2015

An Economic Analysis of Mars Exploration and Colonization

Clayton Knappenberger 15
DePauw University

An Economic Analysis of Mars Exploration and Colonization

Clayton Knappenberger

2015

Sponsored by: Dr. Villinski

Committee: Dr. Barreto and Dr. Brown

Contents

Figures

I. Why colonize Mars?

> "'Why?' The man asks 'why.' George, isn't there anything in your soul but discounts and dividends? Didn't you ever sit with a girl on a soft summer night and stare up at the Moon and wonder what was there?"
> "Yeah, I did once. I caught a cold."
>
> –Robert Heinlein, *The Man Who Sold the Moon*

The systematic study of the heavens is probably as old as agriculture. The periodic features of the sky lent themselves well to the seasonal demands of planting and harvest. And agriculture is inextricably connected with the advent of human civilization - that is a settled and record-leaving society. Thus, the universe has always been tied to humanity's future. Our ancestors' lives depended on their ability to understand the universe in which they abided. Aside from perhaps the moon, no single celestial object has garnered the sustained attention which Mars has. This public attention has ebbed and flowed with the politics of this and the past century, but it has remained present. At the beginning of the twentieth century, the belief that Mars is inhabited was a widely held one. By the end of the same century, we knew not only that Lowell's Martian Canals were merely an optical illusion, but we had actually landed a rover on its surface and sent back images of a barren landscape. The current century is likely to include at least one manned mission to Mars and could very well end with colonies on the Red Planet.

No matter under what marketing this next step for humanity is packaged, few doubt that in the future mankind will be an interplanetary species. Their reasons are varied. Many, including Elon Musk (founder of SpaceX), believe that the prospect of a global catastrophe necessitates the colonization of other planets to preserve the human species. Global climate change, Malthusian limits, asteroid collisions, and nuclear war could each spell the end of humanity. By claiming and colonizing other planets we, as a species, are better equipped to survive the dangers. At the very least disaster on one planet would not mean that humanity is wiped out. This thinking, while correct, is often difficult for many to accept. Judge Richard Posner (2004, pg. 27) reports the findings of the UK Task Force on Potentially Hazardous Near Earth Object that an asteroid 7 km wide would cause the direct destruction of an area approaching the size of Australia, prolonged climate effects, and mass extinctions - while an asteroid 16 km wide would likely cause the extinction of the human species.

The average frequency of collisions is 1 in 10 million years for the smaller sized asteroids (7 km diameter) and 1 in 100 million years for the larger variety (16 km diameter). This makes for a combined annual probability of approximately $1.1 * 10^{-7}$[1]. Few people would bother investing resources today to insure against that happening. No one alive today can reasonably expect the world to end in their lifetimes given such low probabilities. Even if we combine the probability of catastrophic global climate change and other doomsday scenarios, few people find it reasonable to advocate the colonization of other planets as an insurance policy against human extinction. This is a public good which not even national governments can be expected to invest in.

[1] Assuming that asteroid collisions follow a Poisson process with $\lambda_A = \frac{1}{10{,}000{,}000}$ and $\lambda_B = \frac{1}{100{,}000{,}000}$ then the combined annual probability of these two categories of asteroids colliding with Earth equals:
$\frac{1}{10{,}000{,}000 * e^{\frac{1}{10{,}000{,}000}}} + \frac{1}{100{,}000{,}000 * e^{\frac{1}{100{,}000{,}000}}} \cong 1.1E - 7$

Others attempt to argue for colonization using a sort of cosmic manifest destiny for humanity. Zubrin predicts another kind of disaster for humanity without a Martian frontier: "the entire global civilization based on values of humanism, science, and progress will ultimately die." (1996, pg. 297) His fear is that without a Martian frontier, there will be no place to give birth to new cultures. Globalization will continue to break down the walls of cultural diversity leading to the stagnation and decay of human culture. Moreover without the challenges and difficulties imposed by frontier life, technology will also stagnate. This final argument he justifies by saying that if technology had progressed at the same rate as was experienced from 1906 to 1966 we should have videotelephones, solar powered cars, nuclear fusion, inexpensive Earth orbit access, and a host of other technologies (underwater cities).

To the twenty-first century observer such comments cannot help but seem somewhat naïve. Globalization has indeed brought Coca-Cola and McDonald's to Beijing, but it has also brought Miyazaki films, Yoga, and Shawarma to the United States. The trade may seem one-sided for now, but there is no reason to believe that the future of humanity holds no hope for cultural diversity. The early twenty-first century has also seen an explosion in new technologies which cast doubt on Zubrin's technological argument. The internet has given us videotelephones in Skype, Facetime, and Google Hangouts. In addition we are reaching toward the realization of inexpensive Earth orbit access with the development of reusable rocket technologies. We may not have solar powered cars, but we do have electric cars, hybrid cars, and soon self-driving cars. Reports of the death of technological progress on Earth are greatly exaggerated.

A third group will attempt to garner support for space exploration by appealing to nationalism or to economic interests. Space exploration has always been connected to national prestige. The space race of the 1950 and 60s cannot be viewed except through Cold War lenses. Those advocating space exploration on nationalistic grounds soberly inform the reader that other nations are surpassing us in outer space. They point to the retirement of the Space Shuttle and the use of Russian rockets to bring American astronauts to the ISS as a sign of American decline, seeking to generate fear (Sterner, 2011). Then they point out the advances which China, India, Russia, and even Europe are making in space exploration. To them, the success of Chinese Lunar missions and Indian Mars missions are not a success of mankind, but an American failure.

On the economic front, nationalists will remind readers of the many ancillary effects which space technology has had on the economy including: include zippers, cordless appliances, and a wealth of new medical information. Space exploration is therefore a means of ensuring America's preeminence in the 21st century by capturing the political and economic capital which such ventures generate. Never mind that through international trade of goods and ideas those gains are spread across the world. Why should the United States invest billions of dollars into space when we can get many of the technological benefits through trade?

Why then should we colonize Mars? Perhaps surprisingly, even though each of the above arguments has its weaknesses, they each contain a kernel of truth. Refocused and combined, they craft a strong and compelling justification for the colonization of other worlds. While the probability of any single humanity-wide disaster is extremely low, it is irresponsible to ignore a single effort which can help mitigate (and perhaps help prevent) the damages of all these potential disasters. Return for a moment to our potential asteroid disaster. We determined that there is an annual $1.1 * 10^{-7}$ probability of just such a disaster, but we neglected to incorporate the costs of such a disaster. That is to say, we should look at the expected cost of an asteroid collision, not just its probability of occurring. We said that a 7 km diameter asteroid would cause direct damages over an area the size of Australia as well as prolonged climate effects and mass extinctions. Assume that the asteroid does crash into Australia. Then immediately the losses include roughly $1.4 trillion in real GDP that won't be produced in that year. Thus the expected annual cost of asteroid collision is $154,000. Assigning, quite arbitrarily and coarsely, a value of $2.68 million[2] to each of the 23.3 million lives currently inhabiting Australia we arrive at a staggering total loss of $\$6.3844 * 10^{13}$ and an expected annual loss of $7,022,840.

This is in reality a very low estimate of the real expected annual loss. GDP is a flow variable and the absolute destruction of Australia would result in far greater economic losses than the paltry $1.4 trillion of goods and services created in that year – it would mean the destruction of the capital which generates those goods and services *every* year. Moreover the human life is undoubtedly worth more than the $2.68 million earned over a lifetime in the labor force, there are non-pecuniary losses which should be considered too. The prolonged climate change we experience as a result of this asteroid collision would lead to major disruptions in agriculture resulting in further pecuniary and human losses that dwarf the direct costs of such a

[2] An extremely rough estimate of the value of an Australian life would take Australian GDP per capita (roughly $67,000 and multiply it by the say 40 years an Australian will be in the labor force giving us $2.68 million.

collision. But $7 million serves, for our purposes, as a lowball estimate of the minimum damage in our asteroid example (a 7 km collision)[3].

There are a myriad of other ways the world could end. Global climate change could cause similar disruptions to the world economy and if nothing is done to slow or halt its progress, the probability of a human-wide disaster increases every year. Even barring a human-wide catastrophe, the costs of global climate change are mounting as weather becomes more volatile and more extreme. Colonizing Mars will provide a contingency plan for humanity. Even if Earth faces catastrophe, so long as there are people living on Mars, humanity has a future. The exploration of Mars will also provide valuable scientific information about the history of the Martian climate that can help us understand the forces that shape climate change on Earth (F. Taylor, 2010, pg. 212). We believe that the ancient Martian atmosphere was thick with carbon dioxide and other greenhouse gases and the recent launched Mars Maven Orbiter will collect data to help us better understand what happened. (Overbye, 2014) Greater understanding of these mechanisms could provide tools to help us preserve our first planet's future.

Politically speaking there is also much to be gained from colonizing Mars. Space policy has always been a field for national rivalries, but it has also provided opportunities to demonstrate international cooperation. President Nixon used space policy to pursue détente with the Soviet Union through the join Apollo-Soyuz mission. (Ross-Nazzal, 2010) President Clinton chose to advance cooperation with Russia in the creation of the International Space Station as part of an effort to support stability in the newly emerged country. In addition,

[3] Another reason this estimate severely underestimates the true damage of an asteroid collision is that it assumes no population growth and no economic growth. If the Australian economy grows at 2% over the next 1,000 years then its real GDP will skyrocket to astronomical proportions and so would the expected losses.

Clinton hoped to keep Russian rocket technologies from proliferating. (Lambright, 2014, pg. 118) These are two historical precedents for the unifying effect space exploration can have in an international setting. Space policy has traditionally been both a source of national rivalry and a source of international cooperation.

There is no reason why Mars exploration cannot fulfill similar foreign policy objectives. Carl Sagan and the Planetary Society encouraged the idea in the late 1980s that the United States and the USSR should collaborate on Mars exploration. (Lambright, 2014, pg. 100) While the Soviet Union no longer exists as a geopolitical rival to the United States, the underlying benefits of international collaboration in space have not changed. The heir to the Soviet space program, Russia, remains at odds with the United States in many foreign policy arenas. In addition, China has risen as both a space power and as a global power. Sino-American relations are also tense – illustrated by their exclusion from participating in the ISS. Mars can help relieve tensions here on Earth by serving as an area for collaboration between rival nations. Cooperation scientifically not only reduces tensions through détente, but because space policy is typically linked with national security through missile defense systems and communication satellites, cooperation in space can encourage breakthroughs in national security issues as well.

Finally the colonization of Mars could help answer many of the social, economic, and political problems of the future. While there is little to be gained from attempting to calculate the carrying-capacity of the Earth, it is clear that the scarcity of resources and human population growth both have the potential to make human life harder in the future. By opening up the resources of another world, the pressures of population growth and scarce resources might be lifted. Particularly attractive to Mars is its proximity to the asteroid belt. Asteroids contain many of the precious metals which are valuable here for industrial uses and having an abundant supply at hand would prove beneficial to a resource-starved future. By offering new opportunities, Mars can also improve the lives of billions of people. Just as the politically and socially ostracized of Europe colonized the Americas, so will the outcasts of humanity find Mars calling. A Martian society will by needs be politically more open and free from old world politics the political outcasts of the future will be able to find greater agency on the Martian frontier than what was available on Earth.

The absolute difficulty of colonizing Mars will also stimulate technological advances. Those technological advances fuel economic growth by allowing us to do more with less. The primary forces which will act on Martians are the expense of importing products from Earth and the great expense of labor. Transportation costs will make the export and import of many products economically unfeasible and raise the price of whatever is feasible. Thus entrepreneurial opportunities exist both for innovations which decrease transportation costs and for the production of those goods which cannot be transported. Martian labor will also be incredibly expensive relative to Earth. The first settlers are likely to be highly trained scientists and professionals. With a significant scarcity of unskilled labor and a high opportunity cost in using the labor that is available, for almost all production capital will be cheaper than labor. The Martian economy will therefore be incredibly capital dependent and there will exist entrepreneurial opportunities to replace labor with capital on all fronts. These economic pressures on Martian society are likely to result in technological advances and increased productivity. This will shift out the production-possibility frontier for all of humanity allowing us to make more and for less than what was originally possible.

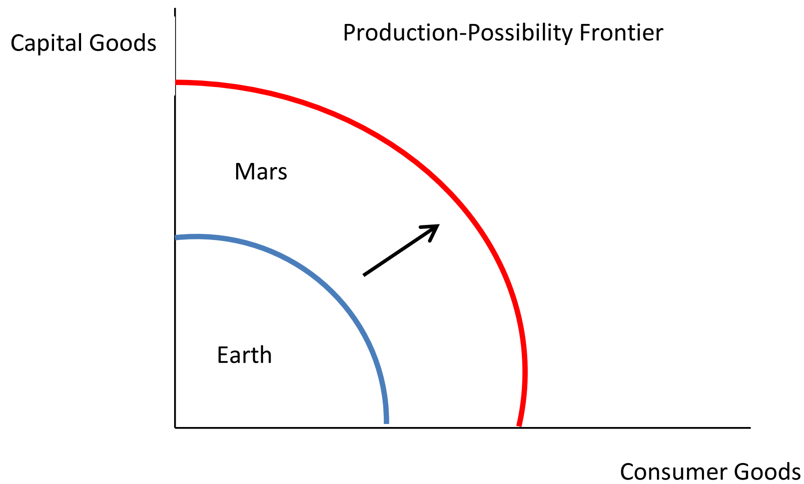

Figure 1 The Production-Possibility Frontier and Technological Innovation

But there is a final and less exciting reason why we should explore and colonize Mars. Doing so will add to the sum of human knowledge and color our understanding of the universe we inhabit. Space advocates often seem hesitant to argue this. They prefer to publicly ground their arguments in the many secondary benefits which come from exploration while personally they are more interested in the discoveries exploration will bring. Scientists and space advocates need not be afraid to say that science is itself a useful endeavor. Even if seeking to put humans on Mars does not add a single new product to the economy or contribute in any way to greater cooperation among the nations of Earth, it will still benefit humanity by providing a continual stream of new information about our universe and mankind's place in it.

This all helps answer why we should colonize space, but it does not necessarily mean that we should colonize Mars. There are other options for human expansion into outer space. Most attractive among these options is the Moon. A Lunar colony would have a significant advantage in communications and in resupply. However, Mars has several advantages that make it a far more attractive destination (especially for extended stays). To begin, lunar crops would have to rely on electric lighting since the Moon roughly 28 day cycles of light and darkness. Mars also has a significant resource and energy advantage over the moon. Industrial society as we know it depends on metals which are deficient on the Moon relative to Mars. In addition, Mars has geothermal energy and the potential for wind energy generation. The clincher, though, is that Mars has water in abundance while the moon has almost none. (Zubrin and Wagner, 1996 pg. 220 - 222)

II. Can We Colonize Mars?

Mars is there, waiting to be reached.

- Astronaut Buzz Aldrin, *Buzz Aldrin: Down to Earth*

For all of our technological advances, many believe that the best days of the American space program are far behind us. We bemoan the fact that 47 years have passed since the first successful Mars mission (Mariner 4 in 1964) and argue that this is far too long a time to have to wait for a manned mission. We point to the constantly shifting timelines as evidence that it will never actually happen. Manned missions for Mars are always 20 to 30 years on the horizon. This seems too long to wait. After all, America was able to land on the Moon 7 years after Kennedy's speech. It is tempting to think that a timeline for Mars should be similar. This metaphoric comparison between the Apollo program and a manned Mars mission is however, flawed.[4] Careful reflection will force us to acknowledge that it has only quite recently become

[4] Apollo is constantly in use as a metaphor within the space community and outside it. Even the creation of the National Cancer Institute during the Nixon administration was considered an Apollo program for Cancer. In most cases this application of the metaphor is inappropriate. Apollo, while successful, was not the end-all model for space exploration. Assuming that what worked once will work again (and in the same way) is damaging to the space program.

realistic to discuss landing humans on Mars. Many of the prerequisites for a manned mission have only recently become realizable.

It is common to hear Mars advocates argue that the technology exists today to send people to Mars. As Zubrin and Wagner, (1996, pg. xix) states "Exploring Mars requires no miraculous new technologies." Beware such bold pronouncements. We may have been able to reach Mars for a long time, but we have yet to bring anything back. Every single Mars mission to date has placed an object in orbit of Mars, landed an object on the planet's surface, or done both. These unmanned missions are trivially simple compared to the difficulties of a manned mission, but the history of unmanned missions is replete with failure. As recent as 1998 and 1999 two subsequent American Mars missions met with failure. A navigation error caused the Mars Climate Orbiter to enter the Martian atmosphere and break up, while the Mars Polar Lander never made contact after its expected landing. (Stooke, 2012) Moreover, we have learned a great deal in the past decade about even greater dangers to sending humans to Mars.

Manned Mars missions are by necessity much longer than the Apollo missions. It takes approximately 6 months to reach Mars and astronauts would then be forced to wait another 20 months before being able to return on another 6 month journey. Clocking in at over 2 ½ years roundtrip this would easily break Sergei Krikalev's time in space record by over 170 days. Kirkalev's record was spread over several missions to *Mir* and the ISS. No human has ever spent that much uninterrupted time in a low gravity environment. Medically, we know from experience aboard the ISS that low gravity has causes muscles and bones to atrophy. Low gravity also causes blood loss which causes the heart atrophy as it does not need to work as hard to pump blood throughout the body. This could cause problems if the astronauts need to perform manual labor upon reaching Mars. While a crew would be able to adapt more quickly to the easy Martian gravity, six months of bone, muscle, and blood loss are going to have their toll. The first few months of a Martian visit could be the most dangerous if it requires the crew to establish a safe and habitable base.

It may be common in Mars advocacy communities to feel as though the time spent on the International Space Station has been an unwelcome distraction from the ultimate goal of a manned Mars mission, but the experience we can now draw on will prove invaluable to the success of sending humans to Mars. Without the ISS we would simply not have the wealth of medical information now available to us. Apollo program mission times were simply not long enough to teach us what we need to know in order to prepare a crew for 2 ½ yearlong missions. Trying to send humans to Mars without that information would have led to costly and tragic mistakes. With a permanent presence in low Earth orbit, it is finally possible for us to study and devise means of addressing the adverse consequences of low gravity on astronauts.

Unfortunately low gravity is not the only challenge facing a manned Mars mission. On Earth we have the glorious luxury of an atmosphere perfect for life. It protects us from solar and cosmic background radiation. Any manned mission to Mars would spend six months traveling through open space where there is no protection from radiation. Then they would spend another twenty months on the surface where there would only be partial protection from Mars' much thinner atmosphere. Until very recently we could not quantify the dangers an astronaut would face on Mars and it was generally assumed that astronauts on the surface would be, mostly, safe from solar and cosmic background radiation. The Curiosity rover paints a much different picture. The average radiation dose of a trip to Mars would be approximately

300 mSv, and while the Martian atmosphere provides some protection against cosmic background radiation, the protection it offers against solar radiation is inconsistent. (Gifford, 2014) The best estimate for a roundtrip dose, including a 500 day excursion to the surface, is 1.01 Sv. This corresponds to a roughly 5% lifetime increased cancer risk. (Wall, 2013) While many astronauts would no doubt willingly accept this increased risk, longer missions are hindered by the present absence of effective radiation shielding.

It is for this reason that the first missions to Mars might skip an orbiting stage. The Apollo landings on the moon were preceded by lunar orbiting missions, but unless radiation shielding is dramatically improved in time for the first mission, the best way to protect astronauts from radiation is to land them on Mars. A landing mission would also be the most useful scientifically and nationally. Having astronauts on the ground dramatically increases the number of opportunities for scientific inquiry. But landing will only partially solve the radiation problem, and adds another problem which might at first seem ridiculous.

How do we land on Mars? Returning to Earth, astronauts take advantage of the Earth's thick atmosphere to slow down before crashing into an ocean or into the steppes of Kazakhstan (as the Soyuz capsule does). In addition to atmospheric drag, capsules landing on the Earth typically rely on parachutes and landing engines. But since the Martian atmosphere is roughly 1/10th that of Earth's, atmospheric drag and parachutes will be less effective at slowing a landing vehicle. In addition, a landing vehicle will be far more massive than the capsules we return to Earth since it must carry things which are essential to the life of the astronauts in addition to scientific equipment and the astronauts themselves[5].

[5] There is the possibility of performing multiple landings to lower the mass of each landing. This would raise costs a

Vikings 1 and 2 were able to make powered descents, but each had a mass of only 600 kg. (F. Taylor, 2010, pg. 41) The rovers Sojourner, Spirit and Opportunity each made use of airbags to assist in their landings. (F. Taylor, 2010, pg. 115). The massive 2,000 pound Curiosity required a new Sky-crane landing system, (Coulter, 2012) but a manned Mars mission will no doubt require far more than just 2,000 pounds of equipment for a 20 month mission. A landing vehicle capable of bringing a crew and the supplies they need to Mars does not exist and has not been designed.

The solutions to each of these problems, and to other problems not mentioned, certainly exist. None of these problems or their sum is insurmountable. These questions mainly require engineering answers and not new theory. Our understanding of medicine, radiation, and reentry mechanics gives us the tools to devise solutions to these questions. But the solutions are not currently available and would require years of design and testing to be deemed acceptable. So can we successfully send people to Mars using today's technology? We probably can, though at great risk to the health of the crew, the success of the mission, and at great cost. This mission could not be carried out at the next launch window though. My sense is that at least a decade long program aimed at answering the above questions would be required before the technologies necessary to bring mission risks down to acceptable levels could be developed and deployed.

It is currently not technologically possible to colonize Mars. The risks associated with interplanetary travel, given today's technology, are such that a Mars colony cannot rely solely

great deal and would also add a different technological problem: how to transport each lander to Mars. Sending multiple landers increases the overall risk of the mission since success of the mission now hinges upon multiple landers arriving safely onto the surface. If a single lander fails the mission may be unsalvageable.

on Earth exports for their every need. Shipments are bound to be lost, and this would jeopardize a colony. In addition the costs of interplanetary travel would make such an approach prohibitively expensive. Things like food and water must, at least to some degree, be provided by those living on the surface. While there remain many questions about how to grow food on Mars, a recent MIT study examining the Mars One mission plan found that a colony growing all of its own food would produce unsafe levels of oxygen for the habitat. (Do et al, 2014) The technology for removing oxygen from a habitat does not currently exist. Colonization also increases the need for radiation shielding. While it may be acceptable to ignore a five percent increase in cancer risk for a single mission to Mars, the risks increase dramatically above five percent if you plan on staying on Mars. The kinds of technologies do not currently exist and may not exist until temporary missions to Mars are able to provide further information about what awaits us on the surface.

III. What would it look like?

Believing that humanity should devote the resources necessary to exploring and colonizing Mars begs the question of how it should be done. So far we have focused on the technical or mission aspects of the Mars program. Very little research has been done on who should be carrying out this program. The limited attention to this question is understandable given that until recently national governments were the only serious actors in outer space. That has and is changing. Not only are private groups reaching outer space independently of government programs, but there is a growing trend toward international programs where several national agencies collaborate on a single project. Examples of this international trend are the European Space Agency and the International Space Station. Thus, asserting that we should explore and colonize Mars obliges us to answer the question: who should do the exploring? The next sections will introduce some of the key institutional features of each option including some of the strengths and weaknesses of employing that institution for space exploration.

A. National Program

> We choose to go to the moon. We choose to go to the moon in this decade and do the other things, not because they are easy, but because they are hard, because that goal will serve to organize and measure the best of our energies and skills, because that challenge is one that we are willing to accept, one we are unwilling to postpone, and one which we intend to win, and the others, too.
>
> - John F. Kennedy, *Speech at Rice University*

The national program is the most familiar and historically the most successful. National programs can take credit for of the firsts in space exploration. A few of the Soviet space program successes include: launching the first artificial satellite, putting the first animal in orbit, putting the first human in orbit, and landing the first rover on the moon. The successes of the American program are even better known but a few worth mentioning are the Mars rovers and the six manned lunar landings of the Apollo program. These successes are an important reason why the national program will always remain a tempting option for future Mars Missions as an arrangement option. There are a few key benefits to choosing to operate through a national government. First, a national government can often find political gains to space exploration in addition to whatever monetary or scientific gains exist. There are the obvious benefits which politicians see in having large space programs in their districts (most notably the high-paying jobs which are sure to follow), but there are also general national political gains to be had.

The national political goals behind space exploration are twofold: national security and prestige. Both of those goals are explicitly named in the initial report issued by President Eisenhower's Scientific Advisory Committee following the Soviet launches of Sputnik I and II. The argument goes that any nation capable of participating in the exploration of outer space must been technologically advanced and economically powerful relative to other nations in the world. Moreover, the technology used in exploring space has military applications and so successful launches demonstrate a nation's military might indirectly. As Eric Sterner, a Fellow at the George C. Marshall Institute put it, "It's kind of pitiful when you think about it: for a Superpower not to be able to put people in space." (Sterner) Space will always remain an arena for demonstrating national power. A Mars mission done as a national program would no doubt be sold, at least in part, with nationalistic rhetoric.

But there are other benefits to choosing a national arrangement for a Mars mission. A national government can harness vast resources toward the completion of a mission and this can be done in a short period of time. Even at historically low levels, the 2014 NASA budget of $17.5 billion is nothing to laugh at. Those funds came out of taxpayer pockets and required no independent financing. The coercive power of government means that taxpayers cannot withdraw their support for a Mars mission that does not meet their demands (or cost-benefit analysis) the way a shareholder or creditor would.[6] In addition, a national program is shielded from many kinds of risk associated with for-profit organization. Namely a national program can suffer disasters without shareholder revolt. While there are certainly political costs to a

[6] Although given time, dissatisfaction from taxpayers could spell disaster for the program through election platforms and results. Space typically is not typically a central election issue, however and so it would be rare to see a candidate lose election over space policy.

disaster, it is often the case that a national program is able to weather the storm because of continued support by politicians who recognize that the setback is temporary and who wish to keep the federal money flowing into their district.

There are downsides, however. Politics does not value efficiency. Congressmen are expected to exercise their influence to bring federal money into their home districts. Federal money brings jobs and is easily used to garner support for reelection. This competition over federal money opens up the federal budget to agencies which are willing to pander to congressmen for funding. Thus the establishment and maintenance of a research center has as much to do with congressional politicking as it does to do with the kind of research being done. One reason Mars missions have had such strong advocacy is that the Jet Propulsion Lab in Pasadena California depends on these missions for its survival (Lambright, pg. 14) and is not afraid to use political influence to further that goal. (Lambright, 2014, pg.86)

Thus research is done not where it might be most effectively done, but where there is sufficient political influence to direct federal money. Once a program finds life, it clings to life even when it might have outlived its usefulness, and because there was sufficient political influence to establish the program in the beginning, that influence often remains nearby to protect it from outright destruction. A profit-maximizing firm would shut down an inefficient facility, while politics can protect and encourage the growth of inefficient facilities. In exchange for political protection against temporary setbacks we lose market discipline.

Sustained public support for space exploration is also hard to muster. Without that governments won’t fund an expensive space program. Public support depends on having a continual stream of exciting successes to report, or on having a sufficiently strong narrative to keep the public's attention. Unfortunately space science is often qualified in its findings. Scientific accuracy demands caution in making claims. Evidence is often ambiguous and reasonable people will disagree. The public however, cannot maintain interest in the face of such ambiguity and the fact of ambiguity is seen as itself a failure of the mission. For this reason, the Viking mission was considered a failure because it did not return conclusive evidence of life on Mars, never mind that it did not produce conclusive evidence of no life on Mar either. (Lambright, 2014, pg. 69)

Finally, political processes can undermine space exploration through changes in leadership and focus. Each President has his own vision for space exploration, and this vision is rarely consistent with the visions of previous Administrations. President George W. Bush wanted the United States to return to the moon, this was scrapped by the Obama administration in favor of an asteroid landing. Inconsistency is particularly damaging in space exploration where programs are rarely accomplished in the four years of a single presidential term and typically not even in the eight years of a reelected President. Launch windows for Mars come every two years and developing a mission often takes much longer than two years. A mission begun at the beginning of one term is unlikely to be ready for launch before the next election. Thus, a change every four years is very sudden in the timespan of a typical NASA program.

Presidential inconsistency is compounded by the fact that the NASA Administrator is a politically appointed office. Not including acting Administrators, the average time each

Administrator spent in the post is approximately 1,645 days or about 4.5 years. Not surprisingly this corresponds almost exactly to presidential terms where an incumbent Administrator is usually replaced at the beginning of a new presidential term. Each Administrator brings his own goals, and often his own personnel, to NASA and this has wide-ranging consequences for every NASA program. Leadership at NASA is therefore inconsistent at all levels to the detriment of space exploration.

B. Private Model

> In the coming era of manned space exploration by the private sector, market forces will spur development and yield new, low-cost space technologies. If the history of private aviation is any guide, private development efforts will be safer, too.
>
> - Burt Rutan in an interview with *Wired* Magazine

Historically, the exploration of outer space has been limited to governmental groups. This barrier has begun to crack and in large part it seems to have been in response to the retirement of the Space Shuttle. Since then, two private companies – Orbital Sciences and SpaceX - have resupplied the International Space Station in fulfillment of their contracts with NASA. In addition, NASA has also contracted with SpaceX and Boeing to provide astronaut transportation to the ISS in the future. Finally, a host of new companies, Planetary Resources and Virgin Galactic for example, have captured the public's imagination with their plans to profit from space-related ventures. It is common now in the news to hear about the privatization of space. There is little historical experience to draw upon, but there are a number of features which are characteristic of private organization and will certainly apply to any private for-profit model of space exploration.

If we assume that a private group is primarily concerned with profit maximization[7] then there is reason to believe that a private group would be more sensitive to risk, uncertainty[8], and the time value of money than an established government agency. This increased sensitivity becomes apparent when we view a private group's decision of whether or not to engage in space-related activities as an investment decision. The for-profit organization will choose the endeavor which maximizes the expected present value of their investment. In general the present value function will take the following form $PV = \sum_{i=1}^{n} \frac{R_i - C_i}{(1+r)^i}$ where R_i are the expected returns in the i^{th} year and C_i are the expected total costs of an investment in that year. Total costs include operational costs and also any organizational costs which must be paid up front. Since the expected costs and benefits are spread over n years they must be discounted by the rate r to account for the implicit opportunity cost of an alternate investment. Uncertainty is not incorporated into the present value model, but is assumed to influence how confident the investor is in his estimation of returns and costs. Hence high uncertainty will make him less confident in his present value calculation, while low uncertainty will make him more confident.

We can begin to see some of the limitations of relying on private investment already. The present value function relies on *expected* returns and so R_i is the expected value (or central tendency) of a distribution of returns. The below chart shows two normally distributed random variables – for the current purposes, let that random variable be a firm's revenue. Both revenue

[7] While much of the discussion which follows devotes itself to private organizations which are profit-oriented, the reader should remember that a private organization can have other objectives than profit-maximization. Founded in 1980, the Planetary Society is a non-public and non-profit organization. In addition to their advocacy work, the Planetary Society is launching LightSail 1 in May, 2015 as the first part of their work in developing a solar sail. (The Planetary Society)

[8] For a discussion of risk versus uncertainty and its effects on commerce, the reader is encouraged to look at the Appendix. Its exposition will no doubt prove useful in the following sections, but those already familiar with the terminology may not require it.

distributions are centered on a mean of zero but the red revenue distribution is riskier – it has a higher standard deviation than the blue revenue distribution. This means that while we would expect each investment to provide the same revenue over a long period of time, the red revenue distribution will have larger year to year fluctuations in the actual amount of revenue returned.

These fluctuations can threaten the entire enterprise. A single accident can cause a great deal of harm to a private group whereas a public program might be better shielded by political backers and already established agency budgets. Look no farther than what has happened to Orbital Sciences following the explosion of an Antares rocket bound for the ISS. Overnight the stock fell almost 18% and the company began to face strong negative media attention over the accident. To make it worse, the negative attention was expanded to all of the private space movement following the crash of Virgin Galactic's Spaceship 2 a few days later. While both groups have successfully rebounded from these setbacks, the future of space exploration is sure to be filled with tragedy.

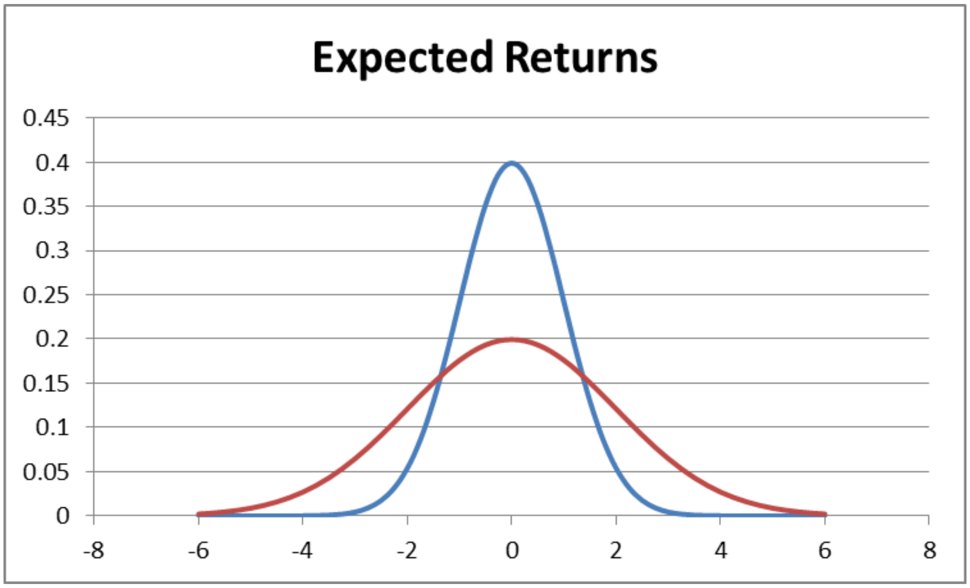

Figure 2 Expected Returns and Riskiness

For-profit firms are also more sensitive to the time value of money than a government agency might be. In our present value calculation, each year's expected net returns are discounted by a rate r. This is reflective of the fact that a dollar tomorrow is not as valuable as a dollar today. Returns expected in the distant future carry a smaller weight to the potential investor compared to what is expected to happen in the near future. Heavy start-up costs will need to be offset by even greater expected returns in the future to make the investment economically attractive. The farther off those expected returns are relative to starting costs, the greater they will have to be. Even if Planetary Resources expects to earn a fabulous return on mining asteroids, the present value of those returns might not be enough to justify the costs which must be devoted to creating spacecraft capable of mining asteroids.

Nevertheless, there are significant benefits to utilizing a for-profit model in space exploration and these benefits arise primarily from competition. The benefits can be broken down into two groups: product market and factor market competitive benefits. It is good to have competition in the market for the goods and service and it is good to have competition in the markets for those factors which allow firms to produce goods and services. The benefits of having competitive product markets are obvious. If the product market is perfectly competitive, then resources are allocated efficiently and are produced at the minimum of the long-run average cost curve. But even imperfect competition is often an improvement over a monopoly. This is because competition incentivizes cost-cutting that will increase the firm's profits.

A benefit arising from factor market competition which is central to our discussion is that it helps solve problems of information-asymmetry where managers and owners only have partially coinciding interests. In an imperfectly competitive manager market[9], a principle-agent problem arises because those running the firm know that they will not receive all the consequences of their decision-making. If, however, there is a competitive market for managerial talent, owners dissatisfied with the performance of their current manager can simply fire and replace him. On the other hand, a successful manager is likely to retain his job. Competition therefore encourages good performance and discourages bad performance.

One feature of private organization is that the market for managerial talent is often more competitive than the corresponding market in public organization. Consider the market for NASA Administrators. As stated earlier, NASA Administrators are typically appointed every four years with presidential elections. Because these appointments require Senate approval, there are significant barriers to changing NASA Administrators. A President not satisfied with how the Administrator is performing would find it difficult to replace his Administrator in the middle of a term. But more significantly, a successful NASA Administrator is likely to be

[9] An imperfectly competitive managerial talent market might be one in which there are simply few people with managerial talents available or in which there are barriers that prevent owners from replacing their managers. Such barriers can include employment contracts which limit the circumstances in which a manager can be fired or the long search time needed to find a suitable replacement.

replaced after a presidential election. The office of NASA Administrator does not seem to be tied to the successes or failures of his tenure, but rather to the Administrator's ability to tie himself to the president. A private space organization would likely have fewer barriers to replacing subpar managers than an agency whose managers are appointed by the political process and it would also be able to reward excellent managers in ways that are not available to public employees.

C. International Model

> The American and Russian capabilities in space science and technology mesh; they interdigitate. Each is strong where the other is weak. This is a marriage made in heaven - but one that has been surprisingly difficult to consummate.
>
> - Carl Sagan, *Pale Blue Dot: A Vision of the Human Future in Space*

Before we begin to discuss international cooperation as a model for space exploration we must consider what is meant by "international cooperation." Clearly cooperation between two or more national agencies falls under the umbrella of international cooperation. An example of this would be the International Space Station which was the result of cooperation between the respective American, Russian, Canadian, Japanese, and European space agencies. But should the European Space Agency (ESA) itself be considered an example of international cooperation within Europe or is it best understood as *the* European agency? For the purposes of our discussion, we will consider the ESA to be a very well-defined and ongoing example of international cooperation. Indeed there are several advantages to having an ESA-like structure over the ad hoc and limited approach of a project like the ISS.

Just like national and private models for space exploration, international cooperation has strengths and weaknesses. The most commonly argued advantage of international cooperation is cost-related. By spreading out the costs of a project over several nations, space activities can be carried out at a smaller per-nation cost than if they were done individually. The cost advantages of international cooperation can either be genuine or illusory. If international cooperation allows for existing national agencies to specialize in specific tasks or if international cooperation allows the nations involved to achieve economies of scale in space activities, the true costs of carrying out space activities will be lower than if each nation acted on its own.

On the other hand, the communication between nations necessary to carry out international cooperation is not costless. Diplomacy is rarely easy and the effort necessary to establish a framework for cooperation can be quite costly since it involves the use of technologies sensitive to national security. One of the most important and least appreciated functions of the United Nations is the reduction of transaction costs for reaching international consensus. Prior to its existence diplomats seeking international consensus had to spend a great deal of energy meeting individually with nations or on building interest in an international conference. Now with committees devoted to any conceivable topic meeting regularly, it is much easier to bring international attention to specific issues and to then build consensus.

Additional costs could come from redundancy. When two separate groups are working to achieve separate objectives of a unified goal, there will be redundancy between them. If these added costs outweigh any specialization or scale efficiencies generated by international cooperation, the total sum of the project may end up being greater than it would have been if carried out at the national level. However, even a more expensive international mission is likely to yield smaller per-nation costs. There are illusory cost advantages because a higher total cost is spread out over more participants so that each is paying less than they would on their own.

Beyond the cost advantages of pursuing international cooperation in space exploration, there are potential foreign policy gains. These include orbital debris mitigation, increased situational awareness, and an increased understanding of the Earth. We will primarily consider those advantages which would result from Mars-related cooperation. Cooperation can cement existing relationships with our allies. Many nations wishing to explore Mars lack the resources to do so on their own. An American commitment to meaningful international collaboration can provide additional opportunities for other nations to explore Mars and can provide options for new nations to explore our neighbor. It sends signals of good relations and of growing cooperation between two nations. (Broniatowski, et al, 2006)

There are also benefits to cooperation with rivals. Collaboration in space with our rivals will build a basis for future collaborations not only in space but in other areas. Historically, space collaboration was used as a means of defusing tension between rivals. This was the case with American/Soviet space collaboration. (Ross-Nazzal, 2010) Furthermore, by building scientific and organizational relations between rival nations, we build incentives to maintain good relations. Taking actions which jeopardize good relations between cooperating nations would mean jeopardizing existing cooperative missions and organizational ties. A nation which unilaterally withdraws from a cooperative effort harms its reputation and credibility. (Broniatowski, et al, 2006)

There are however downsides to an international model. Any international agreement is expensive to negotiate since it must define the goals of collaborating and the responsibilities of each participating nation. Each nation has its own constituencies to satisfy, including domestic industries and scientific communities. Complicating this negotiation are potential tradeoffs between politics and technology. Technologically, the most efficient outcomes might require sharing information between nations. Politically, this technology could very well be classified or subject to export restrictions. These negotiation costs do not disappear after the initial agreement is made. There are sure to be disagreements over the progress of any mission, and because the parties disagreeing are sovereign nations, they will require with some form of international arbitration.

Cooperation between nations also increases the risks of some kind of political disruption in the mission. Budgets are a perennial source of strain in any space agency, and increasing the number of nations increases the risk that over the lifetime of a mission, one of those nations will be unable or unwilling to provide the initially agreed upon funding. An inability to meet this responsibility could be the result of an economic crisis that puts pressure on the national budget, and unwillingness could result from a change in the balance of political power domestically. Within the United States, political disputes over the budget in 2011 created enough uncertainty in NASA's budget that a planned collaboration between NASA and ESA for a 2016 Mars probe was canceled. (Lambright, 2014, pg. 241 – 243) Beyond the possibility of budgets disrupting cooperation, there is also the risk of some international crisis damaging the collaboration. Tensions between the United States and Russia over the Ukraine leaked into space policy when the United States passed high-technology export sanctions on Russia and the Russian Deputy Prime Minister responded by threatening US access to the ISS. (A. Taylor, 2014)

D. The Multiple Model Approach

A clear conclusion of the above discussion is that different organizational models have different advantages and disadvantages. In cases where private capital is insufficient to fund space exploration, a governmental agency or an international collaboration has a clear advantage. In contrast, private organizations reward the efficient allocation of resources. On the other hand, where there is strain on national budgets, international collaboration is attractive. Ultimately, the future of Mars exploration and colonization will require multiple stages of development and multiple forms of organization. This will mean incorporating multiple models for development as they become appropriate and as their unique advantages are maximized. It will also mean devising new private-public partnerships to bring out the best of each model while mitigating the weaknesses of each.

IV. A Model of Institutional Change

Arguing that a Mars mission is technically possible is not the same as saying it is humanly possible. It may be the case that each model for space exploration will be most effective at different stages in the process. We need a model of institutional change in order to examine what the many and varied stages of colonization and economic development will look like. This allows us to predict, using expected exogenous changes in technology and political environments, what the effect will be on colonization. The model also helps us understand why the current stage has its features.

The rest of this thesis is perhaps best considered a speculative exercise in what might be the future of mankind's journey into the greater universe. In succeeding sections I offer a multi-stage approach to Martian exploration and colonization. While I believe that the future will unfold broadly in this pattern, it is not my intention, nor is it within my abilities, to predict the details. This is not to be considered an attempt at Hari Seldon's Psychohistory, but rather a combination and application of existing tools within the social sciences. Specifically, I hope to, using a theory of institutional, change show how exogenous shocks might drive changes in the national/international/private balance of space exploration.

Such a theory is useful in two respects. First it allows us, given a shock, to predict the process by which institutional arrangements and environments might change. That is, if the cost of low earth orbit transportation drops sufficiently to make a new range of enterprises profitable (for example lunar mining of Helium – 3), a theory of institutional change predicts the actors (who will seek change), the form of an emerging institution (individual, voluntary, national, international), and how long this process will take. Second, a theory of institutional change can help us to broadly map out future institutional environments given expected exogenous shocks. Suppose for example, we expect advances in Nuclear Fusion technology that make it an economically viable form of energy production. This exogenous shock would incentivize the mining of Helium-3 on the moon. A model of institutional change would take his expected shock to the institutional environment and produce an expected future institutional state. The model is therefore a forecast of future institutional environments and arrangements. Our model is therefore an exercise in comparative statics.

The basis of our theory can be found in Davis and North (1971) which attempts to create a model of institutional change that can explain episodes in American economic history and also in Ostrom (1991) which seeks to explain why some groups are able to craft lasting institutional solutions to common pool resource problems. Something of a synthesis must take place before we can truly proceed however. Davis and North use very different terminology for their model than Ostrom and so it is necessary to outline both of their models and highlight similarities and differences. From this we will be able to launch our own discussion of a theory of institutional change. I begin by defining some of their terminology below. In the interest of achieving a synthesis between Davis and North and Ostrom, I will occasionally break discussion of one's work to clarify how it relates to the other's work.

A. Definitions

Before we can truly develop a model of institutional change, we must first take some time to clarify what we mean when we speak of institutions along with several other definitions. This section relies heavily on Chapter 1 of Davis and North's *Institutional Change and American Economic Growth* (1971) and also on Ostrom's *Governing the Commons* (1991) and her *Understanding Institutional Diversity* (2005).

1. **The institutional environment** is the set of fundamental political, social, and legal ground rules that establish the basis for production, exchange, and distribution. Davis and North treat the institutional environment as exogenous to their model of economic institutional change.
2. **An institutional arrangement** is an arrangement between actors which governs how they cooperate and/or compete. This arrangement can be formal – a corporation or an organized political party, or it can be informal – a weekly gathering of fishermen to discuss common

interests, but it must serve one of two functions. First it must provide a structure within with members can cooperate/compete for benefits unavailable outside of the arrangement. Second it must provide a mechanism by which the rules governing competition and cooperation can be changed.

These first two definitions comprise two or three levels of institutional analysis according to Ostrom (1991, 2005). In Ostrom's methodology – institutions can be arranged in a hierarchal framework as in Figure 1 below. The institutional environment corresponds roughly with Ostrom's Constitutional-choice situations while the institutional arrangement corresponds with both Ostrom's collective-choice and her operational levels. Constitutional-choice rules (or the institutional environment) are not necessarily exogenous in Ostrom's analysis but instead a deeper (and more costly to change) level of institutional analysis. It is held constant if the analyst is focusing on shallower levels of institutional analysis, but a different analyst may wish to study changes in Constitutional-choice rules. Institutional arrangements can refer to collective-choice rules (Congress is able to pass laws which alter how a corporation can be formed) or to operational situations (the corporation itself).

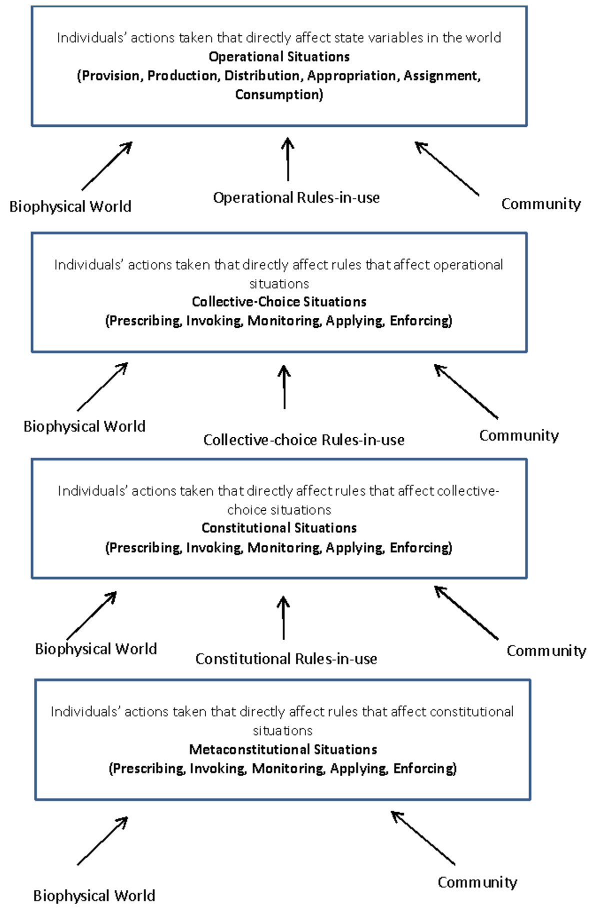

Figure 3 Ostrom's Levels of Analysis and Outcomes (2005, pg. 59)

Using the example of a small fishing community allows us to better illustrate what Ostrom means at each level of the hierarchy. Operational decisions in the fishing community include each fisherman's decision to fish today or not, how many fish to bring in, and what fishing equipment to use. There may be operation rules in place that limit the kinds of decisions that can be made at the operational level. The community may have a quota for how much fish any one fisherman can catch in a day. These rules are set at the collective-choice level. Collective-choice decisions include who is allowed to participate at the operational level (you have to live in the town to fish here) and how operational rules are made (by a vote of everyone who attends the town hall meeting). Of course local authority is often subject to higher levels of rule-making. The constitutional-level is responsible for rules governing how collective-choice decisions can be made. State law might prohibit our fishing town from confiscating fish caught in excess of the quota. How state laws are enacted is in turn partially governed by the US Constitution and federal law acting as a metaconstitutional level in the hierarchy. Article Four of the American constitution requires that each state have a republican form of government. Of course institutions are not fixed within the hierarchy and in another example the state law might be a collective-choice rule and federal law might be a constitutional-choice rule.

3. **The Primary Action Group (PAG)** is a decision-making unit whose decisions govern the process of institutional innovation. The PAG recognizes that there are profits which are currently external to the existing institutional arrangement and so initiates the process of innovation.

4. **The Secondary Action Group (SAG)** is a decision-making group established by a change in the institutional arrangement to help the PAG capture external revenue. New regulatory agencies or government programs can are typically the form which a SAG takes although voluntary arrangements are not unknown (the American Medical Association for instance) – though voluntary SAGs will depend on the coercive power of government in some form in order to be effective in helping the PAG capture income (a government-backed license in the case of the AMA).

Primary action groups are typically at the operational level while the secondary action groups they form are at deeper levels of the hierarchy.

5. **Institutional Instruments** are documents or devices employed by action groups to capture external income when these instruments are applied within a new institutional arrangement structure. Instruments rely on legitimacy and power afforded by a higher level of the hierarchy of institutional analysis. The power to tax is reserved for Congress by the Constitution of the United States – taxation is a collective-choice instrument which derives legitimacy from the collective-choice rules-in-use established by the Constitution. Alternatively, a cease and desist letter is used at the operational level (an operational instrument) to stop a competitor from producing a patented invention. This operational instrument derives legitimacy and power from the operational rules-in-use established by Congress.

B. The Model Developed

Armed with the above definitions we can proceed at outlining a first attempt at a model and theory of institutional change. Succinctly stated the theory is that institutions change because the benefits of change outweigh the costs. Some exogenous shock leads to the emergence of profits[10] which exist outside the existing institutional environment. Once someone recognizes that these profits could be captured through a change in the institutional environment, he will seek to form a Primary Action Group of others who could also profit from the change and who are willing to devote resources to bringing this change about. The recognition of new profits and the formation of a PAG are both subject to lags. In some cases it

[10] Profits in this model should be considered very broadly. As shown in the diagram below, profits can be considered political capital or even the attainment of foreign policy objectives in addition to the traditional dollars that profits usually signify. Political "profits" might include bringing a new demographic group into a party coalition or by achieving some legal change that benefits your party. International agreements can also yield "profits." The Obama administration's recent success in getting China to commit to a climate change agreement might be considered one example. A favorable trade agreement might serve as an example of all three. Such an agreement is likely to increase the international clout of the country able to secure it. The agreement brings political support to the party in power from domestic producers who benefit from it and those domestic producers likely see an increase in profits from the agreement.

may be a long time before the profit potential is even recognized. Where the exogenous change is in technology or in knowledge it may be some time before potential profits are recognized by an entrepreneur. On the other hand, one change can provide encouragement for several subsequent profit sources and successful utilization of a new technology toward one end may encourage its use toward another. Consider the variety of ways the internet has reduced transaction costs and information asymmetry.

Formation of a PAG[11] can also be subject to lags, especially if the group that must be formed is comprised of many people with little prior organizational connections. Forming a PAG will require forging a common identity which previously did not exist. The lag can be greatly reduced if there are existing organizational ties which can be either directed or re-tasked with pursuing the external profit. The process of forming a PAG can be outlined using a form of Ostrom's (1991) model of internal decision-making. In this model individuals decide whether or not to support a change in the rules by discounting the expected net benefits of a rules change. Deciding to support a change in the rules is tantamount to joining the PAG. Internal norms and beliefs about shared norms within the community also play an important role in this process too. These shared norms influences the sets of strategies which will even be considered. For example, if the community and the individual share a strong norm for keeping promises – it may be that strategies of breaking promises are not adopted not because it is more profitable to keep them but because the idea of breaking a promise isn't even considered by the actors. Shared norms can make the formation of a PAG easier by reducing the costs of organization and

[11] The formation of a Primary Action Group is usually seen as a process by which individuals form a group or redirect an existing group, but this need not be the case. Corporations in an industry which together form an industry lobbying group are equally a PAG. Nations often form PAGs. NATO is a Primary Action Group formed to achieve greater defense for its members by institutionally changing how member states cooperate and coordinate militarily.

by reducing free-riding and opportunism. A group seeking to form a PAG which does not have these shared norms is likely to find it harder and more costly to form a group for institutional change.

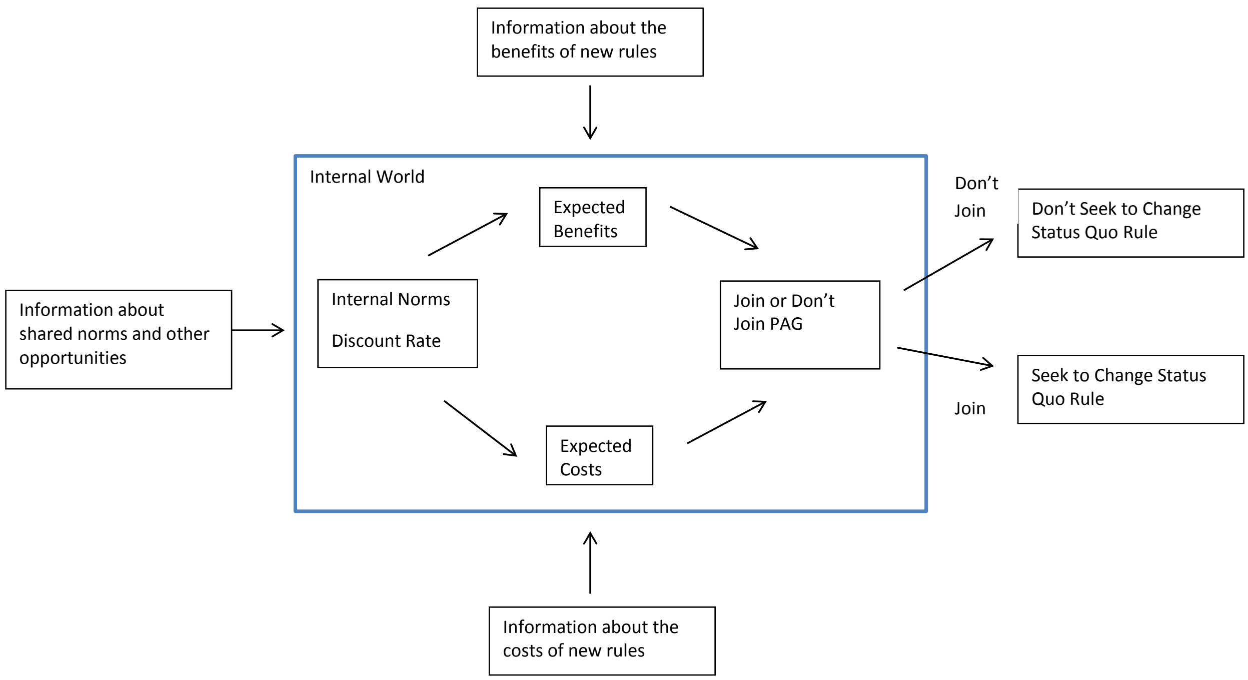

Figure 4 Summary of variables affecting Institutional Choice

Once a PAG has been formed, it falls to them to put forward a concrete plan by which the external profits are to be captured. They do this by scanning the "menu" of available operational and collective-choice forms for one that is suitable to their goals, that is, for at least one option whose present value is greater than zero. The options can be individual, voluntary, national, or international and there are costs and benefits to each[12]. Illegal options are not

[12] Individual arrangements require only self-enforcement and no consensus building. However, an individual arrangement is often not a viable form for larger projects. An individual may form a sole-proprietorship to capture external income but an individual is not likely to pass legislation. A voluntary group has the benefit of greater resources which come from the group but must work to monitor its members' commitments and must work to build consensus. A governmental arrangement has the benefit of being able to coercively enforce its decisions but must build consensus to remain legitimate. International arrangements are similar to voluntary arrangements and

considered – though organization forms illegal in one country might be legal elsewhere. This scanning is done by performing a present value comparison of the different available options. Let R_i be the expected returns and C_i be the operational costs of a new arrangement in year i. In addition to the operational costs of an arrangement there are also organizational costs which must be paid up front. If these organizational costs are high it is possible that an arrangement won't be innovated for this reason alone. Since the expected costs and benefits are spread over n years they must be discounted by rate r. A higher discount rate will make it less likely that small and steady incomes will induce a change. Uncertainty is inversely related to how confident we are in our estimates of the expected costs and returns.

$$PV_1 = \sum_{i=1}^{n} \frac{R_i^1 - C_i^1}{(1+r)^i}$$

$$PV_2 = \sum_{i=1}^{n} \frac{R_i^2 - C_i^2}{(1+r)^i}$$

$$PV_3 = \sum_{i=1}^{n} \frac{R_i^3 - C_i^3}{(1+r)^i}$$

$$\ldots$$

$$PV_m = \sum_{i=1}^{n} \frac{R_i^m - C_i^m}{(1+r)^i}$$

If no alternatives are known it will become necessary for the PAG to innovate a new arrangement. This invention process can be short if analogous arrangements in other industries or countries are known and they can be borrowed or modified for use. This process can also be

must monitor its members' commitments and also build consensus, but it has the benefit of access to the resources of many nations.

made longer if the constitutional and collective-choice environment makes a large number of potential arrangements illegal.

A third lag exists in the menu selection process. If there are a large number of options to pick from, the selection process can take longer – comparative analyses have to be performed. The menu selection lag is shortened by the existence of options which are clearly superior. If one option has a very high present value relative to other arrangement options, then the choice is clear. The smaller the gap between present values in options the longer and harder making a choice will become.

Once an option has been selected, there will be a final startup lag. If the arrangement choice chosen from the menu is an individual arrangement, this startup time can be very short. In addition the more certain and the greater are the expected profits of arrangemental innovation, the greater an incentive there is to devote the resources necessary to actually implementing the arrangement. A voluntary arrangement is likely to face a longer lag if its members don't share equally in the expected profits. This leads to a free rider problem and to delays. The fewer the members of the group the easier and more quickly can this startup occur. A governmental arrangement will be shorter if it can be accomplished within the existing government. If elections have to take place before the startup can occur, this will increase the time and cost of a startup lag. In addition the existence of a large passionate opposition can also increase the lag time. Finally an international arrangement will be similarly shorter if there exists a strong international consensus for the change and if there is no entrenched group opposed.

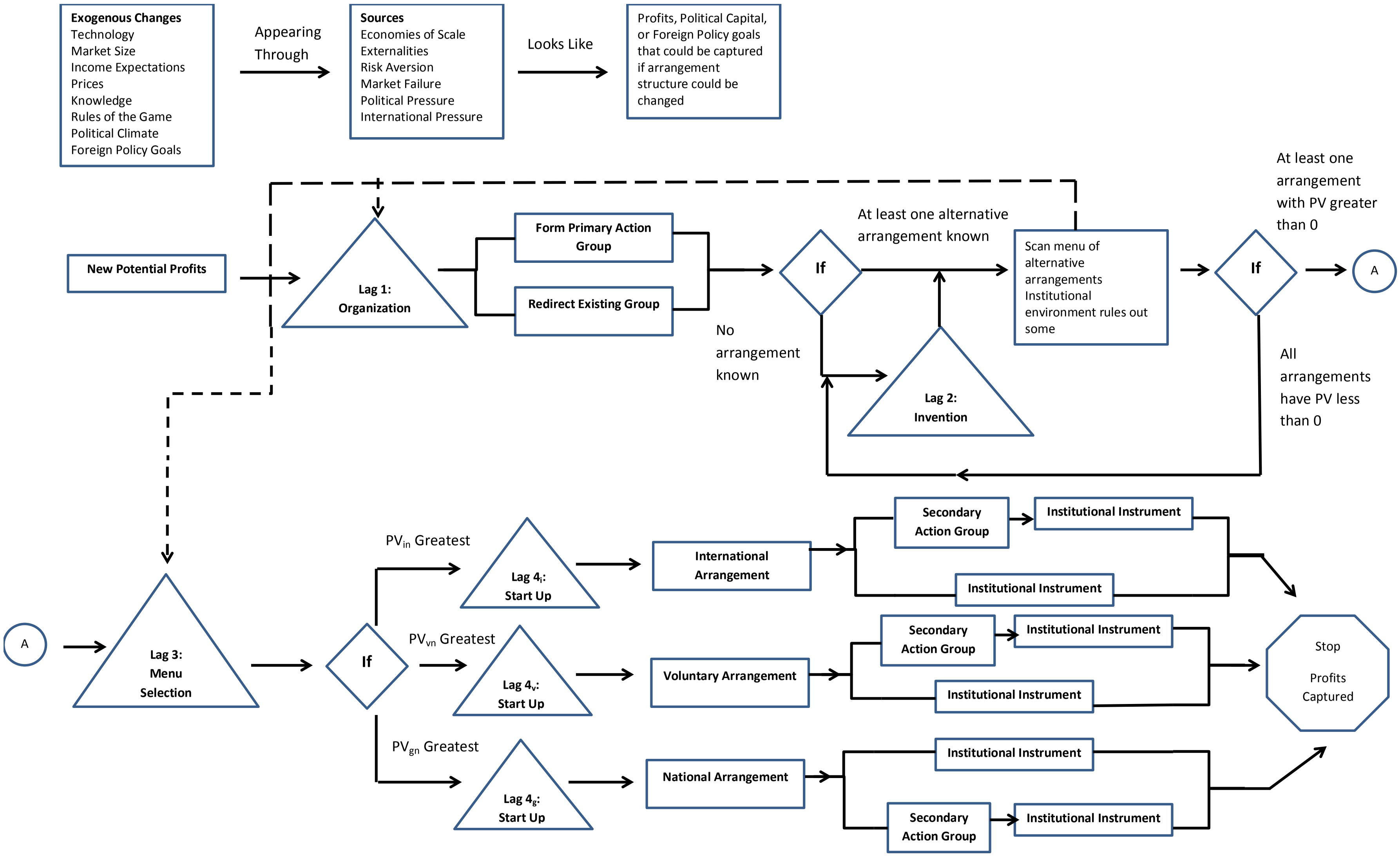
Exogenous Changes
Technology
Market Size
Income Expectations
Prices
Knowledge
Rules of the Game
Political Climate
Foreign Policy Goals
Appearing Through
Sources
Economies of Scale
Externalities
Risk Aversion
Market Failure
Political Pressure
International Pressure
Looks Like
Profits, Political Capital, or Foreign Policy goals that could be captured if arrangement structure could be changed
New Potential Profits
Lag 1: Organization
Form Primary Action Group
Redirect Existing Group
If
At least one alternative arrangement known
No arrangement known
Lag 2: Invention
Scan menu of alternative arrangements Institutional environment rules out some
If
At least one arrangement with PV greater than 0
A
All arrangements have PV less than 0
A
Lag 3: Menu Selection
If
PVin Greatest
PVvn Greatest
PVgn Greatest
Lag 4i: Start Up
Lag 4v: Start Up
Lag 4g: Start Up
International Arrangement
Voluntary Arrangement
National Arrangement
Secondary Action Group
Institutional Instrument
Institutional Instrument
Secondary Action Group
Institutional Instrument
Institutional Instrument
Institutional Instrument
Secondary Action Group
Institutional Instrument
Stop
Profits Captured

Figure 5 Davis and North's Theory of Institutional Change

Now we can apply this model in a few example cases. Let us begin with a set of expected exogenous shocks to the existing institutional environment that includes advances in fusion technology using helium – 3 and a decrease in the cost of orbital transportation. Because the profit potential available from an advance in fusion technology is already known to us and because the scientific literature is publicly available it is likely that there won't be any recognition lag. Organization is also not likely to take very long as there are existing companies which mine radioactive material on Earth. These companies will redirect their efforts toward finding and mining helium – 3 sources here on Earth. Because these companies already exist and have an established corporate form (a voluntary arrangement) it is also likely that the menu selection and startup lags will be short. Once these lags are overcome the external profit is captured and the new helium – 3 industry will form from existing uranium mining companies.

Second, assume that there is a change in the price of reaching low Earth orbit. Assume that since this is a large factor in the cost of space travel that this reduction in ticket price makes lunar mining of helium – 3 commercially viable if property rights can be established on the moon. Again, recognition of this new profit source is not likely to take long since the media coverage that would cover this breakthrough in space travel would be extensive. The organization lag however could be quite long. Uranium mining groups, while no stranger to complex property rights questions, will likely have to form a new group to lobby governments to change the property rights structure of space. Since the costs of such an effort are likely to be expensive, they will need to reach out to other groups who would benefit from a change in the property rights structure. These disparate groups are likely to never have worked together before and will therefore need to devote considerable time to creating collective-choice rules for themselves. The more groups that would benefit from a change in property rights, the harder will organization, but the easier may be lobbying later.

Assuming they are successful in creating some sort of PAG (which is likely to take the form of a voluntary, multinational, special interest group) they must then choose between the two options available on their menu. Their first option is to lobby for a single powerful government to pass a law recognizing property right claims on the moon. This is likely to be cheaper initially, but could prove more expensive later if other nations challenge these property rights domestically. For example: a US-recognized property right might not be effective if the product is sold in China/the European Union/Japan. These countries might seize the product or prohibit its sale within their country leading to costly lawsuits and arbitration.

The second option, operating through an international agreement, is more expensive initially but makes it less likely that the property right will be challenged later. The initial costs can be reduced if the PAG lobbies the United Nations directly instead of working through several governments. Depending on how these costs are structured, either option is possible. While there are significant benefits to solving this problem through the UN, there is also considerable opposition to changing the property rights structure of outer space within the UN. Developing countries are attached to the existing structure of property rights in space. On the other hand, there is existing pressure in the United States already to change property rights in space. However, once an option is chosen, the startup lag could be long. New technologies likely have to be developed for mining in low-gravity and transportation techniques would have to be developed to move sensitive radioactive materials.

This example is not meant to serve as an actual prediction of how mining on the moon will come about, but rather is meant to help illustrate how the model as developed can work. By applying it to a hypothetical example, we have a stronger idea for what its strengths and weaknesses are. In addition we can go forth and confidently apply it to the current institutional environment using exogenous changes we do expect.

V. Putting the Model to Use

While it is impossible to speak with any real certainty about the particulars of the next century of humanity's obsession with Mars, with the help of the above framework, we can begin to lay forth the contours of what our future with Mars will hold. I argue that the path will eventually culminate in human settlement of Mars and furthermore that this path towards human settlement will follow three or four stages. First, a stage involving only robotic missions, a second stage of one or more manned missions, a possible third stage of permanent scientific settlement, and a fourth stage of civilian settlement on the red planet. It is possible that a permanent scientific base on Mars may not be necessary prior to human colonization, but historically this seems the less probable outcome. Consider the International Space Station or the Antarctic Base. Both are the first permanent human settlements in orbit and in the Antarctic, respectively, and both are primarily scientific endeavors. Both are also international scientific endeavors and so the history also seems to suggest that a scientific base on Mars would also be internationally operated. The figure below presents the four stages envisioned in the human settlement of Mars.

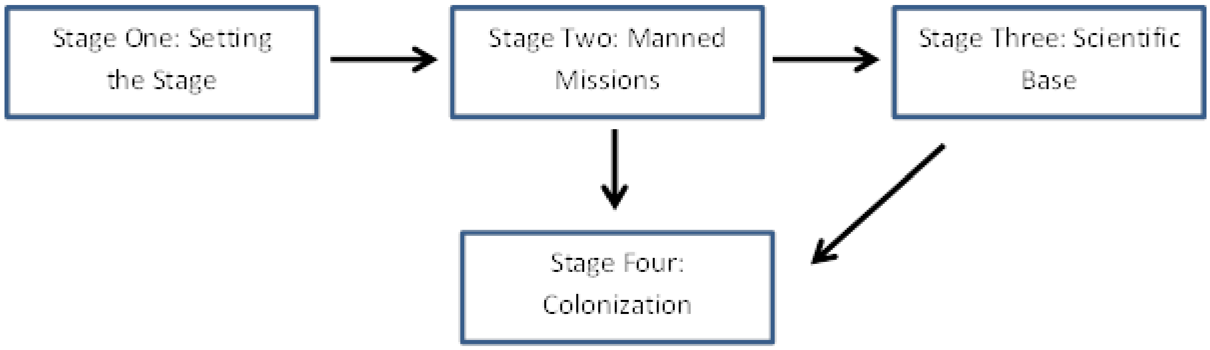

Figure 6 The Stages of Mars Development

I do not present these stages merely as a desired program for human colonization of Mars, but as what I believe to be a necessary progression of events in the journey toward Mars colonization. I argue that these stages will occur in this order not just because it forms a logical progression of what a sustained human presence on Mars should look like, but because there are social, political, and economic forces driving the progression in this direction and through these stages. In the following sections, I use the model of institutional change we have developed to argue that institutions will drive humanity along the path outlined. But in the interest of being upfront and honest with the reader, I begin by examining some of weaknesses inherent in the present analysis.

While this model may be able to say something about the substance of the institutional change involved, it has little to say about the pace of the change. That is, the model is for the most part incapable of offering a timetable for any single stage or even for all the stages combined. While I may be tempted to assert that this process will eat up the better part of the 21rst century, this estimate comes entirely from my gut and a look back at the history of spaceflight. Humans have been able to reach Mars since Mariner 4's November 1968 flyby and the present analysis is unable to say whether the intervening 47 years spent entirely in robotic missions to Mars is a long time for us to be still in the first stage or whether we should expect to remain in this stage for decades to come. Ultimately, the pace at which Mars exploration will proceed is a function of technological progress and of political developments – two things which are exogenous inputs to our model. The problem is somewhat minimized for the earliest stages as we can estimate the timeframe with information exogenous to the model: recent technological breakthroughs or breaking political developments which will either speed up or slow down the pace of Mars exploration.

A partial defense against this criticism lies in the fact that the pace of exploration is in many ways a vacuous concept. It depends on predefined and subjective endpoints. Changing the endpoints changes the pace (somewhat considerably). Consider the Moon landings since this is so often the standard by which other space programs are judged. If we say the Moon Program begins with President Kennedy's 1961 State of the Union, and allow the Moon Program to end with Neil Armstrong's 1969 moon landing, then the entire thing seems to have taken place over about 8 years. But perhaps we really ought to extend the Moon Program to include every Apollo landing. This would extend the Program to 1972 with Apollo 17 and the duration grows to 10 years. But why should we limit this timeline to solely American efforts? The Soviets attempted to land a probe on the moon three times in 1958 (each rocket exploded shortly after takeoff). The timeline roughly stands at 14 years now and it could be extended even earlier (to include von Braun's V2) or later depending on our subjective criteria.

VI. The Path to Mars

A. Stage One: Setting the Stage

> "People will go there. Eventually. They'll settle, they'll terraform – just like you've been dreaming. It might take longer than you were thinking, but you were never going to be one of the ones going anyway, so what's the rush? It'll happen."
>
> \- Kim Stanley Robinson, *Discovering Life*

It now comes time for us to put our model to use. As specified our model should allow us to take a set of expected exogenous shocks to the institutional environment and transform it into a set of expected future institutional states. Let us begin by identifying some of the features of our current institutional environment. Stage one is government dominated. National programs remain the primary agent of space exploration, provide the bulk of the funding, and are the only customer for space activities. Historically, several actors have acted as the Primary Action Group in space exploration, making use of NASA as a secondary action group. These actors include Presidents, the scientific community, outside advocacy groups like the Planetary Society, and NASA Administrators with their own space exploration goals. Traditionally the path of institutional change in space exploration in America has required a primary action group to lobby congress for money through NASA which is then put toward a specific mission. NASA as an existing agency allows primary action groups to save on organization costs and has also served as a means by which government, the scientific community, and private individuals working for NASA can make collective decisions.

There are two major themes of this first stage are institutional changes. The first is the emergence of international collaboration as a viable alternative to national models of space exploration. The second is the development of a new private space industry that will lower the cost of transportation to LEO. These changes can be traced primarily to exogenous shocks of the type we are interested in. The removal of the USSR as a geopolitical rival to the United States removed a key rationale behind large space spending while the recession meant there was less money to spend on many programs. Pressure on NASA budgets and instability in Russia meant that there existed profits (broadly defined) outside the current institutional environment. By the time Clinton was elected President, the *Freedom* space station was in danger of being killed (Lambright, 2014, pg. 117), however instability in Russia meant that cooperation with the new Russian Space Agency on the space station would have a new foreign policy element (demonstrating US support for Boris Yeltsin's government) and a national security element (keeping Russian scientists in possession of dangerous technologies employed). With Clinton onboard a new coalition was formed, saving the space station which was subsequently renamed the International Space Station. In terms of our model - a new international institutional arrangement was innovated by redirecting an existing plan for space station *Freedom* to include Russia. This allowed Clinton to secure of foreign policy victory which was not possible prior to the collapse of the USSR. Moreover, NASA was able to secure for itself the funding needed to complete the space station.

The existence of the ISS has facilitated the emergence of a nascent private space industry, but it was the Great Recession and technological innovations which made such an industry necessary. While there has been for some time commercial demand for launch services[13], NASA and other governmental needs represent a very large market which until

[13] While we do not discuss it, the existing commercial demand for launch services is also thanks to a shock. The advent of the internet and of global positioning technologies led to new and lucrative uses for satellites. This greatly increased demand for launch services and meant that suddenly there would now be room for several new

recently was not open to competition. This market was therefore stagnant and dominated by the joint venture between Lockheed Martin and Boeing, United Launch Alliance. The Great Recession and its attending pressure on budgets forced an already budget-strained NASA to find further cost-savings. This is why we see a conscious effort by NASA to encourage a new space economy by awarding NASA contracts to private firms to meet this need.

In addition this private industry has begun to force itself into the defense market. A great deal of hype surrounds SpaceX and its CEO Elon Musk over his efforts to innovate reusable rockets, but the most important thing to come from SpaceX so far is its legal challenge of the US Air Force's contracting practices. The contract in question, a launch deal with the Boeing/Lockheed Martin joint venture called United Launch Alliance (ULA), is worth $11 billion and is part ULA's monopoly over military satellite launches.

Williamson (1985) argues that transaction costs exist only in circumstances where bounded rationality, opportunism, and asset specificity exist. This certainly describes the present situation. The assets which go into building and launching satellites are incredibly specific. Rockets and rocket scientists have few applications beyond putting objects into orbit. This asset specificity ties together each party of the transaction and makes it very difficult for either to find other options in the marketplace. Williamson's (1985, pg. 53) and the transaction costs approach prediction in cases of this nature is a movement away from market procurement to unified ownership – which in this case would mean the nationalization of ULA. While nationalization has not happened, this was the general thrust of Elon Musk's latest accusations against ULA. In an interview he accused The Air Force of stalling in its certification

launch providers in the market.

of SpaceX so as to avoid offending friends and potential future employers at ULA. (de Selding, 2015) Whether this accusation is true or not is secondary to our purposes. Transaction costs theory suggests that in the bilateral relationship between ULA and the Air Force, we would expect to see higher transaction costs – that is asset specificity makes both parties vulnerable to opportunistic behaviors.

One oddity which exists in the contractual relationship between ULA and the Air Force is the cost-plus pricing system as described by Zubrin (1996, pg. 283 – 284). Prices are set as a fixed markup of whatever the costs end up being. This means that ULA has little incentive to innovate cost-saving measures. Reducing costs will not improve their profitability. The evidence for this stagnancy can be seen in ULA's own actions. The news broke in early 2015 that SpaceX would drop its suit with the Air Force in exchange for the promise of more competitive opportunities in the future (Davenport, 2015). But even before this news was released ULA had replaced its CEO and announced a major restructuring aimed at cutting their costs in half. (Avery, 2014) The timing of this decision seems hardly coincidental.

We now show that if SpaceX is successful in breaking into the defense industry (as currently seems likely) this will actually lower the costs of transportation to LEO. A Bertrand duopoly with product differentiation seems the most appropriate model for this situation as it clear that price per launch is the strategic variable. The number of launches needed is determined exogenously to the competing firms. The Air Force decides it needs a number of launches each year and asks companies to quote a price given that number of launches. In addition, ULA goes to great lengths to emphasize that in addition to price, their customers should consider their success rate when awarding contracts. To show how this model indicates that innovation and cost-cutting is incentivized please begin by assuming that SpaceX and ULA are the only two firms competing over defense contracts and that they face symmetric demand curves.

ULA: $Q_{ULA}(p_{ULA}, p_{SX}) = -p_{ULA} + 0.3p_{SX} + 10$

SpaceX: $Q_{SX}(p_{ULA}, p_{SX}) = -p_{SX} + 0.3p_{ULA} + 10$

Assume that SpaceX's boasts are accurate, that it can provide launch services for much less than ULA can – in our model this is tantamount to arguing that SpaceX has a lower marginal cost of launching satellites than ULA. Let SpaceX have marginal costs of 10 and let ULA have marginal costs of 13 (that is SpaceX is able to launch additional rockets at roughly 76% the cost of ULA). It is a simple matter to show that if SpaceX and ULA are allowed to compete SpaceX will dominate the market unless ULA is able to cut costs. We begin by finding our initial industry equilibrium. SpaceX's reaction function, which is a decision rule that allows SpaceX to maximize its profit given ULA's decisions, is found by taking the first partial derivative of π_{SX} with respect to p_{SX}, setting it equal to zero and then solving for p_{SX}. Since the firms' demand functions are roughly symmetric we're thus able to quickly establish the reaction functions for each firm.

$$\pi_{SX} = TR - TC = p_{SX} * Q(p_{ULA}, p_{SX}) - mc_{SX} * Q(p_{ULA}, p_{SX})$$

$$\frac{\partial \pi_{SX}}{\partial p_{SX}} = -2p_{SX} + 0.2p_{ULA} + 20$$

$$p_{SX} = \frac{0.3p_{ULA} + 20}{2}$$

$$p_{ULA} = \frac{0.3p_{SX} + 23}{2}$$

We can now determine how much each firm will produce and at what price by substituting in each other's best response function and solving for the relevant price. Thus SpaceX will launch at a price of roughly $11.99 with profits of $3.979 while ULA will charge roughly $13.30 and barely cover their marginal costs of launching. Under this competitive pressure United Launch Alliance will therefore need to either cut its costs or exit the market. If ULA is able to adopt the cost-saving practices that SpaceX has innovated, then our best response functions will be as follows.

$$p_{SX} = \frac{0.3p_{ULA} + 20}{2}$$

$$p_{ULA} = \frac{0.3p_{SX} + 20}{2}$$

The results are two-fold. Not only is ULA able to stay in business by charging the much lower price of $11.76, but SpaceX is also forced to match ULA's new lower price. Why would ULA do this though? They were making an economic profit before, albeit a very small one. Cutting costs and lowering prices actually allows ULA to take in an even greater profit. Lowering marginal costs by $3 allows ULA to increase its profits by over 3,371%. Obviously the above is a contrived example. The conclusions are, however, quite robust to changes in the coefficients used. Alter the marginal costs or cross-price coefficients in SpaceX's favor and you will only increase the incentives which ULA faces to cut costs competitively. Lowering ULA's marginal costs relative to SpaceX's will only make the initial industry equilibrium more competitive and granting ULA lower marginal costs than SpaceX only shifts the competitive incentive from ULA to SpaceX.

In broad terms the changes seen over the past twenty years conform to what our model would predict. Prior to the collapse of the Soviet Union, international collaboration in outer space was quite limited and efforts at a US led space station *Freedom* were floundering. While there were certainly benefits to collaborating with the Soviet Union on a space station[14] – it was not possible in the existing Cold War environment. The collapse of the Soviet Union brought about a dramatic change in the international political environment. The political rationale behind *Freedom* as a symbol of Western success that could be compared to the Soviet *Mir* was gone. Moreover the over budget and behind schedule project looked like an easy target for those seeking to cut deficits. However, demonstrating US support for the new Russian government was an important goal of the Clinton administration and so *Freedom* was redirected to include Russia as a larger and ultimately successful international project. With an existing example of how manned international collaboration in space could work, future projects will be less costly to negotiate and more credible to suggest and advocate.

The emerging private industry is also explained by our model. Technological advances increased demand for satellite launching services while an economic recession put pressure on NASA to make additional cuts or sacrifice mission plans. Because the recession was global and due to the regular nature of ISS resupply mission needs, innovating a private industry capable of meeting this need seemed to be the most attractive option. NASA therefore devoted what money it did have toward encouraging a new private industry through the Commercial Crew and Cargo Program. Now we see this nascent private launching industry seeking to open up

[14] Beyond the cost savings benefits which adding another partner would have brought, the Soviet Union was the only other major space power at the time and in addition had a great deal of experience in space stations from their own *Mir* space station. This experience would prove very valuable to the ISS.

new sources of revenue by seeking further institutional innovation: namely the opening up of defense contracts to competitive bidding. Ultimately the result we can expect of these innovations is that the cost of transportation to LEO will decrease dramatically.

B. Stage Two: Manned Missions

"Houston, Tranquility Base here. The Eagle has landed."

- Astronaut Neil Armstrong upon landing on the moon

Now we will begin to apply our model to what might lie next in Mars exploration. It may help us to divide this stage into primarily two different realms: what happens in private industry and what happens on the international stage. Beginning with the private realm, if private industry is able to deliver the cost reductions we expect from stage one, then a wider range of space activities will now become cost viable. This includes tourism to a privately run space station and the mining of near-Earth objects. Elon Musk hopes to be able to reduce transportation costs by a factor of 100 if he succeeds in designing a reusable rocket. According to the SpaceX website, the Falcon Heavy will carry 117,000 lbs to LEO and will be priced at $90 million. Decreasing the cost by a factor of 100 will mean that transportation to LEO would cost roughly $7.69/lb. At that cost, a few days' vacation on a space station would be very expensive, but not out of reach for the upper middle class in the developed world.

This attractive pricing will also encourage government agencies to outsource many of their LEO needs to private companies. Those governmental space agencies which do outsource these missions to private industry will be able to devote time and resources to exploring beyond Earth orbit. However, because escaping Earth orbit requires a rocket that is a great deal more powerful than that necessary to achieve LEO and because there is no real commercial demand at this point for activities beyond LEO there is no reason to expect that private industry will have moved beyond the confines of the Earth's orbit.

While the arrival of vacationers and of private infrastructure to LEO would be cost viable given the reduction in transportation costs, there will remain a great deal of uncertainty which may hinder private development. In Article VII of the Outer Space Treaty of 1967, nations are internationally liable for any damages caused by objects launched within their territory. Simply put, this suggest that if a privately owned object were launched from the territorial confines of the United States and it then damaged something else in orbit, the United States would be held liable. This kind of framework was sufficient when nations were the sole actors in outer space, but nations would be reluctant to expose themselves to increased liability for the actions of non-governmental actors.

Moreover, Article I of the same treaty requires that the "use of outer space, including the Moon and other celestial bodies, shall be carried out for the benefit and in the interests of all countries." While it is likely unenforceable, this provision has been interpreted to either be a guideline that nations appropriating resources from outer space should share them internationally or as a legal imperative that these resources be shared. (Coffey, 2009) Regardless of how it should actually be interpreted, the presence of non-state actors in outer space will provide incentives for both nations and for private actors to seek a change to the international legal regime for outer space. Nations will recognize that they could benefit from reducing their liability and private actors will recognize that they can benefit from a clarification of their rights and responsibilities. The drive to make changes will be balanced by the non-spacefaring nations' desire to maintain the status quo out of fear that a change to the international regime will leave exclude them from meaningful participation.

This conflict could play out one of two ways. A new regime could be negotiated through the United Nations, or a new regime could be negotiated among the space powers exclusively. Because of the costs associated with negotiating an international regime, using the UN is an attractive prospect. But any agreement made through the UN will require the consent of the non-spacefaring nations. The specifics of such a compromise are impossible to determine, but it will likely require some kind of transfer payment from those profiting from outer space activities to those incapable of exploiting outer space. This new legal regime is likely to extend to all areas of outer space much as the current regime makes little distinction between orbit, space, and celestial bodies. The primary reason for this is that physical demarcations (which would form the basis of a legal distinction) separating areas of outer space are difficult, if not impossible, to determine. This change in the international space regime will allow for private enterprise to truly expand into LEO.

Meanwhile, on the international stage, exploration beyond LEO is likely to remain similar to that seen presently. The form which this exploration will take is, however, not entirely clear. On the unmanned front, continued budget pressures might make greater international collaboration more attractive, but the kinds of missions carried out are unlikely to change. The real ambiguity lies in the future of manned exploration. Barring any major changes, America may continue on its plans to land astronauts on an asteroid. This may be done solely by NASA, but an international model might also be utilized. Even with sufficient NASA budgets, the experience the ESA has gained through its Rosetta mission may make collaboration on a manned asteroid mission attractive. Alternatively, NASA or some USA-led international group may target Mars as the next logical step. Any international asteroid mission is however unlikely to include either China or Russia unless there are significant changes in the international political environment in the coming decades. Regardless, without new incentives to push the boundaries of human space exploration, the American space program will not likely go far. Cheaper transportation to LEO may make such missions cost-effective, but nothing thus far has made going to Mars politically attractive.

Tensions between the United States and its traditional rival Russia have reemerged on the international stage, while angst over the rise of Chinese economic and military might has created new tensions. These are unlikely to dissipate. It is far more likely that they will increase in intensity. China especially will continue to seek opportunities to flex and demonstrate its newfound muscles, and their interest in outer space is undeniable. Russia will continue to try to find opportunities to reclaim its national prestige (barring any major domestic political reforms). In seeking these opportunities both nations are sure to conflict with American interests. Demonstrating power almost requires that they do. In addition, being the first to make a major advancement in space exploration carries significant advantages over being second. The first to make it gains all the glory while subsequent successes by other nations achieve only heavily reduced gains. This means that China and Russia will have less to gain through international collaboration since going alone demonstrates their ingenuity and power far better than cooperation would (especially any cooperation with the US). Thus we expect one of these nations to be the most active in space at the beginning of Stage Two.

If either Russia or China were to make a significant space-related achievement, this would reignite the space race. Neither nation has the technological (or indeed economic) capability of landing on Mars. Russian probes to Mars have a very high failure rate while China has yet to independently send anything to Mars. But either, given sufficient resources and political will, could credibly seek to land men on the moon. A Chinese moon landing seems far more likely, but in the end it may not matter which nation actually does the landing. The effect in the US would be quite similar to that of the Soviet launch of *Sputnik*. Immediately, landing Americans on Mars would become a national necessity.

While our model does seem to predict that during this stage a new space race will erupt between the United States and some geopolitical rival (most likely China), several things remain uncertain. First, what kind of achievement by a rival would be necessary to trigger this second space race? The above discussion suggests a moon landing primarily because this is consistent with the stated goals of Chinese space policy and because a moon landing would likely carry a stronger message than their intermediate goal of building a space station. The American presence in LEO is unlikely to end anytime soon and so adding a Chinese space station does not seem terribly shocking to American pride. On the other hand, the United States has not landed on the moon since 1972. Nevertheless, we have little idea of what public reactions to tomorrow's headlines will be.

The second uncertainty lies in who would actually win this space race. That answer depends primarily on future technological differences between the participating nations. If technology is allowed to diffuse freely between these nations, we would expect more or less equality in their chances. Given the tension we expect in this period however, freely flowing technology might seem less likely, however powerful communications technologies promise to make knowledge far more available globally and with fewer restrictions. Unless the international community is able to establish a regime which effectively counters the growing cultural and technological trend to make all information freely available, it is possible that even nations which consider themselves enemies will not be capable of preventing the passage of information from one to the other. Having private access to LEO also reduces some of the technical challenges of reaching Mars. A nation could contract with some private company to bring supplies and personnel into LEO and also to bring returning astronauts to the Earth. The nation would then only be responsible for getting from LEO to Mars and back. Ultimately however, Stage Two ends with the first landing of humans on Mars.

C. Stage Three: Permanent Scientific Base

"It's a fixer-upper of a planet, but we could make it work."

- Elon Musk, *Interview with CBS*

With the arrival of humans on Mars, the second space race will end. This may not be immediately apparent as subsequent missions to Mars may still occur – especially if the losing nations were also close to achieving a landing. But these subsequent trips will likely be made at a loss. The national gains to be made from journeying to Mars will immediately drop. Without the title of "First to Mars" succeeding missions can at best demonstrate a technological parity among the competing nations. If there were a clear next target for space exploration, the losers of the second space race might refocus their efforts and aim to "win the next round."

This would be similar to what happened following the Soviet launch of *Sputnik I.* The United States lost the race to be the first to launch a manmade satellite into orbit, but there were clear targets which lay beyond that feat. Ultimately the United States came to dominate the space race in lunar achievements. It is, however, not clear where humanity should go after Mars. As inhospitable as Mars seems to us now, there is no place in this solar system which would be more supportive of life outside of Earth. Without a clear next target for national competition, the game has to end.

From here the future of mankind on Mars could go one of two ways. This section assumes that a third stage of scientific bases will be necessary prior to the beginning of sustained colonization. The next section considers the possibility that this third stage could be skipped and colonization could begin following the first manned mission to Mars. Private actors are the ones most likely to be the driving force behind the colonization of Mars. Without the nationalistic fervor of a space race, profit is the force most likely to encourage the kind of investment necessary to make a sustained human presence possible on Mars. Therefore the actual course of the future will depend on the state of private industry in space and on our knowledge of Mars itself.

At this point in outer space development, we expect private industry to have expanded into LEO through tourism and the mining of asteroids near Earth. Remember this was made possible through a change in the international space regime to clarify the rights and responsibilities of those acting in space and through a dramatic decrease in the cost of transportation to LEO. We do not expect private industry to have expanded much beyond LEO at this point primarily due to uncertainty over the potential profits to be gained thereby. Assuming that uncertainty does prevent private expansion beyond LEO, then the private sector would be in a weak position to drive the colonization of Mars.

This does not mean that the colonization of Mars is a doomed endeavor. But it will increase the length of time necessary for colonization by a great deal. While the second space race may be over, there will still be political benefits to be had on Mars. National investment in reaching Mars as part of the space race will have meant new spending that flows into districts around the country. As we argued above, once that spending occurs there are strong political incentives to keep the money flowing. Congressmen and industry will continue to lobby for federal spending on the kinds of programs which began during the space race. This lobbying will be at least partially supported by groups like The Planetary Society, and the space science community. Thus the allocative inefficiency of government spending combined with domestic support for space exploration will preserve at least some Mars spending. But without a national security or geopolitical rationale to add to this lobbying, there will once again be pressure on space science budgets.

Those who have a vested interest in seeing further exploration of Mars will therefore see benefits to advocating international models of Mars exploration. This prospect will be especially attractive if it can be adopted as a means of achieving some wider foreign policy goal like détente between two rival space powers. These kinds of projects may begin bilaterally as limited missions between former space rivals, but as other countries continue to develop outer space technologies larger collaborative efforts will become possible. Eventually such efforts will yield a permanent scientific presence on Mars. Thanks to our experience aboard the ISS, negotiating such a multinational and long-term mission will be easier. The parties involved will have a better understanding of what kinds of problems to expect in the building and maintenance of a permanent scientific base.

D. Stage Four: Colonization

> The men of Earth came to Mars. They came because they were afraid or unafraid, because they were happy or unhappy, because they felt like Pilgrims or did not feel like Pilgrims. There was a reason for each man. They were leaving bad wives or bad jobs or bad towns; they were coming to find something or leave something or get something, to dig up something or bury something or leave something alone. They were coming with small dreams or large dreams or none at all.
>
> - Ray Bradbury, *The Martian Chronicles*

With a permanent scientific base on Mars, private industry will be able to expand from LEO to Mars much as it has expanded into LEO in our own time. The scientific base will need regular resupply which could be provided by private industry at lower cost than by a national or international space agency. This would create a stable revenue stream for private infrastructure to be built on. On the other hand, it is possible that the changes to the international space regime we expect in the second stage might have been sufficiently broad to dispel some of the uncertainty surrounding profits that could be made on and around Mars. In addition, experience in mining near-Earth asteroids might be easily translated to asteroids near Mars. If this is the case then private investment in Mars development may already be forthcoming.

But in what projects would private capital be invested? If there is a permanent scientific base already established on Mars, one attractive investment would be projects which can reduce the cost of transportation to and from Mars. A system of cyclers seems to be the most viable means of reducing Mars-Earth transportation costs. Proposed by the astronaut Buzz Aldrin, a cycler uses gravitational forces and orbital mechanics to maintain an orbit around the sun which puts it on a continual trip between Earth and Mars. Increasing the number of cyclers increases the frequency of available “flights” to and from Mars and also allows for regular trade in people and goods between the planets. The primary advantage of using a cycler network is that once placed in its orbit, a cycler will only need to make minor course adjustments to stay in that orbit. This allows you to conduct several trips with one launch. Taxis would complete transfers to and from the passing cyclers. (Oberg and Aldrin, 2000)

If there is no scientific base to rely on, investment would most likely flow to the main asteroid belt. Without a stable revenue stream from resupply missions, mining asteroids in the main belt between Mars and Jupiter would be the most tempting commercial option. Using data primarily from NASA JPL and from the Minor Planet Center, Asterank.com attempts to evaluate how profitable it would be to mine over 600,000 catalogued asteroids. There are a number of problems with Asterank’s calculations. Perhaps the most significant flaw with their calculations is that they use the current market value of the metals being mined. If precious metals from asteroids were to flood into Earth markets, the current market value would certainly drop – making the entire venture less profitable. Nevertheless, our understanding of the composition of these asteroids will continue to improve allowing for more accurate prospecting. It is not at all hard to believe that eventually science, experience, and sophisticated financial models will allow us to remotely estimate the composition and value of an asteroid.

Once such evaluations of asteroids exist, it will be a simple matter of finding the right asteroids. Using the present value function defined above ($PV = \sum_{i=1}^{n} \frac{R_i - C_i}{(1+r)^i}$) firms will be now be armed with all the information they need to make efficient business decisions. Asteroids will be mined if the expected present value of mining that asteroid is greater than zero. The long distances involved mean that any revenues are likely to be far in the future while any costs are largest up-front. Colonizing Mars would then evolve out of a need to solve these two problems. Having a base on Mars would allow firms to reduce costs a great deal and could also reduce the amount of time between costs and revenues. It is less costly to base an asteroid mining operation on Mars than on Earth first because Mars is closer to the asteroid belt and second because the cost of launching from the surface of Mars to an asteroid would be significantly less than that of launching from Earth to the same asteroid. Zubrin (1995) argues that Mars "has a 7-fold advantage in mass ratio over Earth as a port of departure for the Main Asteroid Belt." In addition by having a base on Mars would encourage the development of a Martian commodity market, especially if there was a cycler network in operation. This Martian market could buy what is mined for export to Earth. This would allow for smaller scale mining since the exporter will be able to pool the fruits of several mining operations for more efficient transportation.

A Martian base will require people to staff it. While at first a Martian base might be staffed rotationally, eventually advances in technology and pressures to cut costs will encourage people to remain on Mars as dedicated employees. These Martians will create demand on Mars for all the goods and services which Earth can provide and with a dedicated cycler network connecting Earth and Mars the flow of goods and of people will be possible.

Mars will suffer sever labor shortages, especially for unskilled labor. This will result in higher wages than could be found on Earth. High wages will attract people, and having more people will mean greater demand for goods and services on Mars. The realization of those demands will increase the quality of life for those on Mars and this increase quality of life will create further incentives for people to move from Earth to Mars.

E. What Lies Beyond

"I'm coming back in… and it's the saddest moment of my life."

- Astronaut Ed White at the end of the first American space walk

From where we stopped in Stage Four there are several directions we could take. We could attempt to determine what political institutions might form out of a Mars base. We could seek to answer what economic forces might drive the terraforming of Mars. Or we could even try to examine the broader relationship between Earth and Mars culturally, politically, and economically. I leave these questions to future researchers. Even a speculative exercise as fantastic as the present one must bind itself temporally. The further out we attempt to gaze, the more uncertain our vision becomes. Eventually this uncertainty becomes so great that it becomes difficult to compare the likelihood of too possible outcomes.

We have already begun to hit this wall in the previous two sections. While it may seem likeliest that an international Martian science station is an important and necessary step toward colonizing Mars, we could not neglect the possibility that it would be unnecessary if private capital was forthcoming. Thus we had to examine two branches of outcomes. Because future institutional environments are dependent, at least in part, on prior institutional environments, those branches have a tendency to multiply once they emerge. If these branches grow geometrically then the two branches we have now would grow to four in the next stage and eight in the stage after that. Best to stop before things get out of hand.

VII. References

Avery, Greg. (2014, October 16). Exclusive: ULA plans new rocket, restructuring to cut launch costs in half. *Denver Business Journal*. Retrieved April 8, 2015, from http://www.bizjournals.com/denver/blog/boosters_bits/2014/10/exclusive-ula-plans-a-new-rocket-restructuring-to.html?page=all.

Garber, S. (Ed.). (n.d). Biographical and Other Personnel Information. Retrieved April 8, 2015 from http://www.hq.nasa.gov/office/pao/History/prsnnl.htm.

Broniatowski, D. A., Faith, G.R., and Sabathier, V.G. (2006). *The Case for Managed International Cooperation in Space Exploration*. Publication. Center for Strategic and International Studies. Retrieved April 8, 2015 from, http://csis.org/files/media/csis/pubs/060918_managed_international_cooperation.pdf.

Coffey, Sarah (2009). Establishing a Legal Framework for Property Rights to Natural Resources in Outer Space. *Case Western Reserve Journal of International Law* 41:119. Retrieved April 8, 2015 from, http://heinonline.org/HOL/LandingPage?handle=hein.journals/cwrint41&div=9&id=&page=

Coulter, Dauna, and Phillips, Tony. (Ed.). (2012). Strange but True: Curiosity's Sky Crane - NASA Science. *NASA Science News*. National Aeronautics and Space Administration. Retrieved

April 8, 2015 from, http://science.nasa.gov/science-news/science-at-nasa/2012/30jul_skycrane/.

Davenport, Christian. (2015, January, 23) Elon Musk’s SpaceX Settles Lawsuit against Air Force. *The Washington Post* Retrieved April 8, 2015 from http://www.washingtonpost.com/business/economy/elon-musks-spacex-to-drop-lawsuit-against-air-force/2015/01/23/c5e8ff80-a34c-11e4-9f89-561284a573f8_story.html.

Davis, Lance E., and Douglass C. North. (1971). *Institutional Change and American Economic Growth*. Cambridge, England: Cambridge University Press. Print.

Do, S., Ho, K., Schreiner, S., Owens, A., and de Weck, O. (2014, October) An Independent Assessment of the Technical Feasibility of the Mars One Mission Plan. *65th International Astronautical Congress, Toronto, Canada*. Retrieved April 8, 2015 from, http://web.mit.edu/sydneydo/Public/Mars%20One%20Feasibility%20Analysis%20IAC14.pdf.

Gifford, S.E. (2014, February 18) Calculated Risks: How Radiation Rules Manned Mars Exploration. *Space.com*., Retrieved April 8, 2015 from, http://www.space.com/24731-mars-radiation-curiosity-rover.html.

Knight, Frank. (1921) *Risk, Uncertainty, and Profit.* Boston and New York: Houghton Mifflin Company. Print

Lambright, W. Henry. (2014) *Why Mars: NASA and the Politics of Space Exploration.* Baltimore: John Hopkins University Press. Print.

Malik, Tariq. (2011, July 21). NASA's Space Shuttle By the Numbers: 30 Years of a Spaceflight Icon. *Space.com*. Retrieved April 8, 2015 from, http://www.space.com/12376-nasa-space-shuttle-program-facts-statistics.html.

Oberg, James, and Buzz Aldrin. (2000, March 1). A Bus Between the Planets. *Scientific American.* Retrieved April 12, 2015 from, http://www.scientificamerican.com/article/a-bus-between-the-planets/.

Ostrom, Elinor. (1991) *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge, England: Cambridge University Press. Print.

Ostrom, Elinor. (2005) *Understanding Institutional Diversity*. Princeton: Princeton University Press. Print.

Overbye, Dennis. *Looking to Mars to Help Understand Changing Climates.* The New York Times. 8 Dec. 2014. Web. 08 Apr. 2015. http://www.nytimes.com/2014/12/09/science/looking-to-mars-to-help-understand-changing-climates.html?_r=0

Posner, Richard. (2004) *Catastrophe*: *Risk and Response.* Oxford, England: Oxford University Press, Print.

President's Science Advisory Committee. (n.d.) Introduction to Outer Space. *NASA Historical Reference Collection*, Retrieved April 8, 2015 from, http://history.nasa.gov/sputnik/16.html.

Ross-Nazzal, Jennifer. *Détente on Earth and in Space: The Apollo-Soyuz Test Project.* OAH Magazine of History. July. 2010. Web. http://web.b.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=a7f3c257-7b15-48c4-8aa9-43e5fcafbdda%40sessionmgr198&vid=1&hid=116.

Sterner, Eric. (2011, February 9). Sterner on the Role of Prestige in US Space Policy. *George C Marshall Institute*. Retrieved April 8, 2015 from, http://marshall.org/space-policy/sterner-on-the-role-of-prestige-in-us-space-policy/.

Stooke, Phillip J. (2012) *The International Atlas of Mars Exploration.* Cambridge, England: Cambridge University Press. Print.

Taylor, Adam. (2014, April 30). Russia's Deputy PM tells U.S. astronauts to go to space on a trampoline; the joke may be on him. *The Washington Post*. Retrieved April 8, 2015 from, http://www.washingtonpost.com/blogs/worldviews/wp/2014/04/30/russias-deputy-pm-tells-u-s-astronauts-to-go-to-space-on-a-trampoline-the-joke-may-be-on-him/.

Taylor, Fredric W. (2010). *The Scientific Exploration of Mars*. Cambridge, England: Cambridge University Press. Print.

The Planetary Society. (2015, March 9). New Details Released for Citizen-Funded LightSail Spacecraft Launch. Retrieved April 8, 2015 from, http://www.planetary.org/press-room/releases/2015/new-details-released-for.html.

Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies. New York, January 27, 1967. *United Nations, Treaty Series, vol. 610, No. 8843.* Retrieved April 8, 2015 from, http://www.unoosa.org/pdf/publications/ST_SPACE_061Rev01E.pdf.

Wackerly, D., Mendenhall, W, & Scheaffer, R. (2008). *Mathematical Statistics with Applications*. (7th ed.). Belmont: Brooks/Cole. Print.

Wall, Mike. (2013, December 9). Radiation on Mars Manageable for Manned Mission, Curiosity Rover Reveals. *Space.com*. Retrieved April 8, 2015 from, http://www.space.com/23875-mars-radiation-life-manned-mission.html.

Williamson, Oliver. (1985) *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York: Free Press. Print.

Zubrin, Robert. (1995). The Economic Viability of Mars Colonization. *Journal of the British Interplanetary Society* 48: 407. Retrieved April 11, 2015 from, http://www.4frontierscorp.com/dev/assets/Economic%20Viability%20of%20Mars%20Colonization.pdf.

Zubrin, R., and Wagner, R. (1996) *The Case for Mars: The Plan to Settle the Red Planet and Why We Must*. New York: Simon and Schuster. Print.

VIII. Appendix: Risk and Uncertainty

There is a difference between risk and uncertainty - risk is quantifiable and therefore insurable. This distinction was noted most famously by Knight (1921) and this section is based largely on Chapter VIII of his *Risk, Profit, and Uncertainty.* We know that the probability that a flipped coin will be heads is 50% or that the probability of rolling a 2 on a fair six-sided dice is 1 in 6. We have, either through study or experience a priori knowledge about the probability distribution of an event. Risk can be insured against because the probability distribution of an event is known. Suppose there was a known 10% chance that any rocket launch would encounter some irrecoverable disaster (it might explode shortly launch or its heat shield might fail and burn up in reentry). Because we can say with certainty that there is a 10% chance of this occurring, the event can be insured against its occurrence. This insurance can take many forms, but the most obvious form it can take would be through a third party insurer. We could pay premiums to an insurance firm and in the event of a disaster the insurance firm would pay out an amount according to an insurance arrangement worked out beforehand. Less obviously (but perhaps more frequently practiced), there is always the possibility of self-insurance. We could save a certain amount every month in "self-premiums" to cover an eventual disaster, or simply accept the risk that 1 in 10 launches will meet with disaster.

Uncertainty on the other hand is not quantifiable and therefore is not insurable. The probability distribution is not known. An easily understood example of uncertainty lies in starting a small business. There is a chance that it will take off and be remarkably profitable, there is also a chance that it will go bankrupt, and finally there is a chance that it will end up being somewhere in the middle. There exists a probability distribution which describes how likely it is the new business will be successful - but the actual properties of this probability distribution remains hidden and uncertain. It cannot be said with any real certainty that there is a 5% chance the new business will make money in the first year, nor can it be said with any certainty that there's a 13% chance of the business losing money in that year.

Uncertainty is for this reason a greater challenge to commercial space exploration (and enterprise in general) than risk. To further illustrate the difference between risk and uncertainty, consider the following two questions.

1. What is the probability that out of three hundred shuttle launches four will fail?

2. What is the probability that an asteroid will be comprised of precious metals in great enough quantity to be mined profitably?

The first question can be answered using prior experience to estimate a probability distribution. Since a launch is either successful or not and assuming that the events are independent we can model the question using a binomial distribution. There have been 135 Space Shuttle Launches, 2 of which ended in disaster (Malik, 2011). This gives us $\hat{p} = \frac{2}{135} \cong 0.0148$ and using the binomial distribution for three hundred launches given below we can estimate that the probability that four out of three hundred Space Shuttle Launches will fail is approximately 19%. In addition to being able to point to a specific number as the answer to our question, we can say that we're confident in our answer because it corresponds to real Space Shuttle experiences. We are relatively certain about our estimate of the risk involved.

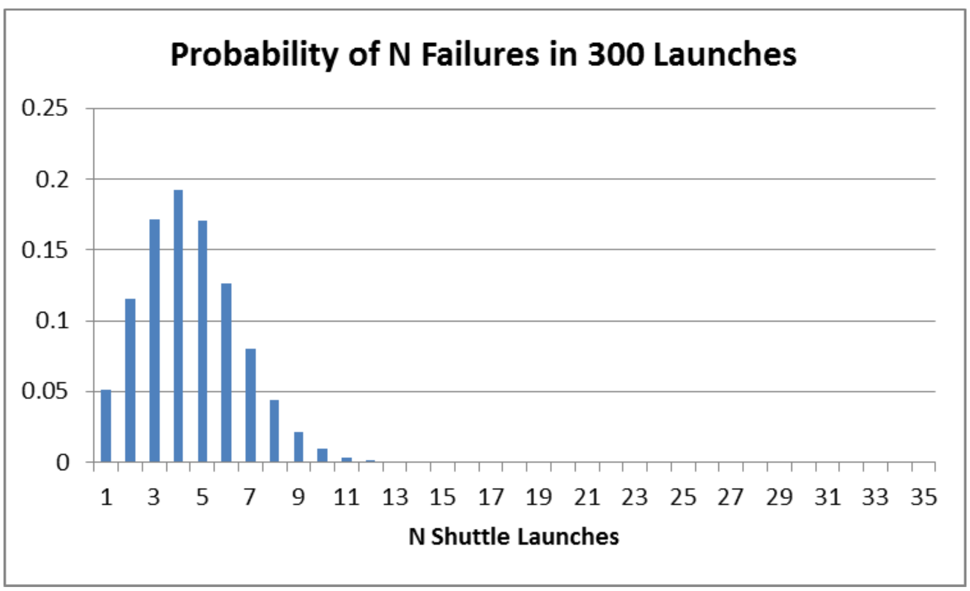

Figure 7 Shuttle Launch Risk

Answering the second question can be done in a similar way, or at least it could be if we knew reliably what the relative frequency of various classes of asteroids in a given astronomical space is. The uncertainty here revolves around both difficulties in determining asteroid composition and also difficulties in discovering asteroids. In fact there is good reason to believe that our understanding of asteroids is biased toward certain classes which have compositions that make them bright enough to be discovered. Transforming this uncertainty into risk requires scientific study of asteroids or experience mining asteroids. By expanding our knowledge of asteroids we may be able to state with greater certainty the relative frequency of different classes of asteroids are.

Transforming uncertainty into risk is itself a moneymaking enterprise. Financial markets are constantly on the lookout for new ways to make this transformation. For a long time mortgage loans were carried out locally. The rationale was that a local bank would be better able to estimate how likely someone was to default since the loan officer lived in the community and had a better idea of its features. The development of credit scores that allow banks to quantitatively estimate the probability of a default for a given set of characteristics revolutionized the housing market.

It is, however, often difficult to incentivize a private organization to conduct this kind of research. The reason for this is that information is a public good and without any kind of intellectual property right to the information - a company cannot derive rents from its generation. If gathering information about asteroid compositions takes ten years and costs millions of dollars in research, but this research is immediately available to the rest of the world, few profit minded groups would bother. However, in cases private groups will conduct the research anyway. These typically follow two forms. First the expected profit is great enough to justify research costs even in the face of freeriding. In this case profits are so large relative to research costs that a single firm is willing to conduct research knowing that the fruits of that research (profit) will be shared by their competitors. In this form, firm size matters. The larger the firm is relative to its competitors, the more likely it is to lead in general research. If firms are of similar size and each are large enough to conduct the research, then it is likely that no one will. There will be a game of chicken to avoid research costs.

Second competitors face barriers to entry. If the research yields some patentable result - for example a new special drill designed to work in outer space then a private company will conduct applied research to develop it knowing that so far as their profits depend on that drill, they are safe from free-riders. In general if firm A can effectively charge other firms for using any information generated by firm A, then such information can be produced privately. This is why newspapers have subscriptions (and why the rise of ad supported internet news agencies is crippling the subscription based model) and also why ratings agencies exist. Where it is possible to exclude people from benefiting from the research, information has by definition become a private good. In many cases multiple firms will require the same kind of research and so we see the emergence of specialized firms to generate and sell that research.

General research, on the other hand is generally not protected by any kind of intellectual property right and it is often difficult to exclude others from its use. It is for this reason that general research is typically carried out by government agencies and by universities. Government institutions are funded by taxes while universities are often funded through a combination of the too. This is however, an imperfect solution. To avoid free-riding we have required that all contribute to the costs of research through taxes, even if the benefits of that research are limited to a subset of the population.