and applying the envelope condition, we have the following expressions for the fleet's marginal value of another satellite and another piece of debris:

$$W_S(S,D) = \pi + \beta[W_S(S',D')(1 - L(S,D) - SL_S(S,D)) + W_D(S',D')G_S(S,D)] \tag{F.5}$$

$$W_D(S,D) = \beta[W_D(S',D')(1 - \delta + G_D(S,D)) + W_S(S',D')(-SL_D(S,D))] \tag{F.6}$$

Rewriting equation F.4, we have

$$W_S(S',D') = \left[\frac{F}{\beta} - mW_D(S',D')\right] \tag{F.7}$$

Plugging equation F.7 into equations F.5 and F.6,

$$W_S(S,D) = \pi + F(1 - L(S,D) - SL_S(S,D)) - \beta W_D(S',D')[m(1 - L(S,D) - SL_S(S,D)) - G_S(S,D)] \tag{F.8}$$

$$W_D(S,D) = (-SL_D(S,D))F + \beta W_D(S',D')[1 - \delta + G_D(S,D) - m(-SL_D(S,D))] \tag{F.9}$$

Define the following quantities:

$$\begin{aligned}
\text{Marginal satellite return: } & \alpha_1(S,D) = \pi + (1 - L(S,D) - SL_S(S,D))F \\
\text{Cost of launch debris' collisions: } & \alpha_2(S,D) = -SL_D(S,D)F \\
\text{Growth-launch fragment balance: } & \Gamma_1(S,D) = G_S(S,D) - m(1 - L(S,D) - SL_S(S,D)) \\
\text{New fragments from current stock: } & \Gamma_2(S,D) = 1 - \delta + G_D(S,D) + mSL_D(S,D).
\end{aligned}$$

These allow us to rewrite equations F.8 and F.9 as

$$W_S(S,D) = \alpha_1(S,D) + \beta\Gamma_1(S,D)W_D(S',D') \tag{F.10}$$

$$W_D(S,D) = \alpha_2(S,D) + \beta\Gamma_2(S,D)W_D(S',D'). \tag{F.11}$$

As long as $\delta < 1$, $\Gamma_2(S,D) \neq 0\ \forall(S,D)$, allowing us to rewrite equation F.11 as

$$W_D(S',D') = \frac{W_D(S,D) - \alpha_2(S,D)}{\beta\Gamma_2(S,D)}. \tag{F.12}$$

Plugging equation F.12 into equation F.10, we get

$$\begin{aligned}
W_S(S,D) &= \alpha_1(S,D) + \beta\Gamma_1(S,D)\frac{W_D(S,D) - \alpha_2(S,D)}{\beta\Gamma_2(S,D)} \\
&= \alpha_1(S,D) + \frac{\Gamma_1(S,D)}{\Gamma_2(S,D)}(W_D(S,D) - \alpha_2(S,D)) \\
\implies W_S(S,D) &= \alpha_1(S,D) - \frac{\Gamma_1(S,D)}{\Gamma_2(S,D)}\alpha_2(S,D) + \frac{\Gamma_1(S,D)}{\Gamma_2(S,D)}W_D(S,D)
\end{aligned} \tag{F.13}$$

Iterating equation F.7 one period backwards and plugging it into equation F.10, we get

$$\begin{aligned}
&\frac{F}{\beta} - mW_D(S,D) = \alpha_1(S,D) - \frac{\Gamma_1(S,D)}{\Gamma_2(S,D)}\alpha_2(S,D) + \frac{\Gamma_1(S,D)}{\Gamma_2(S,D)}W_D(S,D) \\
&\implies W_D(S,D) = \left[\frac{F}{\beta} - \alpha_1(S,D) - \frac{\Gamma_1(S,D)}{\Gamma_2(S,D)}\alpha_2(S,D)\right]\left[\frac{\Gamma_1(S,D)}{\Gamma_2(S,D)} + m\right]^{-1}. \qquad \text{(F.14)}
\end{aligned}$$

Substituting in the forms for $\alpha_1(S,D), \alpha_2(S,D), \Gamma_1(S,D)$, and $\Gamma_2(S,D)$, equation F.11 yields equation 4.9 and equation F.14 yields equation 4.10.

Appendix G

Estimation details

G.1 Measurement error in satellite and debris counts

Limitations of sensor technology suggest that the debris counts are lower-bound estimates. To the extent that this biases the collision probability and debris counts downward, it will bias the estimated decay rate, collision probability parameters, fragmentation parameters, and launch debris weakly downwards. Since downward bias in the physical parameters makes collisions and missile tests appear to cause less congestion than they actually do, the open access and optimal launch rates will be inflated.

Downward bias in the collision probability data will bias the economic parameter estimates weakly downwards as well. This will to some degree offset the inflation in the launch rate caused by the physical parameter underestimation, though the exact extent of the offset is not clear.

In general, measurement error in the collision probability data also causes the nonnegativity constraint on the collision probability parameters (α_{SS} and α_{SD}) to bind in some bootstrap replications. This causes issues, of the type described in (Ketz, 2018), in obtaining asymptotic standard errors.

G.2 Collision probability model misspecification

We assume that the collision probability model has constant parameters. Changes in patterns of satellite placement, construction, and ownership structures lead to changes over time in the physical primitives reflected in α_{SS} and α_{SD}. The "net" convexity or concavity of the time path of the primitives will determine whether the constant approximations over or understate the true time-varying parameters in any period on average. A convex time path — low values initially and high values later on — will be overestimated on average, while a concave time path — high values initially, with slow increases over time — will be underestimated on average.

The misspecification causes two problems with simulation and inference. First, underestimation will inflate launch rate projections and overestimation will deflate them. However, because the deflation affects both open access and optimal launch rates in the same way, the simulated optimal satellite tax will not be affected. Second, underestimation may cause the nonnegativity constraint on the collision probability parameters to bind in some bootstrap replications, causing the same types of asymptotic issues as measurement error.

G.3 Measurement error in returns and costs

We take the returns and costs of satellite ownership from the data used in Wienzierl (2018), which aggregate revenues from all commercial satellites in orbit. By including more than just LEO satellites, the direct returns and costs data overstate the returns to LEO paths. The economic parameter estimates therefore reflect a "LEO share" coefficient on the revenue data between 0 and 1. The LEO share coefficient attenuates the estimates of a_{L1}, a_{L2}, and a_{L3}.

G.4 Returns to scale and economic misspecification

We assume that the LEO satellites aggregate in constellations with constant returns to scale. Decreasing returns to scale will inflate the projected open access and optimal launch rates, while increasing returns to scale will have the opposite effect. While I do not have detailed data on LEO constellation revenues, historical returns and fleet sizes for GEO telecom satellites over 2000—2012 from the TelAstra Communications Satellite Databases (TelAstra, Inc., 2017) offer suggest that both increasing and decreasing returns to scale are plausible.

The data contain returns and constellation sizes for the major commercial entities offering Fixed Satellite Services (using GEO satellites) from 2000 to 2012. Selecting operators with both a constellation size and a returns entry in each recorded period yields the 17 largest constellations in GEO, accounting for approximately 68% of all satellites in GEO in 2009.

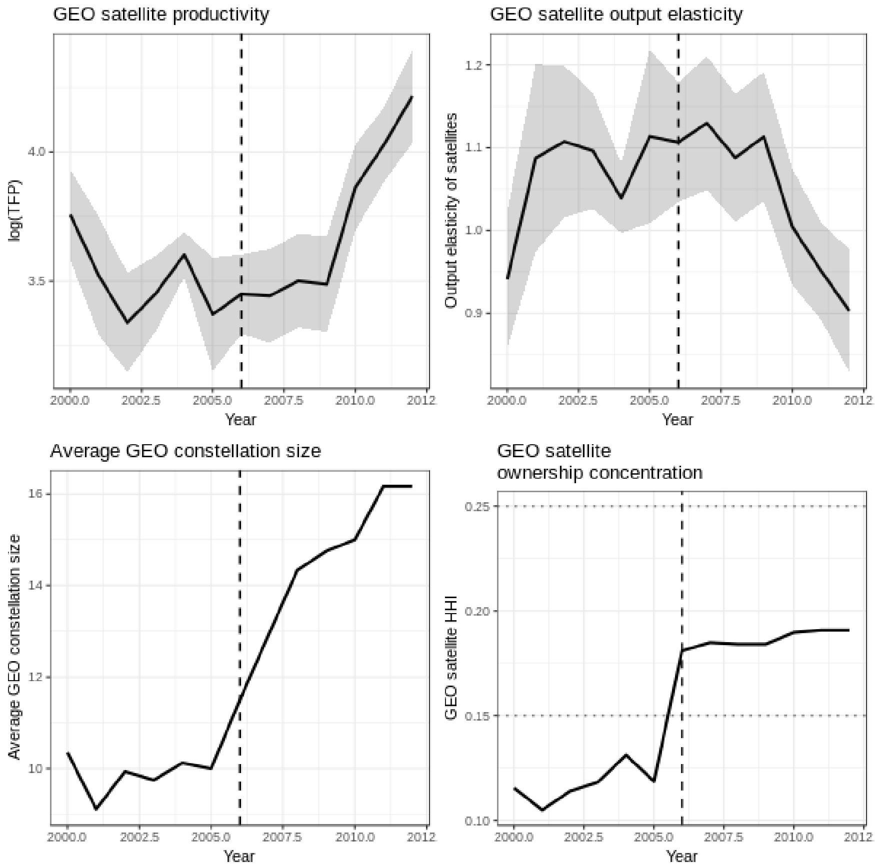

Figure G.1: *Trends in GEO satellite productivity and ownership.*

Upper left: The natural log of satellite Total Factor Productivity.

Upper right: The revenue elasticity of satellites in a constellation, i.e. returns to scale from satellite ownership.

Lower left: The average number of satellites in a GEO constellation over time.

Lower right: The Herfindahl-Hirschman Index for GEO satellite ownership. Higher HHI values indicate more ownership concentration. The vertical black dashed line marks the merger between Intelsat and PanAmSat, wherein Intelsat received all of PanAmSat's satellites.

Standard errors in TFP and output elasticity estimates are shown as shaded regions.

G.5 Projecting the launch constraint

To prevent the model from violating the limited availability of launches, we estimate the launch constraint from the observed historical data and then project it forward. In each historical period, we calculate the maximum number of satellites which can be launched as the cumulative maximum of launch attempts (successes+failures). From the historical calculation, we project the launch constraint forward with a linear time trend and an intercept. Table G.1 shows the estimated coefficients, and figure G.2 shows the estimated and projected launch constraint time paths.

Launch constraint model parameters:	Intercept	Time trend
Parameter values:	30.13	12.5
Standard errors:	16.43	2.65

Table G.1: Parameter values from linear model of launch constraint. All values are rounded to two decimal places. We estimate these coefficients using OLS on the historical launch constraint.

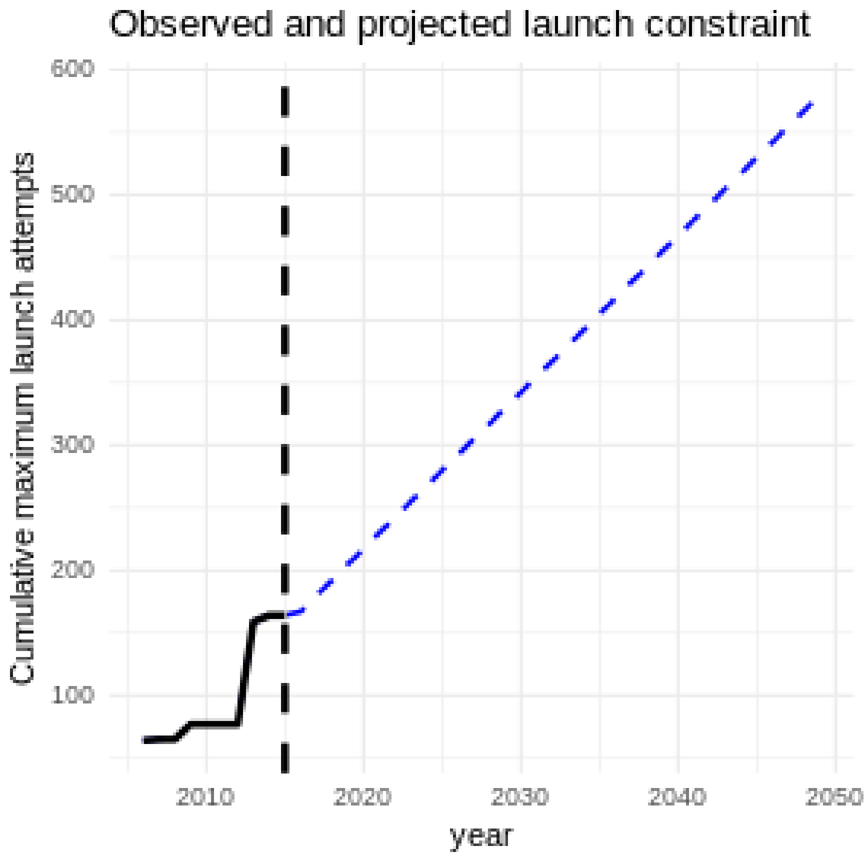

Figure G.2: *Launch constraint, observed and projected.*

The black line shows the observed launch constraint (cumulative max of attempted launches). The blue dashed line shows a linear projection of the launch constraint.

G.6 Sensitivity analyses of physical equation calibration

To study the sensitivity of our conclusions to estimated physical parameters, we conduct a sensitivity analysis of the model simulations given different physical parameter values. We use a residual bootstrap procedure to accomplish this.

First, we estimate equations 4.12 and 4.13 as described above. Then, we sample from the distribution of residuals to generate “bootstrap worlds”. We add these residuals to the estimated models to generate bootstrap world outcome variables. Finally, we re-estimate the model using the bootstrap world outcomes to generate alternate sets of physical parameter estimates, and simulate the model under a random sample of those estimates. The procedure is described in Algorithm 4.

Algorithm 4: Residual bootstrap

1. Estimate equations 4.12 and 4.13, get residuals $\{e_{Lt}, e_{Dt}\}_t$.
2. Sample from distribution of residuals, $\{\hat{e}^i_{Lt}, \hat{e}^i_{Dt}\}_t$.
3. Add sampled residuals to estimated models, generate "bootstrap world" outcomes $\{(S_t L_t)^i, D^i_{t+1}\}_t$
4. Re-estimate equations 4.12 and 4.13 using $\{(S_t L_t)^i, D^i_{t+1}\}_t$

One issue to note is that, because we estimate equation 4.12 with a constrained procedure and the coefficients are near one of the constraint boundaries, the asymptotic properties of this procedure are difficult to obtain. Since our goal is not asymptotic analysis of standard errors but rather to generate alternate parameter sets in a principled way for counterfactual simulations, we select the main model estimates as the mean of the bootstrap world parameters. This ensures that our sensitivity analysis selects parameters around the main model estimates. Ultimately this is inconsequential for the outcomes of interest — collision risk under open access and optimal management, and the resulting satellite tax — since the outcomes are endogenous variables which satisfy economic conditions irrespective of the specific physical parameter values. The physical parameter values affect the specific paths of launches, satellites, and debris, but only such that the collision risk continues to satisfy the economic conditions. Figures G.3 and G.5 show bootstrapped sensitivity analyses of equations 4.12 and 4.13; Figures G.7 and G.8 show bootstrapped projected paths of orbital aggregates under open access; and Figure G.9 shows the resulting bootstrapped optimal satellite tax projections. Figures G.4 and G.6 show the distributions of bootstrapped parameter estimates for equations 4.12 and 4.13.

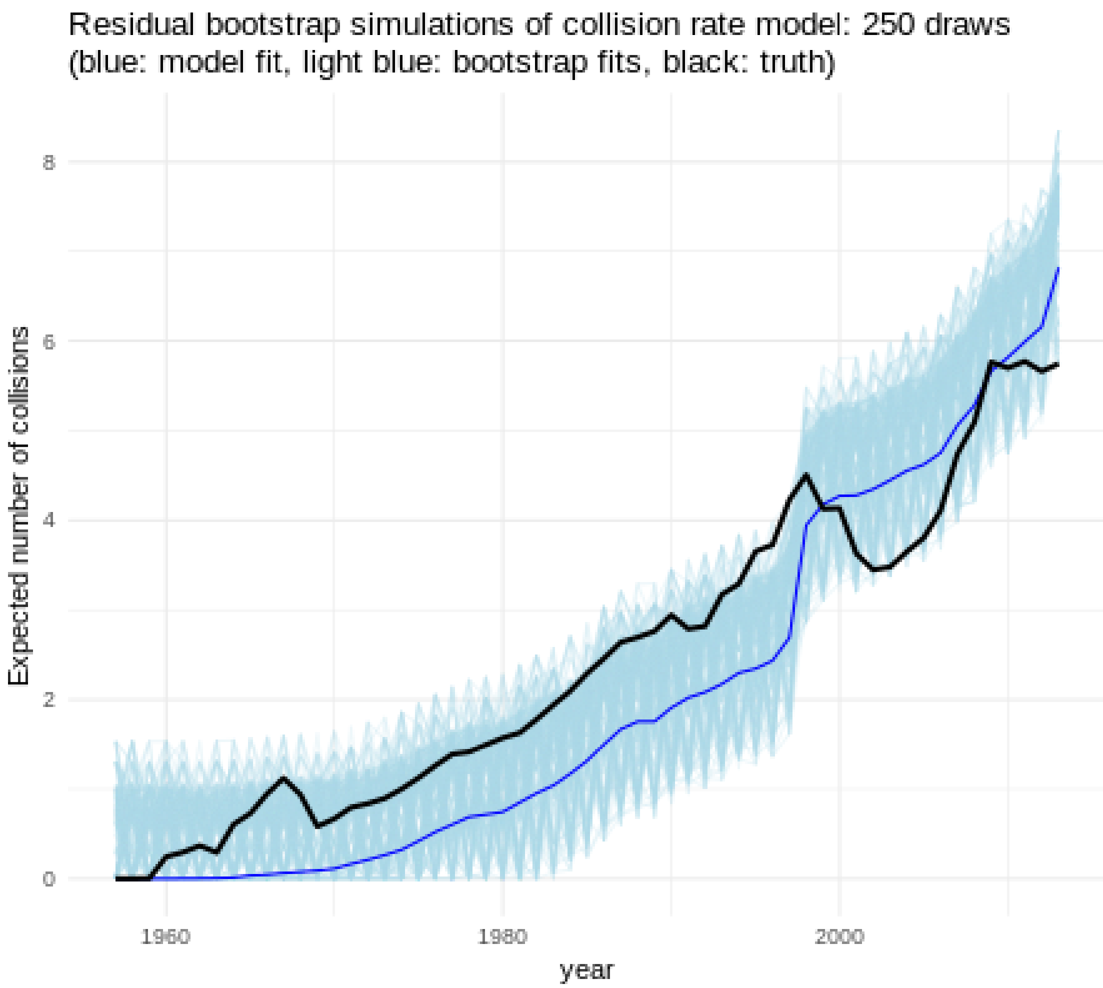

Figure G.3: *Bootstrapped collision rate model projections.*

Equation 4.12 projections generated from residual-bootstrapped parameter estimates.

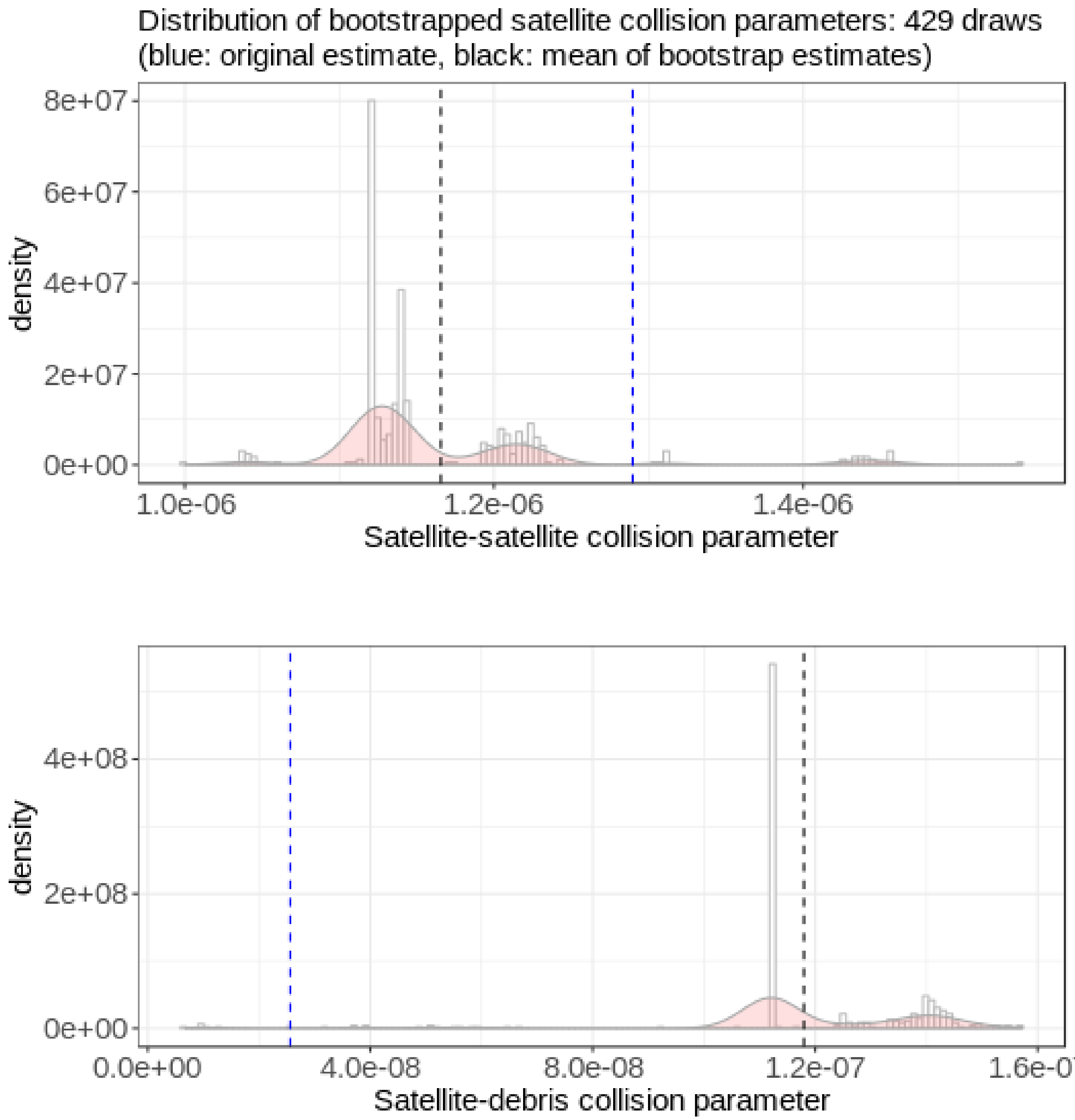

Figure G.4: *Residual-bootstrapped parameter estimates for equation 4.12.* The lack of asymptotic normality reflects the estimation issues described above.

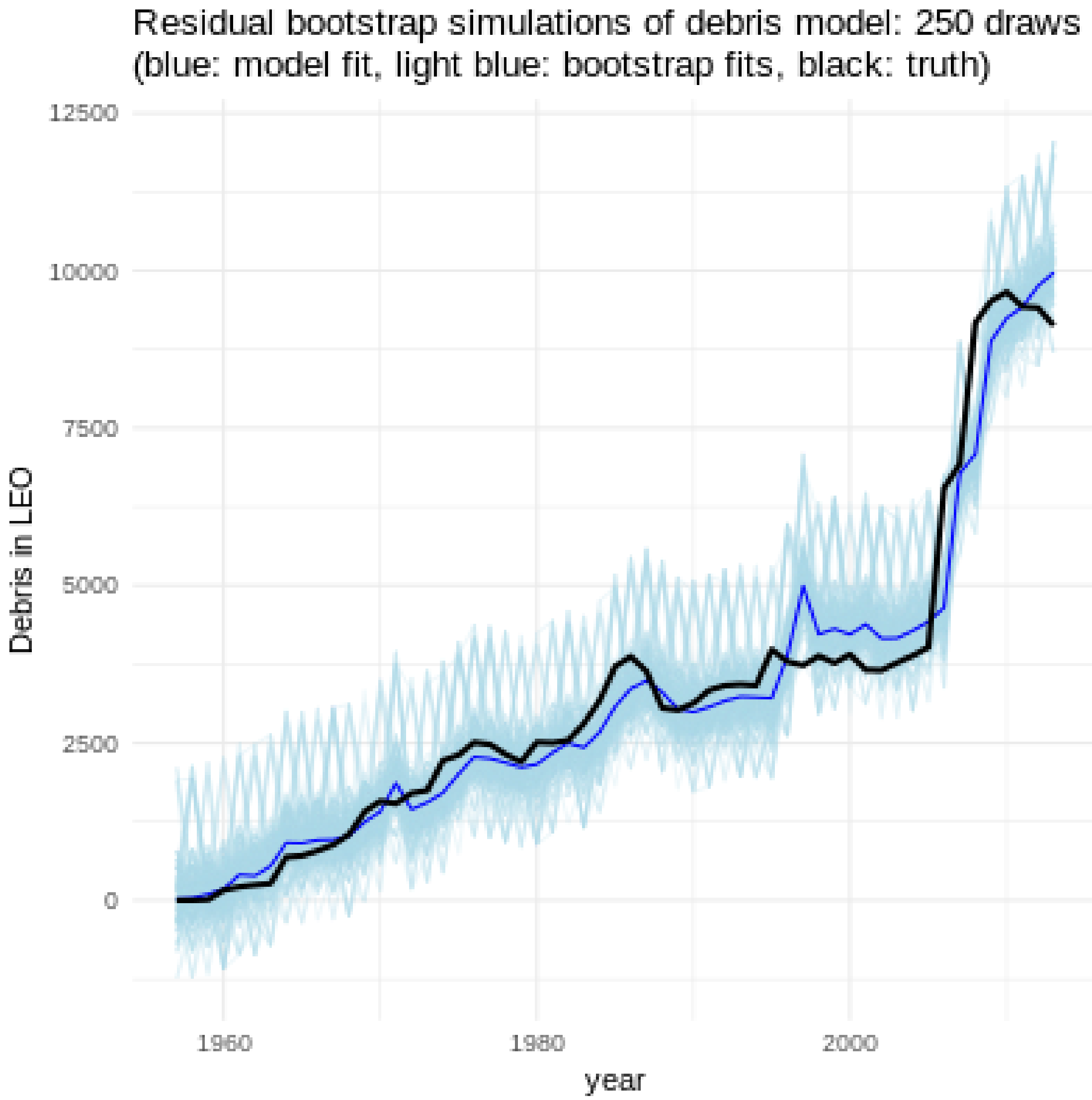

Figure G.5: *Bootstrapped debris law of motion model projections.*

Equation 4.13 projections generated from residual-bootstrapped parameter estimates.

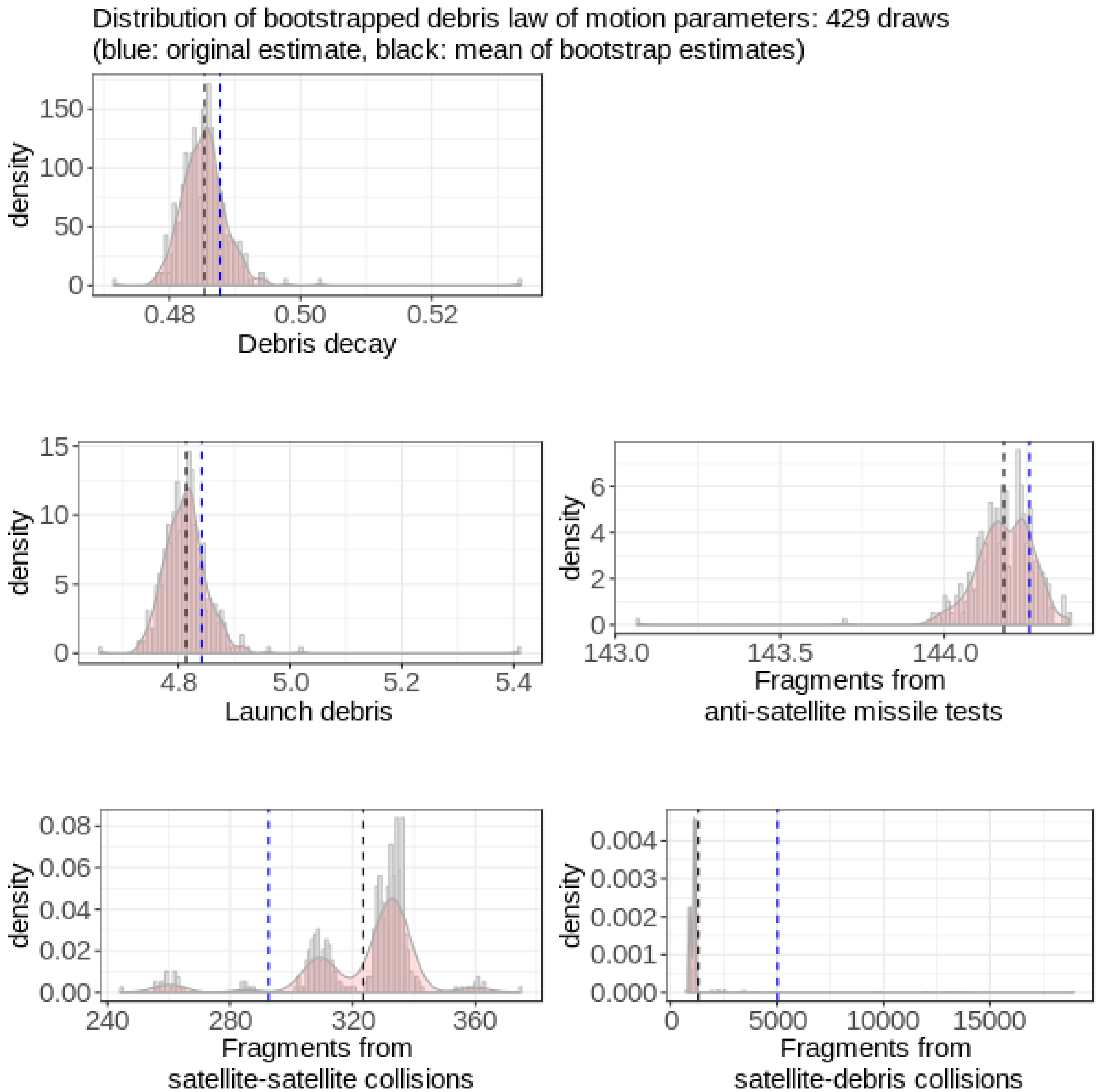

Figure G.6: *Residual-bootstrapped parameter estimates for equation 4.13.* The issues in estimating $\hat{\alpha}_{SS}$ and $\hat{\alpha}_{SD}$ propagate through to the distributions of $\hat{\beta}_{SS}$ and $\hat{\beta}_{SD}$. The remaining parameter distributions are not centered on the main model estimates due to ridge estimation bias.

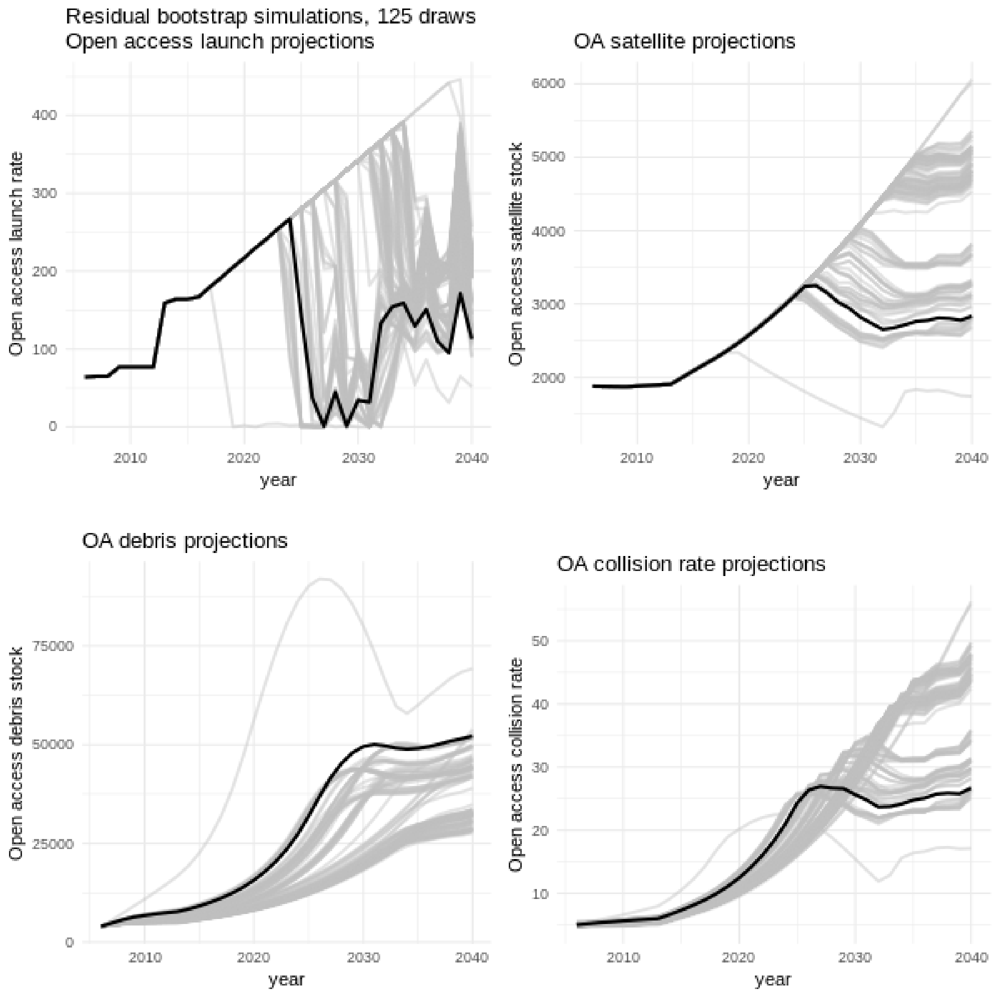

Figure G.7: *Bootstrapped open access orbit use projections.*

Open access orbit use projections generated from residual-bootstrapped physical parameter estimates.

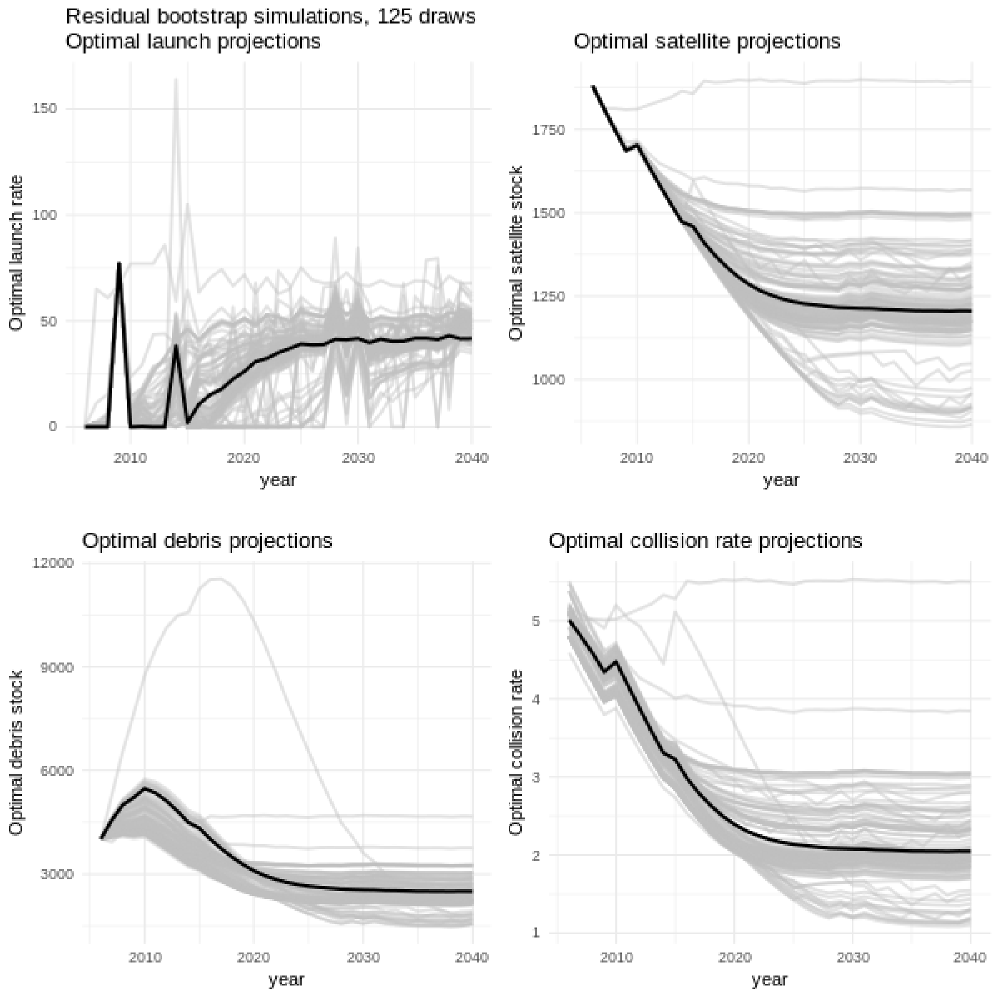

Figure G.8: *Bootstrapped optimal orbit use projections.*

Optimal orbit use projections generated from residual-bootstrapped physical parameter estimates.

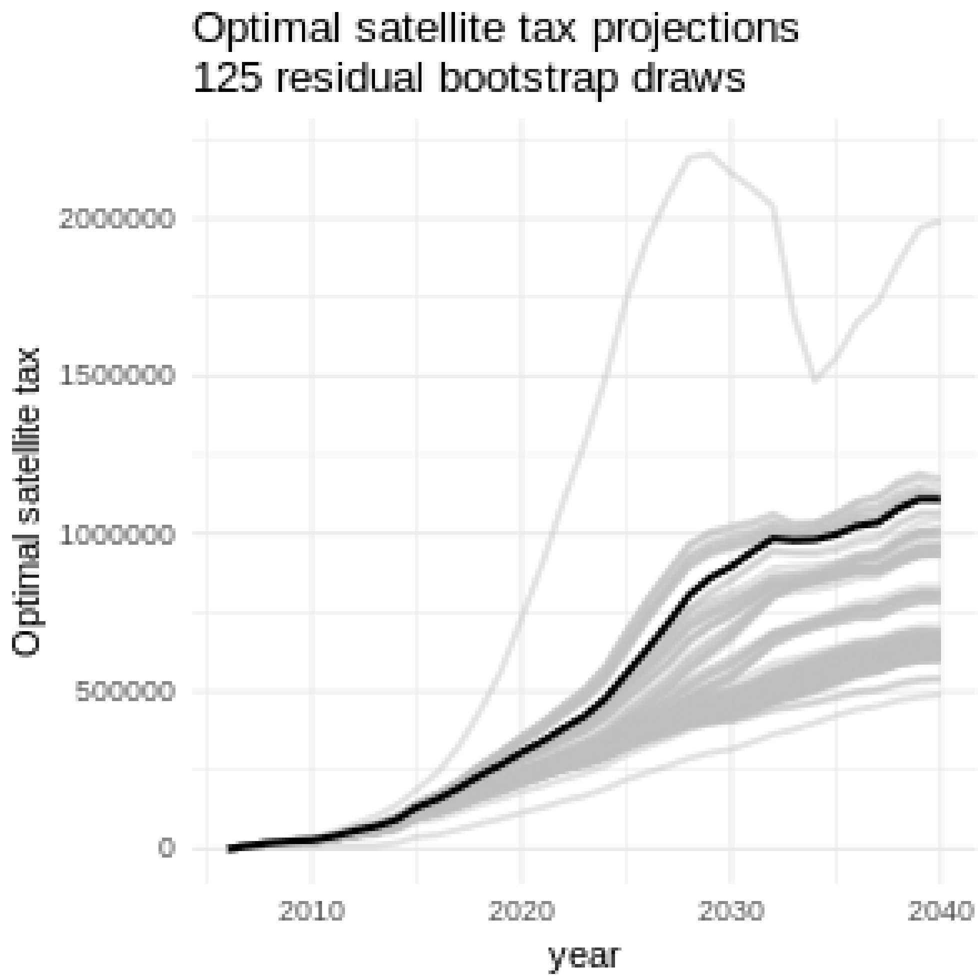

Figure G.9: *Bootstrapped optimal satellite tax projections.*

Optimal satellite tax projections generated from residual-bootstrapped physical parameter estimates.