ARIZONA JOURNAL OF INTERDISCIPLINARY STUDIES

VOLUME 7 - SPRING 2021

confluencenter
for Creative Inquiry

ARIZONA JOURNAL OF INTERDISCIPLINARY STUDIES

VOLUME 7 | SPRING 2021

Contents

Introduction

On behalf of the Confluencenter for Creative Inquiry, we are proud to support the Arizona Journal of Interdisciplinary Studies, and celebrate the seventh issue of this student-directed publication. This issue comes at a historic moment, following a year of ongoing uncertainty and loss due to the coronavirus pandemic. Undergraduate students across the globe have suffered tremendous challenges, such as loss of income, unstable housing situations, and adapting to virtual education models.

Challenges can also bring creativity and adaptation. In the midst of the pandemic, undergraduate Editor-In-Chief Robert Lowell and an undergraduate Editorial Team organized over the summer of 2020 with the desire and motivation to lead the production of a new issue of AJIS. Their dedication to cross-campus collaboration, to intellectual curiosity and rigor, and to supporting and promoting the interdisciplinary work of their fellow students, reflects the core values of this journal and of Confluencenter for Creative Inquiry.

We are proud to be part of this editorial effort, as an example of how innovative and interdisciplinary scholarship can transcend borders, disciplines and intellectual differences.

Join me in applauding the inspiring student editors and authors, and our campus partners, as we continue this collaboration towards supporting innovative, interdisciplinary student scholarship.

Javier Duran, Ph.D.
Director
Confluencenter for Creative Inquiry

Letter from the Editor

Exactly one year ago, Ellen Dubinsky, the Scholarly Communications Librarian here at the University of Arizona, and I began working together to revive the Arizona Journal of Interdisciplinary Studies, which went dormant in 2017. The journal now before you is the product of a venture we have been working on ever since, and I cannot thank her enough for all of her help. In addition, I want to express my sincere gratitude to Dr. Susan Crane, who graciously agreed to be our Faculty Advisor, and to Leona Davis and everyone else at the Confluencenter for Creative Inquiry, for their support.

I also would like to thank each editor on our Editorial Team for working tirelessly for many months throughout the selection and peer review process. Thank you also to all of the undergraduates, even those that were not ultimately selected, who sent us their submissions to be considered for publication; without you, there could be no journal at all. Finally, I would like to thank Rachel Small, who met with me early on to give me guidance about the publication process, as well as Michelle Drake, who encouraged me to start on this journey in the first place.

It was not easy reviving an undergraduate interdisciplinary research journal in the middle of a pandemic. It took a team effort, and I am thankful for everyone who contributed to help make this edition a reality. I am very happy and proud to present to you Volume 7 of the Arizona Journal of Interdisciplinary Studies.

Sincerely,

Robert Lowell
Editor-In-Chief, Spring 2021

Traditional Chinese Medicine and Global Conservation Ecology: Remedies for the Future

Maeve Welter

Introduction

Traditional Chinese Medicine (TCM)[1] is often directly involved in cases of wildlife conservation and trafficking (Chang 2017). Many species currently face extinction as a result of increased demand for animal products used in TCM. However, the relationship between TCM and conservation ecology has yet to be addressed in a manner that places equal emphasis on both the principles behind these animal-based remedies and the ecological impacts TCM has on species and ecosystems around the world. The discovery and furtherance of animal product alternatives in TCM is indispensable, as a decrease in the demand for animal products would lessen their value and diminish the global incidence of poaching. Here, I present several plant-based TCM remedies to be used in place of tiger bone and pangolin scale (see Table 1). The endangerment of wildlife due to TCM is an issue that extends beyond the scope of conservation biologists; it is both an environmental and cultural concern. Working constructively to lessen the strain TCM puts on wildlife requires an understanding of wildlife conservation, as well as the scientific principles behind TCM, reasons for its rise in popularity, and cultural significance it has to those who practice it.

Banning the importation of endangered animal products into China and other parts of Asia has proven largely unsuccessful, due to lucrative global black-market networks (Nowell 2007; Findley 2010, Challender, Waterman, et al. 2014; Bessesen 2018). A ban on all TCM remedies

---

[1] Refers to the modern medical tradition invented in the mid-20th century, rather than the collection of natural remedies recorded in ancient China

MAEVE WELTER | 1

would also be impossible, given the strong governmental and popular support TCM enjoys. Understanding the history and modern role of TCM will help us understand the global conservation issues TCM fuels and find a solution that will be beneficial to both TCM practitioners and the affected ecosystems. This paper seeks to explore TCM, conservation ecology, and the connections between the two by investigating three animals, all endangered as a result of their use in TCM remedies: tiger, pangolin, and totoaba. Due to the limited body of research concerning the pangolin and totoaba, the preponderance of this paper will address tiger conservation within the context of TCM.

Traditional Chinese Medicine

The remedies and principles behind TCM have a long history. Physicians began systematically prescribing medicinal plants, animals, and minerals in China since before the Han dynasty (206 BCE-220 CE). During this time, herbal remedies were combined with older shamanistic healing techniques to strengthen a patient's *qi* (life force or vital energy) and maintain a balance between his or her *yin* and *yang* (feminine and masculine energies, respectively). By the late Han dynasty, the popularity and legitimacy of shamanism had declined, while empirical medicine continued to evolve and expand (Bensky & Gamble 1989). Each remedy accesses a specific part of the body through a meridian, or channel, and has a unique combination of attributes (including taste, thermal property, and direction), which makes it suitable for curing a particular illness or set of illnesses (Bensky & Gamble 1989; Chen, et al. 2012). The number of remedies known to physicians of native medicine in China has grown considerably over time. The earliest text of native Chinese medicine to be discovered by archeologists, *Prescriptions for Fifty-two Ailments* (*Wushi'er Bingfang*), described only around 250 remedies. In contrast, the *Encyclopedia of Traditional Chinese Medicinal Substances* (*Zhongyao Da Cidian*), compiled by the Jiangsu

College of New Medicine in 1977, contains descriptions of 5,767 remedies (Bensky & Gamble 1989).

Though plant, animal, and mineral-based native remedies have endured in China for millennia, TCM is, in fact, a relatively new concept. Prior to the introduction of Western medicine to China in the late 19th century (Croizier 1965), there were no alternatives to native Chinese herbal remedies; "traditional" and "modern" medical traditions in China are both constructions of the 20th century. In the 1930s and 1940s, the leaders of the Chinese Communist Party favored Western medicine, believing "traditional medicine represented all the backwardness, superstition and irrationality of the old society and culture with which they were locked in mortal combat" (Croizier 1965, 2). This intense distrust of imperial China and native medicine on the part of the Chinese Communist Party, as well as the Nationalist government that preceded it, resulted in an increased number of medical students learning, and subsequently, practicing Western medicine in Republican China (Xu 1997). Throughout the 20th century, the Republican period trended towards a notion of "modernity" that spanned medicine, science, and engineering.

After the fall of the Republic of China, the People's Republic of China formed, with Mao Zedong as the chairman of the Chinese Communist Party and leader of the state. During the 1950s and 1960s Mao Zedong sought to unite and industrialize the People's Republic of China with his Great Leap Forward campaign (Hsu 2008). Chairman Mao aimed to project an image of China as a self-reliant state, and so he promoted domestic Chinese medical traditions and created the earliest iteration of TCM, as a symbol of cultural nationalism. This party-endorsed form of medicine incorporated science, while still utilizing traditional remedies. However, Western influences on TCM were still suppressed as much as possible (Croizier 1965).

Since its creation in the 1950's, TCM has evolved, and its role in Chinese society has shifted dramatically. At the time of its conception, TCM was used by the Chinese Communist Party as a tool for rejecting "bourgeois" international influences and ideologies. Government propaganda portrayed TCM as a reflection of Maoist proletarian values and downplayed its connection to imperial China, which was seen by the Chinese Communist Party as contemptible and primitive (Croizier 1965). TCM was meant to be a testament to the success of Mao's modern socialist state. Today, both Western medicine and TCM are present and widely practiced in China. In order to prevent and treat diseases, modern TCM still uses ancient techniques, such as employing *yin*-intensive substances (those that are hot or dry) to counter an excess of *yang* (cold and wet), and vice versa. Because *qi*, *yin*, and *yang* all continue to play important roles in TCM treatments, TCM remains somewhat at odds with Western scientific perceptions (Oshima et al. 2015).

Recently, the Chinese Communist Party, led by President Xi Jinping, has been promoting TCM on a broad scale, embracing it as a "2,000-year-old national treasure" (Xinhua 2014). As a result of party endorsement, consuming TCM products has become a way for Chinese citizens to not only honor Chinese traditions and culture, but also to demonstrate their support of the Chinese state and its leadership. The creation of a new upper-middle class in China has also contributed greatly to the rise in TCM product consumption. Within this newly formed socioeconomic stratum, the demand for animal-based remedies in particular has intensified in the last decade. Animal-based TCM remedies, as well as raw, unprocessed goods, such as rhino horns and ivory, sell for astronomical amounts on the black market, since the obtention, transportation, sale, and purchase of many rare animal products is illegal (Maron 2018). Additionally, Xi Jinping's One Belt, One Road initiative, which emphasizes economic partnerships with Europe and other parts of Asia

(Ferdinand 2016), has brought about increased international demand for Chinese medicinal herbs and animal remedies, due to China's newly strengthened international trade relationships (Xinhua 2014).

Now, China is undergoing an entirely different sort of cultural revolution, and TCM is just one of many aspects of Chinese heritage revived under the current leadership of the Chinese Communist Party. Because the Chinese government now embraces its long imperial history, it is now representing TCM in a way that accentuates the technological advancements China has made over thousands of years. TCM today acts as a crucial part of the nationalist party agenda, an emblem of Xi Jinping's "Chinese Dream".

The Chinese Communist Party's staunch support for TCM does not end at rhetoric. Party representatives have not only extolled the practice of TCM through state-controlled media outlets, but have also passed legislation allowing for more lax regulations on TCM, compared to Western medicine. The National Medical Products Association, formerly the China Food and Drug Administration, now requires fewer regulatory trials of TCM remedies, as long as they are prepared using traditional methods (outlined by the National Medical Products Association). These regulations aim to expedite the production and distribution of TCM products, while potentially putting patients in danger (Cyranoski, 2017).

Since TCM does not need to adhere to the standards that Western drugs are held to, it would be appreciably easier for TCM manufacturers to produce and profit off of fake TCM products or genuine remedies laced with potentially harmful additives. Both spurious and contaminated TCM remedies would be deleterious to a patient's wellbeing. In a best-case scenario, these remedies might fail to effectively cure a patient's illness, while in a worst-case scenario, they might result

in worsening of a patient's symptoms, development of new injurious symptoms, or even death.[2]
In addition, prospective TCM practitioners can obtain their medical license without the completion of a standardized test, and establish a TCM clinic without authorization from the National Medical Products Association (Cyranosky 2017). The promotion of TCM at the expense of patient safety is distressing. Permissive regulations surrounding TCM has resulted in increased accessibility and demand for remedies, including those made from illegal animal products.

Unfortunately, TCM is once again steeped in nationalism, and China's strict censorship policies prohibit Chinese citizens from publicly speaking out against TCM. This means TCM is likely to be an indomitable force as long as the Chinese Communist Party supports it, despite the problematic nature of the state's regulations and the impacts TCM has on conservation ecology in other parts of the world.

Case Study 1: Tiger

In order to investigate the impacts TCM has on endangered species, I will explore three case studies in depth. The first case study, the tiger (*Panthera tigris*), is prominent in the contexts of both TCM and conservation. Tigers are poached extensively for their skins, but also for their body parts, almost all of which have a use in TCM. Tiger bone (*Hu Gu*) is the most common tiger-based remedy. Usually administered in the form of pills and wine (Nowell 2000), tiger bone is an acrid, sweet, and warm substance used for countering "damp" and "cold" illnesses, such as joint weakness, stiffness, bone and muscle weakness, lower back pain and knee pain. The first mention of tiger bone as a medicinal substance was recorded in the *Collection of Commentaries on the Classic of the Materia Medica* (*Ben Cao Jing Ji Zhu*) (Bensky & Gamble 1989; Chen, et al. 2012).

---

[2] Generally, TCM is safe when practiced by a licensed professional with genuine, uncontaminated medication; side effects, toxicity, and proper dosage of medically recognized TCM remedies have been established over centuries and can be found in official TCM texts (Bensky & Gamble 1989)

This text, the second substantial ancient Chinese medical record, was compiled during the North and South Kingdom Period (420–589 CE) by physician Tao Hongjing. It consists of an annotation of its predecessor, *The Divine Husbandman's Classic of Materia Medica* (*Shen Nong Ben Cao Jing*),[3] as well as descriptions of 365 new medicinal substances not found in the *Husbandman's Classic* (Zhao, et al. 2018). Since tiger bone was added centuries after the Han dynasty, it is reasonable to assume that while its use as a TCM remedy stems from ancient medical literature, tiger bone was not a founding part of the Chinese pharmacopeia.

As a result of numerous factors, including habitat loss and extensive poaching, tigers are classified as an endangered species, according to the International Union for the Conservation of Nature. Of the nine subspecies of tiger, only six are thought to be extant. The Amur Tiger (*P. t. altaica)*, Northern Indochinese Tiger (*P. t. corbetti*), and Bengal Tiger (*P. t. tigris*) are all endangered, while the Malayan Tiger (*P. t. jacksoni*), Sumatran Tiger (*P. t. sumatrae*), and South China Tiger (*P. t. amoyensis*) are critically endangered.[4] The other three subspecies, the Bali Tiger (*P. t. balica*), Javan Tiger (*P. t. sondaica*), and Caspian Tiger (*P. t. virgata*), have all been declared extinct since the mid-20th century (Goodrich, et al. 2015).

In China, tigers have had cosmological significance since ancient times. They have been thought to possess mystical powers as "representative[s] of heaven that could bring justice to the aggrieved, aid the righteous in times of need, or impose a reign of terror on wrongdoers" (Coggins 2004, 2). In Chinese folklore, the tiger's presence is commonly associated with fluctuating economic, social, and political conditions. During times when the Mandate of Heaven[5] was

---

[3] Compiled between 200 and 250 CE

[4] The South China Tiger has not yet been declared extinct; however, there have been no live wild specimens recorded in over 40 years.

[5] Refers to the notion that Chinese imperial rulers had a divine right to rule, which was rescinded in the case of natural disasters or other turmoil believed to represent celestial wrath at a ruler's incompetence or unworthiness

satisfied, tiger attacks occurred less frequently. These trends cemented the tiger's mythological importance (Coggins 2004).

Tigers, being large predators, require a copious amount of meat to survive. They usually kill large prey about once a week and eat 18–40 kilograms of meat off a carcass each day for up to six days, or until the carcass has been fully consumed (Nowell & Jackson 1996). Tigers rely on a consistent food supply and play a crucial role in their ecosystem, checking populations of large grazing animals, and therefore, preserving the health of the jungles and grassland they inhabit. Over time, tigers' ranges have diminished considerably. Previously, tiger populations had been widely distributed across Asia, into Southwest Asia and Russia. Today, tigers occupy less than 6% of their natural range (Sanderson, et al. 2006).

Although habitat loss was previously thought to be the primary cause of tigers' endangerment, poaching of tigers and their prey has proven to be "an imminent threat to the survival of the species" (Norwell 2000, vi). Killing tigers is not a new phenomenon in Asia. Independent of the influences of TCM or other native medicines, hunting tigers throughout their range has been occurring for centuries.

In India, the reputation of tigers as "man-eaters" dates back to the 17th century, and since then, tiger attacks[6] have provided justification for killing tigers both in defense and for sport (Nowell & Jackson 1996). In China, the Mao-led Chinese Communist Party labeled the South China Tiger a "pest", and ordered its extermination, which was almost completely successful by the end of the 1960's (Nowell 2000). Because of tigers' folkloric significance, their disappearance was seen as an augury of peace and prosperity; with the rise of Mao's leadership, the heavens no

---

[6] Tigers typically attack humans to defend cubs, or when surprised by the presence of humans on their territory, and even then, they attack only after giving warning sounds and/or rushing the intruding human. Eating humans is a last resort for old or incapacitated tigers when they have lost the ability to hunt effectively for regular prey (Nowell & Jackson 1996).

longer needed to send tigers to keep equipoise on earth (Coggins 2004). After being hunted to near extinction, the remaining population of the South China Tiger was federally protected, but since then, the species has become critically endangered, and has not been sighted in the wild since the 1970's. After the 1980's, Chinese demand for tiger bones from other countries increased, presumably due to the lack of tigers that could be killed domestically. Tiger poaching in other parts of Asia escalated, and in 1993, the Chinese government banned the domestic trade of tiger bone, resulting in decreased consumption of tiger bone products in China. However, the presence and demand for tiger bone products has yet to be eliminated (Goodrich, et al. 2015). Presently, the black-market tiger trade is robust, fueled by both illegal hunting in the wild and intensive farming programs throughout Asia, including domestic facilities, billed as tourist attractions, which house upwards of 5000 tigers (Nowell 2007; Abbott & Kooten 2011).

Tigers bought and sold on the black market are used for multiple purposes: tiger bones are used for TCM, penises are used in aphrodisiacs, skin is sold as a luxury item, and live animals are sold as exotic pets. Between 2000 and 2015, 1755 trafficked tigers[7] were confiscated by authorities in 801 separate seizures (Stoner, et al. 2016). This figure does not accurately reflect the total number of tigers globally smuggled during that time, since the majority of animal trafficking goes unnoticed by law enforcement.

Much of the tiger bone TCM products sold today are counterfeit. These misleading remedies contain very little tiger bone, or none altogether. Often, the crushed or powdered bones of bovines, bears, and other big cats serve as a replacement for tiger bone, and without the help of forensic testing, there is no way to discern whether or not a given medicine contains genuine tiger bone. This preponderance of fake tiger bone medicine may be, in part, beneficial to tigers, but it

---

[7] The total global population of tigers during this time was around 3890 (Stoner, et al. 2016)

has also played a role in increasing consumers' demand for genuine tiger products (Nowell 2000). As of October 2018, the sale of tiger products is illegal, but the State Council of the People's Republic of China has authorized the use of tiger bones by qualified, state-registered practitioners of TCM (English.gov.cn 2018).

Setting aside protected areas across Asia would lessen the threat of habitat destruction; however, protecting tigers' habitat would not ensure their total safety. Because the illegal tiger trade is so well-established and widespread, the idea that government policy and wildlife stewardship programs alone will solve the issue of tiger poaching is unrealistic. In the case of tigers, trafficking will only come to a halt if demand for tiger-based TCM products is curbed across Asia. Non-medicinal tiger products have a much shorter history and weaker cultural influence in China, such that associated trafficking of these commodities account for a far smaller proportion of the overall issue (Nowell 2000). Reducing demand for tiger bone medication can be accomplished most efficaciously by finding and promoting viable alternatives, making tiger bone obsolete.

Fortunately, among the conditions tiger bone treats, none are life-threatening nor are they hard to treat using alternate means. In fact, there are tiger bone alternatives within TCM (see Table 1). Acanthopanax root bark (*Wu Jia Pi*) performs similarly to tiger bone; it also dispels wind and dampness in order to strengthen weak bones and sinews (Bensky & Gamble 1989) and ameliorate lower back and knee pain (Chen, et al. 2012). Acanthopanax root bark was recorded in *The Divine Husbandman's Classic of Materia Medica*, which makes it an older remedy than tiger bone. Additionally, it is native to China, and can be harvested in Hubei, Henan, Sichuan, Hunan, and Anhui (Bensky & Gamble 1989). Another alternative to tiger bone is Pyrola (*Lu Han Cao*), which also dispels wind-dampness and treats bone and tendon pain. Pyrola was first mentioned in Li Shi-

Zhen's *Materia Medica*,[8] published in 1578 (Chen, et al. 2012). Both of these herbs are accessible to TCM practitioners, and are available for purchase online (Nuherbs 2019).

There is another option for those looking to directly replicate the effects of tiger bone without the help of plant-based remedies. Pig bones have long been noted as a slightly less effective equivalent to tiger bone (Chen, et al. 2012), which means there is nothing particularly exceptional about tiger bone, save for its high concentrations of calcium and collagen. There have been numerous cases recorded of synthetic tiger bone, enriched with calcium and collagen, used in place of tiger bone to improve the negative musculoskeletal side effects of aromatase inhibitors in women with breast cancer (Li, et al. 2017). Such synthetic tiger bone replacements would be cheap to produce and distribute on a large scale, at no cost to tigers.

Acanthopanax root bark, Pyrola, and synthetic tiger bone compounds are all viable and readily obtainable substitutes for tiger bone. If the Chinese government encouraged TCM practitioners and consumers to respectively discontinue prescribing and purchasing illegal tiger bone medicine in favor of any of these alternatives, then the demand for tiger bone, as well as the financial opportunity attached to tiger poaching, would significantly diminish, lessening the harmful impact of TCM and allowing tiger populations across Asia to stabilize.

Case Study 2: Pangolin

The second case study, the pangolin (family Manidae), is significantly more obscure than the tiger, yet its animal trafficking rates are the highest among all mammals. Pangolin meat is eaten in some parts of Asia and Africa (Actman 2017), but pangolins are primarily targeted for their scales, used in TCM to treat a variety of maladies. Pangolin scale (*Chuan Shan Jia*) is a salty and cool substance that promotes blood circulation, drains pus, stimulates lactation, and lessens

---

[8] Also referred to as *Compendium of Materia Medica* (Zhao, et al. 2018).

swelling and stiffness (Bensky & Gamble 1989; Chen, et al. 2012). Largely believed to be miraculous cure-all, pangolin scale is also used in folk medicine to cure illnesses not mentioned in the TCM *Materia Medica*, including possession by devils and ogres, deafness, nervousness, malarial fever, and uncontrollable crying ("Chinese Medicine and the Pangolin" 1938). Like tiger bone, this remedy was discovered by Tao Hongjing; it first appeared in the *Miscellaneous Records of Famous Physicians* (*Ming Yi Za Zhu*) (Chen, et al. 2012).

Pangolins are anteaters native to Africa and Asia, notable for their keratinous scales, strong claws, and prehensile tails ("Pangolin" 2012). Of the eight species of pangolin, four are endemic to Africa, while the other four are native to Asia. According to the International Union for the Conservation of Nature, all four African pangolin species, the Black-bellied Pangolin (*Phataginus tetradactyla*), White-bellied Pangolin (*Phataginus tricuspis*), Giant Ground Pangolin (*Smutsia gigantea*), and Temminck's Ground Pangolin (*Smutsia temminckii*), are classified as vulnerable (Pietersen, et al. 2014; Waterman, et al. 2014a; 2014b; 2014c). Two of the four Asian pangolin species, the Indian Pangolin (*Manis crassicaudata*) and Philippine Pangolin (*Manis culionensis*), are endangered, and the other two, the Chinese Pangolin (*Manis pentadactyla*) and Sunda Pangolin (*Manis javanica*), are critically endangered (Baillie, et al. 2014; Challender, Baillie, et al. 2014; Challender, Nguyen, et al. 2014; Lagrada, et al. 2014).

Pangolins inhabit a wide variety of tropical and subtropical habitats, including rainforests, grasslands, dry forests, and tree plantations. The four Asian species of pangolin have a natural range that stretches from Southeast Asia (including the Philippines) to Pakistan, spanning India, Nepal, and southern China. The four African species are more densely distributed across Western, Eastern, and Southern Africa. All species of pangolin, across both Asia and Africa, are slow-breeding, solitary, mostly nocturnal insectivores. Pangolins use their long claws to rip into ant

nests and termite mounds and lay their tongues out to catch insects. Their thick skin and scales act as armor, protecting them from ant bites and stings.

Pangolins have a reputation for being elusive, which is surprisingly not an outcome of the recent spike in poaching. Described in 1871 as "not often seen even by the natives" ("The Pangolin" 1871, 142), pangolins remain shy, mysterious, and under-researched, while progressively becoming more imperiled by human action. They are crucial constituents of their ecosystem, preventing the overspread of colonial insects that would otherwise decimate plant life in the forests, grasslands, and plantations they inhabit (Challender, Waterman, et al. 2014).

Like tigers, pangolins also have spiritual significance in China. As a result of their strong connections to the earth, pangolins have long been viewed as deified protectors of the mountains they live on. According to local customs, hunters must perform an incantation to protect themselves from the pangolin before harvesting its meat or scales. These chants, used to neutralize a pangolin's evil retributory spirit, must all include a deferential greeting to the pangolin, a declaration of the hunter's intentions, and an explanation of why the hunter deserves to take the pangolin's life (Coggins 2004). Still performed in pockets of China where pangolins remain, this long-held ritual demonstrates an intense spiritual connection between hunter and the pangolin, as well as a deep respect, lacking in the indiscriminate, industrial-scale pangolin poaching that occurs today.

Aside from their role in TCM, pangolins are also threatened by domestic bushmeat markets, electric fences, and traditional African medicine (Challender, Waterman, et al. 2014). Though TCM is the preeminent threat to pangolins across the globe, indigenous African medicine also plays a sizable role in diminishing local pangolin populations (Soewu & Ayodele 2009; Boakye, et al 2014).

Pangolins rely on their scales for protection against both animal and human predators. When threatened or startled, they roll into a ball, making it easy for poachers to collect and sell them. It has been illegal to trade Asian pangolins since 2000 (Challender, Waterman, et al. 2014), and in 2016, the Convention on International Trade in Endangered Species of Flora and Fauna (CITES) banned trade of African pangolins as well, though many local governments do not have programs in place to actively expunge poaching efforts and enforce this ban (Actman 2017). Despite their internationally protected status, pangolins are hunted in Asia and Africa and traded in vast numbers to satisfy both local and international demands for meat and scales.[9]

The largest seizure of pangolin scales to date occurred in April 2019, when Singaporean authorities intercepted and confiscated an enormous 13-ton consignment of pangolin scales, shipped from Nigeria and destined for Vietnam. The value of this shipment was estimated at $38.7 million USD (Singapore Government 2019). Though exorbitant, this sum indicates the value of only one shipment of pangolin scales, affording a narrow glimpse into the lucrative nature of the pangolin trafficking industry.

In Asia, pangolin scales are processed before being sold or administered to patients. Pangolin scales are most commonly fried, ashed, or roasted ("Chinese Medicine and the Pangolin" 1938), and can be taken in the form of a powder or decoction (Bensky & Gamble 1989; Chen, et al. 2012). These pangolin scale medicines are readily available and widely trusted, yet pangolin scales are made exclusively of keratin, the same material that makes up human hair and nails (Spearman 1967), and are devoid of any apparent miraculous properties.

Raising public awareness, establishing strict enforcement of international laws, and mobilizing local communities does help suppress poaching and trafficking of pangolins and their

---

[9] Over one million individuals were thought to have been trafficked in the last 10 years (Challender, Waterman, et al. 2014)

scales, but in order to tackle the root of the pangolin's problem, it is necessary to find alternatives to pangolin scales (see Table 1). Vaccaria seed (*Wang Bu Liu Xing*), also known as cowherb seed or cow soapwort seed, is an ideal replacement for pangolin scales, as it also reduces swelling, disperses blood flow, promotes lactation, and drains abscesses. The discovery of vaccaria seed, first recorded in the *The Divine Husbandman's Classic of Materia Medica* (Bensky & Gamble 1989; Chen, et al. 2012), predates that of pangolin scale by at least 200 years. Vaccaria seed can be harvested in Jiangsu, Hebei, Henan, and Shaanxi (Bensky & Gamble 1989), or procured online (Nuherbs 2019).

Existing conservation programs, international trade bans, and public awareness, among other things, are all crucial to the pangolin's survival. However, the fate of the pangolin lies ultimately in the hands of the Chinese government, TCM practitioners and consumers. Pangolin scales have no value if there is no demand for them. A long-term, government-backed boycott of pangolin scales would result in a dramatically diminished market for pangolin scale, a loss of incentive for poachers and, most importantly, an opportunity for global pangolin populations to recuperate.

Case Study 3: Totoaba (& Vaquita)

The final case study, the totoaba (*Totoaba macdonaldi*), is perhaps the most obscure out of all animals addressed in this paper. Totoabas are an unassuming species of drum fish native to Mexico (NOAA Fisheries Office of Protected Resources 2012). They have no cultural or historical significance in Asia, yet their swim bladders, the organs used to regulate buoyancy, are consumed on a prodigious scale in China and Hong Kong, in the form of fish maw soup, an expensive delicacy thought to improve joint pain, circulation, pregnancy-related pain, and skin elasticity (Besessen 2018; Leung 2018). The newfound value of their swim bladders has exacerbated

existing patterns of overfishing in the area, and the vaquita (*Phocoena sinus*), the world's rarest cetacean, often ends up as bycatch (Valenzuela-Quiñonez, et al. 2011, Joyce 2016; Leung 2018). Prior to the decline in availability of Chinese bahaba (giant yellow croaker) swim bladders, there was no history of eating totoaba in China, nor an existing market for the fish (Bessesen 2018; Leung 2018). Totoaba is a new, completely invented remedy, to which no iteration of the TCM *Materia Medica* can lend credibility.

Both the totoaba and the vaquita are critically endangered, hovering on the margins of extinction (Findley 2010; Rojas-Bracho & Taylor 2017). They both have highly localized populations, concentrated in the Gulf of California,[10] and suffer from illegal fishing and gill net entanglement, respectively. Totoabas are largely unknown to those not studying, fishing, or consuming them. They were classified as endangered in 1986, and have been critically endangered since 1991 (Findley 2010). Vaquitas, being the single most endangered marine mammals, are slightly more familiar. They were first listed by the IUCN as vulnerable in 1986, but by 1996, they were declared critically endangered (Rojas-Bracho & Taylor 2017).

Totoabas are the largest member of the Sciaenidae family. Like other drum fish, the totoaba is notable for its ability to produce sonorous vibrations during mating season (Bessesen 2018). Since adult totoabas have been known to reach a length of 2 meters (Findley 2010), they have previously been an attractive target for fisheries and sports fishermen alike. Commercial totoaba fishing started during the 1940's, when totoabas still inhabited the Gulf of California in large numbers, but eventually collapsed in the 1970's, due to reduced fish stocks (Findley 2010; NOAA Fisheries Office of Protected Resources 2012; Valenzuela-Quiñonez, et al. 2015). These totoaba fisheries provided the economic foundation on which communities in the area were built

---

[10] Otherwise known as the Sea of Cortez

(Valenzuela-Quiñonez, et al. 2015, Bessesen 2018;), and in the absence of legal totoaba fishing opportunities, poaching has become an enticing way to bring in extra income. One totoaba swim bladder alone is worth over $14,000 on the black market (Joyce 2016; Wade 2017), making it equally valuable as cocaine. Especially large swim bladders have been sold at $250,000 (Bessesen 2018). If current fishing patterns continue, the totoaba will soon become extinct.

Currently, poaching and habitat degradation are the primary factors threatening totoabas' survival. It has been illegal to fish for totoaba since 1975 (Findley 2010), but the remunerative overseas market still incentivizes local fishermen to leave gillnets out, in the hopes of harvesting a swim bladder (Findley 2010; Valenzuela-Quiñonez, et al. 2011). There are other species of fish in the Gulf of California that are legal to harvest, and much of the illicit totoaba poaching is committed under the guise of legal fishing (Bessesen 2018).

Totoabas are also threatened by the changing composition of their nursery waters. Young totoaba are born in the formerly brackish Colorado River Delta, and they move down the Gulf of California as they mature. This transition occurs a result of the juvenile totoabas' preference for low-salinity environments. However, since the Hoover Dam was built, much of the Colorado River's fresh water gets diverted away from the Gulf of California. This has significantly altered the totoaba spawning cycle (NOAA Fisheries Office of Protected Resources 2012; Valenzuela-Quiñonez, et al. 2015). Evidence from totoaba otoliths (ear bones) indicate that totoabas now spend considerably more time maturing in the Colorado River Delta and grow to only half the size of their pre-dam counterparts (Wade 2017).

The vaquita is a small porpoise that inhabits shallow waters and feeds on fish, crustaceans, and squid (Rojas-Bracho & Taylor 2017). Vaquitas themselves have never been the target of fishermen, but their fate has always been tied to totoaba fishing. The switch from small-scale hook-

and-line fishing to commercial gillnet fishing in the late 1940's proved to be the advent of both species' devastation. Like other cetaceans, vaquitas must surface for air every few minutes. When they get ensnared in gillnets intended for totoabas or other fish, vaquitas panic and thrash, which eventually leads to suffocation. Their bodies are usually found hours later, when fishermen come back to check on their nets, and their deaths are not consistently reported, due to the potential repercussions of killing a critically endangered animal, even by accident. Despite the termination of commercial totoaba fishing efforts in the 1970's, vaquita entanglement has persisted, if not worsened. Gillnets are still used en masse to catch totoaba and other fish, and unmanned "ghost nets" left out in the water are equally hazardous (Bessesen 2018).

Over the past 60 years (approximately three generations), vaquita populations have seen a 95% decrease (Rojas-Bracho & Taylor 2017), with a 34% rate of decline per year between 2011 and 2015 (Taylor, et al. 2016). The current vaquita population is unknown, as vaquita births and deaths are not comprehensively recorded, but estimates place the total number of living vaquitas at around 10 or fewer (Leung 2019).

The Mexican government has repeatedly sought to initiate total gillnet bans and introduce alternative fishing gear, looking to mimic the success of TEDs (turtle excluder devices), which allow turtles to pass through shrimp nets without getting enmeshed (IUCN 2016). Unfortunately, these policies have faced staunch opposition from local fishermen, whose incomes are substantial improved[11] with the sale of a totoaba swim bladder.

Though the extinction of the totoaba seems inevitable at this point, a complete cessation of totoaba poaching could allow populations to rebound to levels sufficient to ensure its future existence. A successful boycott on fish maw soup in China and Hong Kong would mitigate the

---

[11] In Mexico, one pound of shrimp goes for $9, while a pound of totoaba swim bladder is worth $4000 (Bessesen 2018)

high demand for totoaba swim bladder and take some of the stress off the remaining totoaba population. In order to reduce demand for totoaba swim bladder, consumers would need to use other substances to achieve the same results. Because totoaba is not recognized by official TCM texts, it is difficult to suggest exact replacements.

The following table consolidates all remedies mentioned in this paper, their medicinal uses, and plant-based alternatives within TCM.

Table 1: Animal-based remedies and their alternatives

Animal Product	Use	Alternative(s)
Tiger Bone	Improves joint pain, bone and muscle weakness	Acanthopanax root bark, Pyrola
Pangolin Scale	Promotes circulation, lactation, lessens swelling	Vaccaria seed

Conclusion

Today, TCM is in a position of substantial influence. Support from the Chinese Communist Party has revived TCM and made it nearly invincible, while the burgeoning middle class in China and other parts of Asia has developed a tremendous appetite for expensive animal-based miracle drugs that are more status symbol than they are effective medication. These two factors have engendered a monumental surge in animal poaching and trafficking around the globe, pushing some species hazardously close to extinction. The current black-market trade of exotic animal parts threatens not only individual species, but also the overall health of their ecosystems. Many species hunted for TCM, including the ones represented in this paper, are fundamental to the biodiversity and well-being of their environment.

Conservation activists have been instrumental in educating the public and establishing in-situ anti-poaching efforts, but they have been unable to confront the crux of the issue: the overseas demand for illegal TCM remedies. It is necessary to reform regulations and common practices of TCM, so that poaching, smuggling, and selling endangered animal parts are no longer profitable activities. In order to do so, the Chinese Communist Party must be actively involved in building domestic awareness around conservation, establishing (and stringently enforcing) policies that protect endangered animals, and denouncing the use of illegal animal products in TCM. TCM practitioners have a secondary role in shaping public perceptions and promoting animal product substitutes.

Now more than ever, conservation biology and the related study of anthropogenic impacts on our planet must be prioritized, as the window for reversing our damage draws increasingly shorter. Further study of contemporary TCM trends, as well as a larger body of research concerning the pangolin, totoaba, and vaquita are necessary in order to continue protecting these unfamiliar and highly imperiled species.

References

Abbott, B., & Kooten, G. C. (2011). Can domestication of wildlife lead to conservation? The economics of tiger farming in China. *Ecological Economics, 70*(4), 721-728. doi:10.1016/j.ecolecon.2010.11.006.

Actman, J. (2017, August 02). The World's Most Trafficked Mammal Just Got Desperately Needed Help. Retrieved from https://news.nationalgeographic.com/2016/09/wildlife-watch-most-trafficked-mammal-just-got-help/.

Baillie, J., Challender, D., Kaspal, P., Khatiwada, A., Mohapatra, R. & Nash, H. (2014). *Manis crassicaudata. The IUCN Red List of Threatened Species 2014*: e.T12761A45221874. http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12761A45221874.en.

Bensky, D., & Gamble, A. (1989). *Chinese Herbal Medicine: Materia Medica*. Seattle: Eastland Pr.

Bessesen, B. (2018). *Vaquita: Science, Politics, and Crime in the Sea of Cortez*. Island Press.

Boakye, M. K., Pietersen, D. W., Kotzé, A., Dalton, D. L., & Jansen, R. (2014). Ethnomedicinal use of African pangolins by traditional medical practitioners in Sierra Leone. *Journal of Ethnobiology and Ethnomedicine,10*(1), 76th ser., 76. doi:10.1186/1746-4269-10-76.

Challender, D., Baillie, J., Ades, G., Kaspal, P., Chan, B., Khatiwada, A., Xu, L., Chin, S., KC, R., Nash, H. & Hsieh, H. (2014). *Manis pentadactyla. The IUCN Red List of Threatened Species 2014*: e.T12764A45222544. http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12764A45222544.en.

Challender, D., Nguyen Van, T., Shepherd, C., Krishnasamy, K., Wang, A., Lee, B., Panjang, E., Fletcher, L., Heng, S., Seah Han Ming, J., Olsson, A., Nguyen The Truong, A., Nguyen Van, Q. & Chung, Y. (2014). *Manis javanica. The IUCN Red List of Threatened Species 2014*: e.T12763A45222303.http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12763A45222303.en.

Challender, D., Waterman, C., and Baillie, J. (2014). Scaling up pangolin conservation. IUCN SSC Pangolin Specialist Group Conservation Action Plan. Zoological Society of London, London, UK.

Chang, J. (2017). China's Legal Response to Trafficking in Wild Animals: The Relationship between International Treaties and Chinese Law. *AJIL Unbound, 111*, 408-412. doi:10.1017/aju.2017.104.

Chen, J. K., Chen, T. T., & Crampton, L. (2012). *Chinese Medical Herbology and Pharmacology*. City of Industry, CA: Art of Medicine Press.

Chinese Medicine and the Pangolin. (1938). *Nature, 141* (3558), 72-72. doi:10.1038/141072b0

Coggins, C. (2004). *The Tiger and the Pangolin : Nature, Culture, and Conservation in China*. Honolulu: University of Hawaii Press.

Croizier, R. C. (1965). Traditional Medicine in Communist China: Science, Communism and Cultural Nationalism. *The China Quarterly, 23*, 1-27. doi:10.1017/s0305741000009863.

Cyranoski, D. (2017). China to ease regulations on traditional medicine. *Nature,551*(7682), 552-553. doi:10.1038/nature.2017.23038.

English.gov.cn. (2018). China to control trade in rhino and tiger products. Retrieved April 28, 2019, from http://english.gov.cn/policies/latest_releases/2018/10/29/content_281476367121088.htm.

Ferdinand, P. (2016). Westward ho—the China Dream and 'One Belt, One Road': Chinese foreign policy under Xi Jinping. *International Affairs, 92*(4), 941-957. doi:10.1111/1468-2346.12660.

Findley, L. (2010). *Totoaba macdonaldi. The IUCN Red List of Threatened Species 2010*: e.T22003A9346099. http://dx.doi.org/10.2305/IUCN.UK.2010-3.RLTS.T22003A9346099.en.

Goodrich, J., Lynam, A., Miquelle, D., Wibisono, H., Kawanishi, K., Pattanavibool, A., Htun, S., Tempa, T., Karki, J., Jhala, Y. & Karanth, U. (2015). *Panthera tigris. The IUCN Red List of Threatened Species 2015*: e.T15955A50659951. http://dx.doi.org/10.2305/IUCN.UK.2015- 2.RLTS.T15955A50659951.en.

Hsu, E. (2008). The History of Chinese Medicine in the Peoples Republic of China and its Globalization. *East Asian Science, Technology and Society,2*(4), 465-484. doi:10.1215/s12280-009-9072-y.

IUCN. (2016, May 19). Between a gill net and a hard place: More pressure on Vaquita. Retrieved March 19, 2019, from https://www.iucn.org/content/between-gill-net-and-hard-place-more-pressure-vaquita.

Joyce, C. (Writer). (2016, February 9). Chinese Taste for Fish Bladder Threatens Rare Porpoise in Mexico [Radio series episode]. In *Goats and Soda*. NPR.

Lagrada, L., Schoppe, S. & Challender, D. (2014). *Manis culionensis. The IUCN Red List of Threatened Species 2014*: e.T136497A45223365.

Leung, H. (2019, March 15). One of About 10 Remaining Vaquitas Found Dead in Mexico. Retrieved from http://time.com/5552189/sea-shepherd-vaquita-porpoise-endangered-mexico/

Leung, R. (2018, June 2). Is Hong Kong's HK$2 billion love of the dried seafood fish maw worth endangering a species? *South China Morning Post*. Retrieved May 7, 2019, from https://www.scmp.com/news/hong-kong/health-environment/article/2148776/hong-kongs-hk2-billion-love-dried-seafood-fish-mawhttp://dx.doi.org/10.2305/IUCN.UK.2014- 2.RLTS.T136497A45223365.en.

Li, Y., Zhang, Z., Cui, F., Liu, J., Wang, Y., Jiang, J., . . . Lu, W. (2017). Traditional Chinese Medicine Bionic Tiger Bone Powder for the Treatment of AI-Associated Musculoskeletal Symptoms. *Evidence-Based Complementary and Alternative Medicine,2017*, 1-8. doi:10.1155/2017/2478565.

Maron, D. F. (2018, December 20). Will a boom in traditional Chinese medicine threaten wildlife? Retrieved April 17, 2019, from https://www.nationalgeographic.com/animals/2018/12/traditional-medicine-decision-threatens-wildlife/.

NOAA Fisheries Office of Protected Resources. (2012, December 5). Totoaba (Totoaba macdonaldi). Retrieved March 19, 2019, from https://web.archive.org/web/20160129212957/http://www.nmfs.noaa.gov/pr/species/fish/totoaba.htm.

Nowell, K. (2007). Asian big cat conservation and trade control in selected range States: evaluating implementation and effectiveness of CITES Recommendations. TRAFFIC International, Cambridge, UK.

Nowell, K. (2000). Far from a cure: The tiger trade revisited. TRAFFIC International, Cambridge, UK.

Nowell, K. & Jackson, P. (1996). *Wild Cats. Status Survey and Conservation Action Plan*. IUCN/SSC Cat Specialist Group, Gland, Switzerland and Cambridge, UK.

Nuherbs. (2019). Nuherbs Chinese Herbs. Retrieved April 27, 2019, from https://nuherbs.com/t/bulk-herbs/nuherbs-and-reg-herbs.

Oshima, E., Adams, J. D., Wang, J., Lien, E. J., Stiles, B., Thurston, D. E., . . . Peplow, P. (2015). *Traditional Chinese Medicine: Scientific Basis for Its Use*. Cambridge: Royal Society of Chemistry.

The Pangolin. (1871). *The Phrenological Journal and Science of Health (1870-1911), 52*(2), 142. Retrieved from http://ezproxy.library.arizona.edu/login?url=https://search-proquest-com.ezproxy2.library.arizona.edu/docview/137914113?accountid=8360.

Pangolin. (2019). In *Encyclopædia Britannica*. Retrieved from https://academic-eb-com.ezproxy3.library.arizona.edu/levels/collegiate/article/pangolin/58261.

Pietersen, D., Waterman, C., Hywood, L., Rankin, P. & Soewu, D. (2014). *Smutsia temminckii. The IUCN Red List of Threatened Species 2014*: e.T12765A45222717. http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12765A45222717.en.

Rojas-Bracho, L. & Taylor, B.L. (2017). *Phocoena sinus. The IUCN Red List of Threatened Species 2017*: e.T17028A50370296. http://dx.doi.org/10.2305/IUCN.UK.2017-2.RLTS.T17028A50370296.en.

Sanderson, E., Forrest, J., Loucks, C., Ginsberg, J., Dinerstein, E., Seidensticker, J., Leimgruber, P., Songer, M., Heydlauff, A., O'Brien, T., Bryja, G., Klenzendorf, S and Wikramanayake, E. (2006). Setting Priorities for the Conservation and Recovery of Wild Tigers: 2005-2015. The Technical Assessment. WCS, WWF, Smithsonian, and NFWF-STF, New York and Washington, DC, USA.

Singapore Government, Singapore Customs; National Parks Board. (2019, April 4). *National Parks Board and Singapore Customs seized largest shipment of pangolin scales in a single haul globally in recent years* [Press release]. Retrieved May 5, 2019, from https://www.customs.gov.sg/-/media/cus/files/media-releases/2018/for-website-04042019-largest-seizure-of-pangolin-scales.pdf.

Soewu, D. A., & Ayodele, I. A. (2009). Utilisation of Pangolin (Manis sps) in traditional Yorubic medicine in Ijebu province, Ogun State, Nigeria. *Journal of Ethnobiology and Ethnomedicine,5*(1), 39th ser. doi:10.1186/1746-4269-5-39.

Spearman, R. I. (1967). On the nature of the horny scales of the pangolin. *Journal of the Linnean Society of London, Zoology, 46*(310), 267-273. doi:10.1111/j.1096-3642.1967.tb00508.x.

Stoner, S., Krishnasamy, K., Wittmann, T., Delean, S., and Cassey, P. (2016). Reduced to Skin and Bones Re-examined: An analysis of Tiger seizures from 13 range countries from 2000-2015. TRAFFIC International, Selangor, Malaysia.

Taylor, B. L., Rojas-Bracho, L., Moore, J., Jaramillo-Legorreta, A., Hoef, J. M., Cardenas-Hinojosa, G., . . . Hammond, P. S. (2016). Extinction is Imminent for Mexico's Endemic Porpoise Unless Fishery Bycatch is Eliminated. *Conservation Letters, 10*(5), 588-595. doi:10.1111/conl.12331.

Valenzuela-Quiñonez, F., Arreguín-Sánchez, F., Salas-Márquez, S., García-De-León, F. J., Garza, J. C., Román-Rodríguez, M. J., & De-Anda-Montañez, J. A. (2015). Critically Endangered totoaba Totoaba macdonaldi: Signs of recovery and potential threats after a population collapse. *Endangered Species Research,29*(1), 1-11. doi:10.3354/esr00693.

Valenzuela-Quiñonez, F., García-de-León, F. J., De-Anda-Montañez, J. A., & Balart, E. F. (2011). La Totoaba del Golfo de California: ¿Una Especie en Peligro de Extinción? *Interciencia,36*(9), 664-671. Retrieved May 7, 2019.

Wade, L. (2017, December 08). Ending poaching may not save the doomed totoaba. Retrieved May 6, 2019, from https://www.sciencemag.org/news/2017/03/ending-poaching-may-not-save-doomed-totoaba.

Waterman, C., Pietersen, D., Soewu, D., Hywood, L. & Rankin, P. (2014a). *Phataginus tetradactyla. The IUCN Red List of Threatened Species 2014*: e.T12766A45222929. http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12766A45222929.en.

Waterman, C., Pietersen, D., Soewu, D., Hywood, L. & Rankin, P. (2014b). *Phataginus tricuspis. The IUCN Red List of Threatened Species 2014*: e.T12767A45223135. http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12767A45223135.en.

Waterman, C., Pietersen, D., Hywood, L., Rankin, P. & Soewu, D. (2014c). *Smutsia gigantea. The IUCN Red List of Threatened Species 2014*: e.T12762A45222061. http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T12762A45222061.en

Xinhua. (2014, August 24). "Initiative Helps Boost Traditional Chinese Medicine," August 24, 2014). *China Daily*. Retrieved April 11, 2019.

Xu, X. (1997). 'National Essence' vs 'Science': Chinese Native Physicians' Fight for Legitimacy, 1912-37. *Modern Asian Studies,31*(4), 847-877. Retrieved April 5, 2019.

Zhao, Z., Guo, P., & Brand, E. (2018, April 10). A concise classification of *bencao* (*materia medica*). Retrieved April 23, 2019, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5894148/.

A "Perfect Dictatorship": The PRI, Corruption, and Autocracy in Mexico

Mohammad Azeem Khan and Andrea Lara-García

Abstract

Mexico's celebrated democratic transition in the year 2000 belied the centuries of authoritarian rule that preceded it. From the start of Spanish colonialism in 1519, Mexico has been a monarchy and a personalist dictatorship, and was most recently a single-party dictatorship helmed by the Partido Revolucionario Institucional (PRI) for over seventy years. The PRI's decades-long rule may be surprising to observers given Mexico's proximity to the United States, a country that has long sought to promote democratic rule in the region. This paper explores the reasons behind the longevity of the PRI's rule, especially given the failure of other authoritarian forms of government in Mexico's history. We found that the structural conditions following the Mexican Revolution favored the creation of a party-based regime, as many former revolutionaries later became party elites. These findings form a basis through which to understand the current political environment in Mexico, as well as a cautionary tale for those who wish to protect the democratic institutions in this fledgling democracy.

Introduction

For Mexicans, the year 2000 was more than just the start of a new millennium. It also proved to be the start of a new political era. After 71 consecutive years in power, the country's ruling political party, the Partido Revolucionario Institucional (PRI), lost the popular vote and allowed an opposition candidate to assume the office of the presidency. Mexico's peaceful democratic transition has made the country a model case study in the Latin American context,

where democracy has historically struggled to take hold (Santiso, 2006). Despite these achievements, corruption marks the country's democracy as one that is still in development and may be hampering economic and social development (Emmerich & Benitez Manaut, 2010).

All of this begs the question of why Mexico has struggled so much to attain and maintain a democracy in the first place. After all, Mexico shares the world's most frequently crossed border with one of the world's largest democracies. The United States has repeatedly intervened in Latin American politics to ostensibly ensure democratic outcomes; in fact, when the French attempted to reestablish a monarchy on Mexican territory in 1865, the American government offered to send its armed forces to assist Mexican freedom fighters (*Our Documents - Monroe Doctrine (1823)*, n.d.). How is it that Mexico, with its proximity to one of the most vocally democratic countries in the world, came to have an authoritarian government so deeply entrenched and efficient that it was once referred to by Nobel Prize-winning poet Mario Vargas Llosa as the "perfect dictatorship" (Langston, 2017)?

As it turns out, the roots of authoritarianism run deep in Mexico, and a multitude of factors—including colonial history and the circumstances surrounding the Mexican revolution—prevented an effective democratic transition until 2000. We briefly outline Mexico's history to illuminate the foundations upon which the country's political institutions were built. From there, we analyze the strategies that the PRI used to remain in power and why they were so effective in the Mexican context. Finally, we discuss why Mexico was more vulnerable to a single-party autocracy than to other forms of authoritarian governance and how seven decades of PRI rule continues to affect the social, economic, and political environment to this day.

A Political History of Mexico

The Olmecs, the earliest documented civilization in the territory now known as Mexico, laid the political foundations for many of the indigenous societies that followed. Their political leaders are believed to have depicted themselves as demi-gods who exercised control through their supposed connection to the divine (Kirkwood, 2000, p. 18). Similarly to the Olmecs' framework, the pre-Colombian civilization in Teotihuacán is thought to have controlled its population through a sophisticated hierarchical political apparatus, with the divine ruler and priest class at the top and slave laborers at the bottom (Kirkwood, 2000, p. 20). As Teotihuacán elites began to lose their grip on power in the surrounding areas, they relied on a highly developed military to terrorize their subjects into submission (Kirkwood, 2000, p. 21). Later Mexican civilizations followed similar practices: politics and religion were inextricably intertwined, and warfare was frequently used to subjugate the populace. Since many of the details surrounding these civilizations are lost in the annals of history, one hesitates to classify these political systems as monarchies or military dictatorships. Regardless, it is important to consider how these pre-Colombian proto-autocracies shaped Mexico's political history in centuries to come, especially given the willingness of the Spanish crown to exploit these political systems for their own benefit.

Though the arrival of the Spanish in 1519 brought about monumental change, in some regions extant political institutions persisted. Indigenous elites ("caciques") pledged their allegiance to the Spanish crown and were granted some degree of autonomy over their subjects; in return, they collected tributes for the Spanish and adopted European customs (Pohl, 2012). Despite this continuity, it is undeniable that this era is where we start to see the development of the "state" in the Western sense; the Spanish consolidated these fragmented civilizations under their rule and developed centralized institutions to extract a wealth of

natural resources and transport it to market. During this time, Mexico can clearly be classified as a monarchy, since it operated entirely as an extension of the Spanish Empire.

Mexico gained independence from the Spanish in 1821 and began to shape its own political identity after centuries of colonial rule. In 1824, Mexico drafted its first constitution, which was modeled heavily after the American Constitution and contained numerous provisions to promote the development of a democratic state (Macías, 2011). Regardless, the state struggled to permanently establish democratic governance and periodically fell back into authoritarianism, such as in two attempts to re-establish a monarchy in the country (Macías, 2011). During one period of democracy, General Porfirio Diaz made several attempts to overthrow the presidency and install himself as the head of state; though these coups were always unsuccessful, he was eventually elected president in 1867 and would remain in office for the next three decades, a period historians refer to as the Porfiriato (*Mexico During the Porfiriato - The Mexican Revolution and the United States*, n.d.).

The Porfiriato would be defined by remarkable leaps in infrastructure development and economic expansion, though the vast majority of Mexicans would never enjoy the benefits of these advancements. Diaz surrounded himself with a technocratic elite that he referred to as the "cientificos," men who served as policy advisors and were lauded for their economic prowess; regardless, the Diaz regime relied heavily on authoritarian techniques of political repression to keep critics in line, including false imprisonment, assassinations, and threats of military force (*Mexico During the Porfiriato - The Mexican Revolution and the United States*, n.d.). Though Diaz frequently cited his military experience as one of his greatest assets as a leader, he did not actively involve military leadership in decision-making. As such, the Porfiriato would most likely be classified as a personalistic regime with heavy support from the military, or perhaps an evolutionary military regime. The latter term refers to a regime that begins as one type and later transforms, or "evolves", into another; in the case of the Porfiriato, the argument can be made

that the Porfiriato began as a military dictatorship but transformed into a personalist regime with time.

Try as he might, Porfirio Diaz was ultimately unable to stem the tide of revolution in Mexico. He was finally ousted in 1910 amidst an outbreak of conflict fueled by decades of inequality; the constitutional republic that emerged from the seven-year-long struggle was, on paper at least, strongly democratic and radical in its provisions for social and political rights (Bantjes, 2011). Still, democracy continued to struggle to find a foothold in the country. Though the revolution officially ended in 1917, political instability continued to plague the country; after president-elect Alvaro Obregon was assassinated in 1928, his hand-picked successor, Plutarco Elias Calles, took the reins amidst the chaos. Ultimately, Calles decided that the best way to stop the infighting was to consolidate various revolutionary organizations under a single party in 1929: the Partido Nacional Revolucionario (PNR). This party would later come to be known as the Partido Revolucionario Institucional (PRI) (Langston, 2017).

The single-party regime stewarded by the PRI was famously referred to as a "perfect dictatorship" by poet Mario Vargas Llosa; indeed, the party was able to maintain a stranglehold over Mexican politics for seventy years (Langston, 2017). The president, who served as the party's leader, would have the final say in selecting his successor. Institutionally, most power was concentrated under the PRI's executive party leadership instead of the general rank-and-file membership so as to prevent fragmentation (Langston, 2017). The PRI maintained power through political patronage, election-rigging, and in some instances, violence. This system proved effective, so long as the party could deliver spoils to its patrons and economic stability to the general public; however, Mexico experienced several economic crises in the last two decades of the 20th century, leading to increased scrutiny of the party's ability to lead the country. Fearing a popular revolt, party leaders began gradually allowing opposition candidates to win and quietly reformed the country's electoral system (Langston, 2017).

In 2000, the PRI was finally ousted from the office of the presidency in the country's first truly democratic election (Langston, 2017). In the two decades since, the presidency has been mostly free of accusations of electoral corruption, and elections in the country have been classified as "generally free and fair" by scholars of democracy and electoral integrity (Emmerich & Benitez Manaut, 2010). While Mexico's democracy is far from perfect, it is strong enough that it is unlikely to slip into authoritarianism in the near future (Emmerich & Benitez Manaut, 2010).

Analyzing Mexico's Political System under the PRI

The PRI exemplifies the longevity of rule that party-based regimes can attain, as it effectively monopolized Mexico's political arena for seventy years. Understanding the political history of the state, we can now analyze the specific apparatuses that allowed for the PRI to endure as long as it did. First and foremost, we must acknowledge the bedrock of the PRI's power—co-optation—as one of the many strategies autocratic regimes employ to consolidate their rule.

Co-optation can take two primary forms: economic and political. Within the former derivation, the regime buys the loyalty of its supporters, incorporating politically salient groups into their operation by securing for them a direct line of rents, such as exclusive government contracts of preferential tax breaks, to ensure their complacency (Haber, 2006). In this context, "rents" are primarily economic and political capitulations granted to the backers of a regime. On the other hand, there is political co-optation, or as Haber calls it, "organizational proliferation", the regime either aligns the incentives of its rank-and-file membership with its leadership in order to ensure their loyalty, or it makes the leadership coordinate with other organizations newly-mandated by the regime itself, like paramilitary forces or dispute arbitration mechanisms (e.g., the Italian "Blackshirts" under Mussolini, the Chilean military tribunal courts under

Pinochet). This is all done to raise the cost of collective action, disincentivizing the subversion of the regime's institutions by rogue actors, as they may compromise their own position and, thus, their own payoffs coming from the regime (Haber, 2006). In Mexico under the PRI, we see both derivations of co-optation at play.

Economically, a few of the main currencies by which the PRI co-opted their support base were property rights and public policies, which they doled out selectively to their winning coalitions (Haber, 2008). This clientelist system, distinct because of its informal business-state relationship, dates back to the Porfiriato, during which Diaz would reward his coalition of landowners, bankers, and industrialists with tangible benefits delivered through policy (Haber, 2006). According to Haber, regimes would generate these economic rents by the institution of policies like regulatory barriers (2006). For example, Diaz would provide the industrialists in his coalition protective import tariffs and financial barriers to market entry, sheltering his circle from foreign and domestic competition alike (Haber, 2008). Similarly, in the oil industry, Diaz awarded drilling concessions, tax exemptions, and property rights to federal lands to incentivize oil magnates in order to benefit from any attracted investment coming off of the fruitful oil revenue (Haber et al., 2003). After Diaz's deposition, the PRI would continue the informal working relationship of the state and businesses for the pursuit of profit (Thacker, 2012).

By ensuring stable relations with Mexico's "big business", the PRI established itself as necessary for the acquisition of profit and, thus, stabilized itself, co-opting the economic sector into relying on the politicians to keep the rents and concessions flowing (Purcell, 1981). An example of this symbiotic relationship can be found in Mexico's relation to its sugar industry. In the 1940s, the PRI government passed decrees that bound sugar cane farmers to work for the sugar mill owners by fixing their wages to the wholesale sugar prices that the mill owners set (Purcell, 1981). This allowed the sugar magnates to profit from exploiting their labor base with

menial wages, and, in turn, the government co-opted those same magnates into controlling the farmers at their behest. Mexico's financial sector offers another example of this mutually beneficial and informal relationship. As was common during the Porfiriato, the PRI established alliances with Mexican banks (Haber, 2008). The terms of this alliance included the government allowing bankers to write their own rules when it came to their lending practices, while in return banks kept a line of credit selectively open for PRI politicians; this created a "revolving door" effect by which politicians would receive rents from bankers or become bankers themselves (Haber, Razo, et al., 2003). As demonstrated, the PRI managed to stay relevant and stable in the political arena because of its strategic economic co-optation of key industries in the economy.

Politically, the PRI engaged in co-optation by shoring up the membership within its ranks to ensure loyalty. They were able to do this by maintaining a corporatist, patrimonial structure that "coordinated and controlled disparate societal interests under auspices of the revolutionary party" (Edmonds-Poli & Shirk, 2009, p. 192). By cultivating the concept of the inclusive "revolutionary family," the PRI was able to co-opt different interest groups under the same wide tent and control them by handpicking their leadership, manipulating the operations and objectives of said interest groups (Edmonds-Poli & Shirk, 2009). Many of the PRI's rank-and-file were of urban middle-class origin, gaining access to their positions via "kinship ties or shared educational and occupational experiences" (Haber, 2008, p. 27). These members, admitted into positions that allowed for ample personal gain, then agreed to certain norms about power-sharing within the regime; for instance, norms on office rotations stabilized the PRI as it ensured access to higher level positions for politicians with progressive ambition, keeping them loyal to the party as they waited to access the upper offices and their corresponding benefits (Haber, 2008) This system allowed the PRI to acts as the hegemony of Mexican politics until 2000.

By no means was the PRI completely without opposition, however, and it was also not beyond the application of political violence to subdue dissenters. Opposition parties continued to exist, but because the PRI was able to form a diverse coalition drawing from a wide berth of constituencies by banking on its nationalist rhetoric and policies, it was able to safeguard its own dominance (Haber, 2008). When it was threatened by opposition groups that it could not co-opt politically, it would use violence selectively, such as clamping down on labor protests or intimidating critical journalists (Bartman, 2018; Haber, 2008). The PRI was also prone to commit electoral fraud by manipulating electoral results and using the media and state resources to mislead their citizens about oppositional candidates and the party's own positions (Edmonds-Poli & Shirk, 2009). They were able to accomplish this because of the tentacle-like nature of the PRI apparatus, which reached into various sectors of Mexican political life and maintained absolute political hegemony with their personnel and resources. In sum, while economic and political co-optation helped the PRI ascend to dominance in Mexico's political system, by no means were these strategies sufficient to maintain their undemocratic grasp on power. Like any other authoritarian regime, the PRI did not shun the use of state violence and electoral fraud to suppress those who would not follow the party line.

Mexico's Structural Conditions and the PRI

In this section we consider the structural factors of Mexico that allowed for the rise of the PRI as a party-based regime. We begin by explaining why other authoritarian regime types such as militarist, personalist, and monarchic regimes could not crop up in Mexico given the country's structural conditions. This section will end with our postulation of the primary factor that made Mexico a party-based authoritarian regime.

What is remarkable about Mexican governance when compared to the rest of Latin American politics is its stark lack of military coups (Hachemer, 2017). The closest Mexico ever came to a military rule was indeed the Porfiriato, but even then, it was an evolutionary—or a transitioning—regime, trending towards a personalist dictatorship. This means that the Porfiriato was never going to just be a military dictatorship, but rather a regime headed by a single dictator at the forefront as opposed to a cabal of military officers. When compared with other developing states, Mexico succeeded in implementing the formula of civilian supremacy over the military apparatus, incrementally reducing the power of the military by subordinating it to civilian leadership (Kurzer, 2015). Among the cited reasons for this phenomenon is Mexico's professionalized civil service corps as a competent leadership group, as well as the hegemonic nature of the PRI as a revolutionary party (Serrano, 1995). Regarding the latter, it is argued that the legacy of the Revolution of 1910 promoted a climate favorable to peaceable relations between civilian and military elite, as both arose out of the same roots in the populations and worked together to bring about revolutionary change; thus, the military esteemed themselves as "guardians of the Revolution" and of the revolutionary party, the PRI (Serrano, 1995, p. 428). We conclude that, because of this unique civil-military relationship, a military regime could not be born on Mexican soil.

Mexico arguably was a personalist regime during the Porfiriato, which means it was characterized by the reign of a single despotic leader. This is why the PRI, as a revolutionary party, could not adopt that same form without losing their legitimacy. One million Mexicans lost their lives not only in the insurrection against Porfirio Diaz but in the resulting civil wars as well (Edmonds-Poli & Shirk, 2009, p. 39). Diaz's regime was characterized by high amounts of inequality despite the high economic growth rates; many suffered from harsh working conditions, food shortages, and rampant squalor (Edmonds-Poli & Shirk, 2009). What followed

the overthrow of the Porfiriato was a brutal civil war between various liberal and conservative factions. The PRI rose into power amidst this chaos by carrying with them, in many former revolutionaries' eyes, the mantle of the spirit of the original Revolution of 1910. Tied into this rhetoric is a mythologized view of the civil war that emphasizes Mexican nationalism, statism, social justice, and political liberalism as the revolutionary goals the PRI, as a whole, worked to bring about (Kurzer, 2015). Reneging on these revolutionary goals by instituting a personalist regime would run counter to the PRI's legitimacy. Thus, Mexico's structural condition as a "revolutionary state" prevents personalism from occurring within the PRI.

Mexico could have been considered a monarchy in its colonial past, as under the Spanish crown, a viceroy was appointed to oversee the colony and report back to the crown (Kurzer, 2015). In 1865, French forces attempted to interfere in Mexican politics by propping up the Emperor Maximilian and his wife Carlotta as the monarchs; they were promptly deposed and the would-be king was executed by government forces (*Our Documents - Monroe Doctrine (1823)*, n.d.). Because of this legacy, it would have been highly unlikely that Mexico's citizenry would have ever accepted a monarchy, given the brutal colonial legacy of exploitation the Spanish and French left in the country (Martínez, 2004). Another possible explanation may reside in the Mexican state's relationship with the church. During the Spanish reign, the crown and the Catholic Church reached an agreement that allowed Catholicism to be the dominant religion of "New Spain"; Porfirio Diaz continued such a relationship but tempered its power by forcing the church to relinquish power in regards to politics (Kurzer, 2015). As a monarch tended to derive their legitimacy from some sort of divine right, the church played an important political role in bestowing that divine right to a ruler. But since the Catholic Church had exchanged its political power for dominance in the cultural and religious sphere of Mexico, it would not have been able to regain that role without severe consequences. Thus, Mexico after

the Revolution did not devolve into monarchy.

This leads us to consider what structural condition *did* lead to the creation of a party-based regime. We ultimately believe that the answer lies in *how* the PRI came to power: revolution. The Porfiriato perpetuated a degree of wealth inequality so unacceptable that a left-wing revolutionary force had to depose it altogether with violence. Diaz's strategy of economic development favored the wealthy classes, cronies who propped him up as a ruler; his regime was also "repressive, favorable to the interests of foreigners, and politically antiliberal [dictatorial]" for most Mexicans (Kurzer, 2015, p. 211). This led to an uprising led by oppositional groups from the northern state of Sonora. After beating Diaz's forces on the battlefield, the victors forced Diaz to abdicate and went on to hold the election of 1913, the "fairest election in Mexican history," which saw the victory of Francisco Madero (Kurzer, 2015, p. 211).

Madero was assassinated soon after by Diaz loyalists, and this cycle of violence continued to repeat until Plutarco Calles founded the PNR in 1929 as a forum for all of the veterans of the Revolution to organize and pool their power and resources after the assassination of President-elect Alvaro Obregon (Brenner, 1971). Even though they came into formal power via election, the PRI was composed of the revolutionaries that fought against the old regime. That means the Revolution and the crises it wrought fostered cohesion amongst the revolutionaries, formalizing their cooperation as political elites and producing leadership seen as legitimate in the eyes of the Mexicans for whom they fought for (Levitsky & Way, 2012). Because of the violent and bloody road to victory, the founders of the PRI thought it prudent to bind together and create the "revolutionary family" as it is known today.

Impact on Human Rights and the Economy

Invariably, the PRI's system has created long standing effects in two major areas of Mexico's politics: narco-violence and the economy. Narco-violence, or drug violence as a result of the Mexican war on drugs, has had an adverse effect on human rights in the country (Human Rights Watch, n.d.). Alternatively, the PRI's statist approach to economic development helped sustain what many economists regard as the "Mexican miracle". How the authoritarian apparatus of the PRI interacted with the daily life of the party's constituents can be better discerned by analyzing these two areas of concern.

Mexico's drug trafficking organizations (DTOs) are the most significant players in the country's organized crime arena (Beittel, 2019). What makes these DTOs so threatening is their utilization of violence in not only an inter-organizational setting, but also against "the government, political candidates, and the media" (Beittel, 2019, p. 3). Many commentators have also argued that the Mexican DTOs can be "excessive" and "exceptional" in their displays of violence (Beittel, 2019, p. 3). In recent years, many government actors, including those of the PRI, have been implicated in criminal collusion with the DTOs; for instance, Javier Duarte, the Governor of Veracruz, was arrested in 2017 on the charge of criminal involvement with the DTOs that led to the "forcible disappearances" of more than five thousand persons (Beittel, 2019, p. 5). Given the PRI's proclivity towards co-optation, it stands to reason that many actors under the PRI umbrella would be prone to collude with the DTOs that hold immense power in the region, using their positions in the state to benefit from the illicit activities the DTOs carry out (Shirk & Wallman, 2015).

This complicity between the state and the DTOs proliferated the drug trade and its resulting violence in Mexico. According to Shirk and Wallman, because of the PRI's extended tenure in Mexico's high offices since the 1940s, "high-level corruption created a

blanket of protection throughout the political system. The illicit arrangements between

Mexican drug traffickers and state authorities might be better understood not as criminals

corrupting the state but criminals as subjects of the state" (2015, p. 1359). The PRI also kept

corruption within law enforcement itself hidden from public eye "through control of the

media and bureaucratic reshuffling of abusive police from one set of forces to another"

(Davis, 2006, p. 64). Under the PRI, the military also became corrupt due to its longstanding

relationship in combating the DTOs (Davis, 2006).

This led to Mexico's democratic turnaround of 2000, in which a member of the

opposition party (PAN), Vicente Fox, was elected to office, signaling a regime change into

something more democratic. In this transitional period, however, violent intrastate conflict has

worsened (Davis, 2006; Shirk & Wallman, 2015). With law enforcement's former master, the

PRI, relinquishing its old absolute power, corrupt police officers have fled to the criminal

underworld (Davis, 2006). As a result, presidents since Fox have heightened the violence in the

drug wars by centralizing their own police forces and increasing their usage of coercive power.

As a result, the state has overseen many human rights abuses such as "extrajudicial killings,

enforced disappearances, and torture" in their struggle against the DTOs (Human Rights Watch,

n.d.). This was all a consequence of the PRI's authoritarian regime cooperating with the violent

cogs of Mexico's organized crime.

The PRI put forth a system that saw heavy intervention in the economy by the state,

including the nationalization of many key industries such as railways, agricultural holdings, and

petroleum (Cline, 1962). This particular route of state-economy relationships was pursued due

to the nationalist nature of the Revolution of 1910, where Mexican nationals felt that their

newfound identity as a state warranted a more hands-on approach to the economy to better

deliver social welfare services to the citizenry (Kurzer, 2015). Rather than encroaching upon the

market's ability to function and thereby throttling it, the state under the PRI managed to align its interests with those of its private sector (Camp, 1989). An example of this can be found in the relationship between Mexican state and its banking industry that began in the 1920s. The state intervened in the banking industry by creating "a dependency of the expanding manufacturing sectors on government-supplied capital…a key channel for an exchange of private- and public sector leadership.a model in later years for state involvement in manufacturing and mining" (Camp, 1989, p. 17). Policies such as this manufactured an environment in which the relationship between public and private sector resembled that of a "mentor-disciple", for when the state "consolidated power, it doled out resources to nurture the nascent private sector" (Camp, 1989, p. 16). This was to set the tone of the state's interactions with the economy in the decades to come.

By 1940, these interactions had become formally institutionalized, and the PRI leadership began to turn towards a policy of industrialization to usher in economic growth. In order to incentivize the private-sector to target developmental goals, they created tax breaks and directed funds into domestic industries through providing them loans from state-sponsored banks with below-market value interest rates (Camp, 1989). By the administration of Miguel Aleman in 1946, a more technocratic approach to running the economy was adopted by the PRI's newer breed of educated administrators that ushered in the fastest period of social and economic change in Mexico. This time period saw a cooperative relationship between the government and the private sector, given Aleman's affinity towards private initiative (Camp, 1989).

The relationship would continue with relative stability until the 1960s and 1970s, when political agitation began to cast doubt on the system of the PRI as well as the 1940s' style of economic development. This strategy of development was none other than import-substitution industrialization, which is where domestic industry is jump-started by substituting

foreign imports for domestic goods; the high growth rates that occurred during this era were what dubbed this era the "Mexican miracle" (Antonio, 1977). By the 1970s, however, the state and the economy increasingly fissured apart because of more pressures on the PRI to liberalize politically; in order to appease growing opposition, presidents during this time implemented more populist policies. Despite still maintaining an alliance with the private sector, the PRI government in 1977 turned to international borrowing and deficit spending (Camp, 1989). Debt in both sectors increased, leading to an economic crisis in Mexico in the 1980s.

Eventually, the PRI looked to absolve the economic conundrum it found itself in by opening up its markets to trading partners, its sights set on the neighbor to the north: the United States (Cline, 1962). Eventually, the PRI government would stabilize the economic situation by pursuing neoliberal policies and free trade-agreements with its neighbors, although poverty levels in the country have not been alleviated (Kurzer, 2015). As can be seen, the PRI's direct role in the Mexican economy ushered in fast growth from the 1940s to the 1970s, but this party also holds the responsibility for the economic crises and growing inequality the country has faced and continues to face.

Conclusion: Prospects of Democratization

Those tasked with stewarding Mexico's fledgling democracy cannot afford to forget the centuries of authoritarianism that precede them. Though the PRI has not given any indication that it plans to forcefully retake power, echoes of its favored modes of power continue to ring through Mexican politics. High levels of corruption in the country, especially at the state and municipal level, harken back to the co-optation strategies that the PRI used to keep local officials in check; low levels of citizen participation and generalized distrust in the police are reminiscent

of the days when one's safety was only guaranteed by not rocking the boat (Emmerich & Benitez Manaut, 2010). At this point, the PRI's tentacular reach into the fabric of Mexican society as a megalithic party has entrenched corruption to the point that revolutionary steps would need to be taken to undo the damage. Mexico's current president, Andres Manuel Lopez Obrador, was elected on this very principle; one of his campaign promises was to stop political corruption at every level of governance (Malkin & Villegas, 2018). Of course, reality has proven to be much more complicated than Obrador bargained for, and his administration has failed to make significant headway in the fight against corruption in Mexico (Malkin & Villegas, 2018). Only time will tell if Mexico's democracy can uproot the "perfect dictatorship."

References

Antonio, L. (1977). Import Substitution in Mexico: Past and Present. *PAST AND PRESENT*, 13.

Bantjes, A. A. (2011). *The Mexican Revolution* (pp. 330–346). Wiley-Blackwell. https://doi.org/10.1002/9781444391633.ch19.

Bartman, J. M. (2018). Murder in Mexico: Are journalists victims of general violence or targeted political violence? *Democratization*, *25*(7), 1093–1113. https://doi.org/10.1080/13510347.2018.1445998.

Beittel, J. S. (2019). *Mexico: Organized Crime and Drug Trafficking Organizations*. 35.

Brenner, A. (1971). *The wind that swept Mexico;the history of the Mexican revolution, 1910-1942.* Austin,. http://hdl.handle.net/2027/mdp.39015019045601.

Camp, R. (1989). *Entrepreneurs and politics in twentieth-century Mexico* (p. 306).

Cline, H. F. (1962). *Mexico;revolution to evolution, 1940-1960.* London,. http://hdl.handle.net/2027/txu.059173022895307.

Davis, D. E. (2006). Undermining the Rule of Law: Democratization and the Dark Side of Police Reform in Mexico. *Latin American Politics and Society*, *48*(1), 55–86. JSTOR.

Edmonds-Poli, E., & Shirk, D. A. (2009). *Contemporary Mexican politics*. Rowman & Littlefield Publishers.

Emmerich, G., & Benitez Manaut, R. (2010). The State of Democracy in Mexico. *Norteamérica*, *5*.

Haber, S. (2006). Authoritarian government. *The Oxford Handbook of Political Economy*, 693–707.

Haber, S. (2008). *Mexico since 1980*. Cambridge University Press. http://ezproxy.library.arizona.edu/login?url=http://www.netLibrary.com/urlapi.asp?action=summary&v=1&bookid=234363.

Haber, S., Maurer, N., & Razo, A. (2003). When the Law Does Not Matter: The Rise and Decline of the Mexican Oil Industry. *The Journal of Economic History*, *63*(1), 1–32. https://doi.org/10.1017/S0022050703001712.

Haber, S., Razo, A., & Maurer, N. (2003). *The Politics of Property Rights: Political Instability, Credible Commitments, and Economic Growth in Mexico, 1876–1929*. Cambridge University Press. http://ebookcentral.proquest.com/lib/uaz/detail.action?docID=218088.

Hachemer, P. (2017). Mexico: A Civil-Military "Pact" Unravelling? In *Reforming Civil-Military Relations in New Democracies* (pp. 83–101). Springer.

Human Rights Watch. (n.d.). *Mexico*. Human Rights Watch. Retrieved April 29, 2020, from https://www.hrw.org/americas/mexico.

Kirkwood, B. (2000). *History of Mexico*. Greenwood Publishing Group, Incorporated. http://ebookcentral.proquest.com/lib/uaz/detail.action?docID=3000600.

Kurzer, P. (2015). *Comparative politics in the age of globalization*. Mcgraw Hill.

Langston, J. (2017). Democratization and Authoritarian Party Survival: Mexico's PRI. In *Democratization and Authoritarian Party Survival*. Oxford University Press. https://www-oxfordscholarship-com.ezproxy2.library.arizona.edu/view/10.1093/acprof:oso/9780190628512.001.0001/acprof-9780190628512.

Levitsky, S. R., & Way, L. A. (2012). Beyond Patronage: Violent Struggle, Ruling Party Cohesion, and Authoritarian Durability. *Perspectives on Politics*, *10*(4), 869–889. https://doi.org/10.1017/S1537592712002861.

Macías, F. (2011, February 24). *The History of the Mexican Constitution | In Custodia Legis: Law Librarians of Congress* [Webpage]. //blogs.loc.gov/law/2011/02/the-history-of-the-mexican-constitution/.

Malkin, E., & Villegas, P. (2018, October 20). Mexico's President-Elect Rethinks His Campaign Promises. *The New York Times*. https://www.nytimes.com/2018/10/20/world/americas/mexico-lopez-obrador-campaign-promises.html.

Martínez, M. E. (2004). The Black Blood of New Spain: Limpieza de Sangre, Racial Violence, and Gendered Power in Early Colonial Mexico. *William and Mary Quarterly*, *61*(3), 479. https://doi.org/10.2307/3491806.

Mexico During the Porfiriato—The Mexican Revolution and the United States. (n.d.). Retrieved April 25, 2020, from https://www.loc.gov/exhibits/mexican-revolution-and-the-united-states/mexico-during-the-porfiriato.html.

Our Documents—Monroe Doctrine (1823). (n.d.). Retrieved April 30, 2020, from https://www.ourdocuments.gov/doc.php?flash=false&doc=23.

Pohl, J. M. D. (2012, September 24). *The Spanish Conquest and the Archaeology of the Colonial and Republican Periods*. The Oxford Handbook of Mesoamerican Archaeology. https://doi.org/10.1093/oxfordhb/9780195390933.013.0068.

Purcell, S. K. (1981). Business-Government Relations in Mexico: The Case of the Sugar Industry. *Comparative Politics*, *13*(2), 211–233. JSTOR. https://doi.org/10.2307/421623.

Santiso, J. (2006). *Democracy in (Latin) America*. *27*. https://doi.org/10.1787/245130057361
Serrano, M. (1995). The Armed Branch of the State: Civil-Military Relations in Mexico. *Journal of Latin American Studies*, *27*(2), 423–448. JSTOR.

Shirk, D., & Wallman, J. (2015). Understanding Mexico's Drug Violence. *Journal of Conflict Resolution*, *59*(8), 1348–1376. https://doi.org/10.1177/0022002715587049.

Thacker, S. C. (2012, January 13). *Big Business, Democracy, and the Politics of Competition*. The Oxford Handbook of Mexican Politics. https://doi.org/10.1093/oxfordhb/9780195377385.013.0014.

Exploring the Roots of
Poor Medication Adherence

Parisa Mosallai

ABSTRACT

Background. Approximately one-half of U.S. adults do not take their medications as prescribed. This increases healthcare costs, reduces patient health, and gives rise to safety concerns.

Objectives. We assessed patient medication adherence and identified the most common barriers for this.

Methods. We called Medicare Part D beneficiaries from University of Arizona SinfoniaRx Medication Management Center. Subjects qualified to participate in the study if they were identified as nonadherent. Nonadherence was identified using patient self-reports and prescription refill data which defined nonadherence as proportion of days covered (PDC) less than 80%. Data were collected from March 2018 and April 2018.

Results. When a gap in drug refill history is found, SinfoniaRx's computer system creates an alert in a patient's profile. The sample included 85 calls and 42 adherence alerts were observed. Of 42 alerts, 52.4% of alerts were denied by the subjects, and 95.3% of subjects reported adherence issues unrelated to the triggered alert. No alerts were identified for 56 subjects, and 72 adherence issues were identified via self-report only. The two most common reasons for nonadherence were cost (19.75%) and poor health literacy (22.2%).

Conclusions. Medication nonadherence is an important public health crisis because it reduces health and wellbeing as well as increases the risk for life-threatening events. Noncompliance also increases healthcare costs. Cost and poor health literacy were the most frequent barriers to compliance in our small study. Furthermore, most subjects were not aware of their

noncompliance. To address noncompliance, we suggested lower cost drug alternatives and provided patient education.

INTRODUCTION

Here, we present findings about medication nonadherence researched from the sampling of telephone calls at SinfoniaRx between February 2018 and April 2018. We investigated common reasons cited for nonadherence and potential solutions for such noncompliance.

SinfoniaRx

SinfoniaRx, an institute of the University of Arizona College of Pharmacy (formerly the Medication Management Center), is a call center that provides Medication Therapy Management (MTM) services to Medicare Part D beneficiaries. Medicare Prescription Drug Plans (Part D) include prescription drug coverage and are offered by private insurance companies approved by Medicare. Many Medicare Part D patients are eligible to qualify for the MTM program. Additional qualifications include having multiple chronic conditions, taking multiple Part D drugs, and spending a certain threshold of annual out-of-pocket costs. MTM is a medical service provided by pharmacists to improve therapy and optimize patient therapeutic outcomes.[1] Over the telephone, pharmacists target nonadherence, gaps in care, cost-saving opportunities, safety concerns such as therapeutic duplications or drug-drug interactions, and medication questions. SinfoniaRx provides MTM services on behalf of various insurance companies it is contracted with. Data were collected at SinfoniaRx by researchers listening to the calls. During a call, the caller reviews the patient's medications, inquiring if the drug is being taken and how often. At the end of the call, the patient is transferred to a pharmacist who reviews the list of medications again to ensure no problems and to answer remaining patient questions.

Medication Nonadherence: The Silent Epidemic

Medication adherence is the extent to which patients take their medicines as prescribed by their healthcare providers.[2–5] Often, the simple act of taking a medication according to directions is, to some adults, a complex task.[2] Medication adherence is more than reading the directions on a label, but it is an intricate interaction that involves a social environment, the patient, his/her disease, and the healthcare professionals.[3,6] Examples of nonadherence include but are not limited to skipping drug doses, taking drugs at the wrong time, or failing to fill prescriptions.[2] Two categories of medication nonadherence are recognized: primary nonadherence by which the patient fails to initially fill a written prescription and secondary nonadherence by which the patient fails to follow directions or refill prescriptions.[2,5] A patient may decide to be non-adherent purposefully, or the patient may be unintentionally non-adherent. Three phases of medication adherence exist: 1) the patient fills the prescription written by a healthcare provider (initiation); 2) the patient takes the drug according to instructions (implementation) and 3) continues to take the drug appropriately with no premature cessation (discontinuation).[7] Nonadherence may occur at any one of these three phases.

Approximately half of the 3.2 billion annual prescriptions in the U.S. are not taken as directed.[2,8] This is a concern because medication nonadherence can harm the patient and increase general healthcare burdens. Because drugs are the most common type of therapy used in healthcare, maximizing their use is required for populations who rely on long-term pharmacotherapy.[9] Noncompliance reduces treatment efficacy, limits chronic disease management, and increases costs due to poorer health outcomes.[6–7,10] Annually, an estimated 125,000 deaths in the U.S. are due to medication nonadherence and poor adherence is thought to account for ~10% of hospitalizations and ~64% of hospital readmissions within 30 days.[2,10] Noncompliance causes

revenue losses for pharmaceutical companies and pharmacies when prescriptions are not filled, and when prescriptions are filled but not used, drug wastage is significant.[8,10] These costs total $100 billion annually so that drug adherence experts declared this problem a public health crisis.[4,10]

METHODS

Measuring Medication Adherence

Various tools are used to measure adherence, both directly and indirectly as well as quantitatively and qualitatively. Direct methods include use of biomarkers, determination of the concentration of a drug in blood or urine, dried blood spot analysis (DBS), and direct observed therapy (whereby a healthcare provider watches people take their medications).[2,10] Indirect methods to measure adherence are pill counts, prescription refill rates, electronic monitors (which consist of a microprocessor placed in a medication container which records the date and time the bottle was opened), and patient self-report (which includes medication diaries, patient questionnaires, and patient interviews).[2] Some examples of standardized questionnaires are the Morinsky Medication Adherence Scale (MMAS), Medication Adherence Questionnaire (MAQ), and the Medication Adherence Report Scale (MARS).[8,11]

Technologically advanced measurement methods include mobile phone alerts and monitoring, electronic cap event monitoring (MEMS caps), smart pills, computerized logbooks, and ingestible medication markers.[2,10] However, these methods take time to get approval from the FDA and are generally too expensive for the average patient to utilize.[8] From refill rates and pill counts, adherence can be numerically calculated in many ways for claims-based adherence measurement. Medication possession ratio (MPR) is calculated by dividing days' supply

obtained by the refill interval, and the proportion of days covered (PDC) calculation is defined as the number of medication-available days divided by the number of days in a specified time period.[1,12–13]

Disadvantages of Methods

There is no perfect standard to measure adherence because each method has its disadvantages.[1] Self-report, although easy and inexpensive to perform, has been found to overestimate adherence because patients often say what they believe their healthcare provider would like to hear.[9,14] In addition, self-report relies on patients' perception of their behavior and is thus subject to recall and reporting bias.[8] Prescription refill records and pill counts do not verify whether a patient is actually taking their medication and whether they are doing so correctly.[8] Electronic devices may not always be precise and do not indicate the amount of medicine ingested.[14] Indirect methods are more common because they are generally easier to conduct, less expensive, and noninvasive but often less accurate and do not provide all the required information to measure adherence.[8,15] Direct methods are generally more costly, more time-consuming, and more invasive but more accurate.[11] Many require the patient to go to a healthcare clinic, which is particularly difficult for the elderly who are disproportionately affected by physical limitations.[2,16]

Data Collection

SinfoniaRx has access to drug refill histories so gaps can be identified. Once a gap is found, the computer system automatically creates an alert in the patient's profile and each alert identifies one medication. The gaps in refills are calculated using the PDC method with nonadherence defined as PDC less than 80%.[1] Both refills and patient self-reports of Medicare Part D participants who are offered MTM services from SinfoniaRx were sampled to measure

adherence. Alerts indicated a gap in refill history (quantitative data), and the subject's conversation was self-reported qualitative data. Data were recorded when an alert was triggered and when a participant verbalized concern(s) about taking or having access to a medicine. Direct quotes were also recorded. All participants verbalized consent to be recorded and monitored.

RESULTS

Reasons for Nonadherence

Many factors affect medication adherence.[7] Factors include patient, treatment, or healthcare system factors.[2,10] Patient aspects include forgetfulness, personal beliefs on medications, physical limitations, health literacy, demographic factors such as race or education, limited English proficiency, poor social support network, behavioral/psychological factors such as depression, and socioeconomic factors.[7,10] Treatment factors consist of regimen complexity (number of medications and intake frequency), actual or perceived side effects, duration of therapy, changes in medication regimen, and severity of symptoms.[2,15] Healthcare system factors include a weak provider-patient relationship, poor provider communication skills or poor information provision, high drug costs and copayments, limited access to care, insurance complications, and pharmacy refill issues.[2,9] For this study, factors raised in subject phone calls were recorded. Thus, 14 factors were developed and these appear in Table 2.

Findings

Data appear in Tables 1 and 2. Adherent patients with gaps in refill history were due to hospitalizations during which medications are provided by the hospital, use of samples sourced from a healthcare provider and as such will not appear in refill history, and a change in dose or medication whereby the patient stops the

Table 1 Adherence Issues Identified from Calls

Number of calls	85
Total number of alerts in system	42
Adherence issues detected with no alerts in system	72
Alerts with adherent patients	11
Patients with no alerts	56

Hospitalized (H); Use of Samples (UOS), Change in Dose or Medication (CDOM)

previous regimen as directed by their provider, causing a gap for the previous regimen. We identified 18 subjects with 1 alert, 9 subjects with 2 alerts, and 2 subjects with 3 alerts. The most frequent adherence barriers are depicted in Table 2. Thus, based on this study, cost and poor health literacy were the most common reasons for nonadherence. Past research of the Medicare Part D beneficiary population suggests that cost and side effects were the most common barriers.[1] However, forgetfulness is also cited as a common cause of noncompliance.[2,16] The most frequent obstacles to adherence differ among populations. For example, some minority groups are more likely to be non-adherent due to language barriers or cultural beliefs about medicines.[7] Cost-related nonadherence may disproportionately affect elderly individuals who have limited income.[10]

Table 2 Adherence Issues

Adherence issue	Occurrences	Direct Quotes	Solutions Discussed
Forget	5 (6.2%)	"I won't say that I haven't missed one every now and then."	Suggest patient to use a pill box, alarm, sticky note, or calendar. Keep medications easily visible.
Cost	16 (19.75%)	"I can't afford that kind of money. They'll get every nickel they can from me."	Suggest patient to get medications via mail order, get longer supplies, get the generic instead of the brand, pay their deductible, get samples, try alternatives (call insurance, talk to prescriber, talk to pharmacist).
Personal Beliefs	5 (6.2%)	"I don't like 'em but I know I got to take 'em." "I won't take them tablets."	Suggest patient to talk to prescriber and to tell prescriber before stopping medication(s). Transfer patient to pharmacist.
Physical Limits	4 (4.9%)	"I don't get out often cause they took my car from me."	Suggest patient to get caregiver(s), easy open caps, pharmacy delivery service.
Poor Health Literacy	18 (22.2%)	"I just take it. I don't pay attention to what it is."	Inform and educate patient.
Behavioral/Psychological	6 (7.4%)	"I'm dying of liver cancer, that's my only concern." "My having fun is going to the doctor's office."	Provide empathy.
Physician Factors	8 (9.9%)	"Nobody ever give me an explanation."	Suggest patient to talk to prescriber.
Lack of or Poor Social Support Network	1 (1.2%)	(jokingly) "They're only helping me cause of my will."	Suggest patient to get caregiver(s).

Table 2 Adherence Issues

Adherence issue	Occurrences	Direct Quotes	Solutions Discussed
Regimen Complexity	2 (2.5%)	"I just can't handle taking four more pills."	Explain the importance of taking medication(s). Suggest patient to ask pharmacist for help.
Side Effects	5 (6.2%)	*NQO	Suggest patient to talk to prescriber if having side effects with medicines before changing the dosage on their own or stopping the regimen. Suggest alternatives and talking to a pharmacist. Educate patient on importance of taking medication(s).
Change in Symptoms	2 (2.5%)	"I didn't feel like it was helping."	Educate patient on importance of taking medication(s). Suggest patient to talk to prescriber if having issues with medicines before changing the dosage on their own or stopping the regimen.
Pharmacy Issue/Refill Problem	5 (6.2%)	*NQO	Suggest patient to continue to take medication(s) if possible or take alternatives, call prescriber, call pharmacy, request refills in advance.
Insurance Problem	1 (1.2%)	"What's the difference between copay and insurance?"	Suggest patient to call insurance and prescriber for help.
Limited English Proficiency	3 (3.7%)	"No speak English"	Language services are provided. Use a sheet saying how to say "one moment please" in several different languages, then transfer patient.
Unable to Recall	22	*NQO	NA

Table 2 Adherence Issues

Adherence issue	Occurrences	Direct Quotes	Solutions Discussed
No Reason	3	*NQO	NA

*NQO: No Quote Obtained

DISCUSSION

Cost: Healthcare System Factor

Nonadherence due to high cost is a healthcare system factor and this may be due to low income, limited insurance coverage for outpatient medications, and/or increased out-of-pocket costs.[10] When copayments increase, patients with chronic illnesses are likely to skip or stop taking their drugs.[2] Sometimes patients enter the 'donut hole', a coverage gap that occurs once the patient has reached the drug plan coverage limit. Copayments drastically rise in the donut hole and two subjects in this study expressed concerns about this. SinfoniaRx has interventions to help people pay for their medicines such as using mail order, samples from a provider, and three-month versus one-month supplies. Patients are also encouraged to use generic versions which are less expensive. Two patients were advised to file an appeal to their insurance to get a coverage rule waived for them. Finally, patients are advised to obtain lower cost alternatives by talking with their provider or a pharmacist. Communication is vital with nonadherence; patients are often advised to discuss their issue(s) with their physician.

> "I can't afford that kind of money."
>
> "They'll get every nickel they can from me."

Poor Health Literacy: Patient Factor

Another common cause of medication nonadherence is poor health literacy which is a patient factor. Health literacy is the ability to understand health information and services to make appropriate health decisions.[17] Patients must read and understand medical information and each is expected to self-manage chronic conditions using basic literacy skills such as listening, speaking, writing, reading, and numeracy. Health literacy requires general knowledge of health topics as well as the nature and causes of one's chronic condition(s).[2,17] Poor health literacy is evidenced by unawareness of the need for a medication or inability to understand dosage instructions when they have been explained by a healthcare provider.[15]

> "I just take it. I don't pay attention to what it is."
>
> "I didn't know that. I'm glad I'm reading that!"

Many instruments are developed to measure health literacy such as the Short Test of Functional Health Literacy in Adults and the Rapid Estimates of Adult Literacy in Medicine.[18] According to the National Assessment of Adult Literacy conducted in 2003, 77 million U.S. adults (35%) have basic or below basic health literacy, whereas only 26.4 million (12%) have proficient health literacy.[15,19] Poor health literacy is most common in immigrants, racial/ethnic minorities, older patients, and those with less education.[19]

At SinfoniaRx, interventions are implemented to educate the patient using pictorial and audiovisual educational material instead of written instructions and creating a 'shame-free

environment' which may increase the likelihood of an admission of limited literacy.[2] Other

solutions may empower patients to ask questions such as the teach-back method which allows

healthcare providers to confirm patient understanding by asking him/her to explain in his/her

own words what they need to know or do.[15] Given that 13.5 million Americans were identified to

have limited English proficiency (LEP) in 2011, providing information in the patient's native

language can secure medication comprehension for these patients.[20] SinfoniaRx has translators to

provide verbal information in a native language when a patient has LEP and this may address

medication nonadherence issues for these patients.

Denying Gaps in Adherence

"Yes I'm taking my medicines; I want to live!"

Of the 42 alerts that appeared in the system, 52.4% were denied by the subjects. Only 4.7% of

patients reported adherence issues related to the triggered alert; the other issues were not related

to the alert. This frequent failure to recognize nonadherence is consistent with the results of a

previous study performed at SinfoniaRx.[1] Subjects deny nonadherence for many reasons so

open-ended questions are used to generate discussion.[11] For example, asking "Do you take your

medication as prescribed?" is a closed-ended question and typically generates binary and false

answers compared to open-ended questions "How often do you take your medication?"[8]

Normalizing the likelihood of nonadherence increases the accuracy of self-reported medication

use.[8] For example, asking "How often do you forget to take your medicine?" may cause patients

to overestimate their adherence.[11] Phrasing the question, "People often forget to take their

medicines; in the past week, how many days did you forget to take your medicine?" can result in

more accurate responses.[2,11] The second question reduces potential guilt patients may feel

because they view nonadherence as normal and not entirely their fault when missing doses. Thus, medication adherence measured via self-reports requires question phrasing to minimize response bias. At SinfoniaRx, staff used different wording to address alerts and this variety in phrasing may account for some of the patient denials.

In addition, nonadherence gaps may be due to cash-paid prescriptions because these are not recorded in drug refill histories, but this can be addressed by the patient during the telephone conversation. Furthermore, beneficiaries of MTM services must often manage multiple chronic conditions and several medications; thus, missing a few doses every week is likely and may indicate poor management of conditions. As a result, patients with comorbidities may not notice their own gaps in therapy.

LIMITATIONS

This study has several limitations. Assessing adherence qualitatively is difficult because nonadherence is typically not straightforward but is due to complex interactions of multiple factors. Next, the methods of the study were limited to self-reporting and refill history. Self-reporting is subject to bias, and consistent question phrasing was not used. Alerts were frequent, but because only 4 of 85 subjects recognized their refill gap, little quantitative data were obtained. As a result, this study largely depended on self-reporting to evaluate adherence. However, subjects rarely clearly stated they were non-adherent with a medicine but they implied it or vaguely referred to it. Because the call was not performed by the author, subjects could not be asked to clarify adherence and it was not possible to use structured questionnaires to measure adherence, which may have produced more accurate results. Unfortunately, ambiguous cases

were frequent. At times the patient did not know if they were taking a medication or claimed their physician told them to take a lower dosage of the drug. Some patients had a complicated story for the reason for refill gap(s), which for the sake of time could not be pursued in depth.

CONCLUSION

Medication nonadherence is an important public health issue because it lowers quality of health and causes avoidable costs to the healthcare system.[2,10] Unfortunately, nonadherence is often overlooked which can be addressed by offering adherence training for providers.[15] However, providers cannot shoulder this burden entirely. Rather, drug recipients should be aware of the medication benefits they have. Many called in this study knew little about the MTM program for which they were qualified. We must provide better benefit information to patients so that these resources can be more frequently used.

Our results confirmed that cost and poor health literacy are the most common causes of nonadherence in a Medicare Part D beneficiary population. Based on previous research, cost, side effects, and forgetfulness are other common reasons for U.S. adults.[2,10] The most common barriers to adherence differ with different populations such as racial/ethnic minorities; healthcare providers should consider this phenomenon to identify adherence issues more effectively.

At SinfoniaRx, the interventions for cost concerns and poor health literacy include finding lower cost alternatives and educating patients. General interventions include providing written action plans and written instructions for patients, reducing regimen complexity such as creating the simplest dosing schedule, and individually delivering provider-to-patient interventions.[2,8] In this way, patient-centered communication can be achieved, and individual adherence issues can be

identified and resolved. Finally, because medication adherence is a multifaceted issue, it requires interventions at multiple levels of the healthcare system.[3]

ACKNOWLEDGEMENTS

REFERENCES

1.   Hincapie, A.L., Taylor, A.M., Boesen, K.P., Warholak, T. (2015). Understanding Reasons for Nonadherence to Medications in a Medicare Part D Beneficiary Sample. *Journal of Managed Care & Specialty Pharmacy, 21*(5), 391–399.

2.   Bosworth, H. (2012). *Enhancing Medication Adherence: The Public Health Dilemma.* Springer Healthcare.

3.   Hugtenburg, J.G., Timmers, L., Elders, P.J.M., Vervloet, M., & Van, D.L. (2013). Definitions, variants, and causes of nonadherence with medication: A challenge for tailored interventions. *Patient Preference and Adherence,* 675–682.

4.   Chisholm-Burns, M.A., & Spivey, C.A. (2012). The 'cost' of medication nonadherence: Consequences we cannot afford to accept. *Journal of the American Pharmacists Association, 52*(6), 823–826.

5.   Lee, S.Q., Raamkumar, A.S., Li, J., Cao, Y., Witedwittayanusat, K., Chen, L., & Theng, Y.L. (2018). Reasons for Primary Medication Nonadherence: A Systematic Review and Metric Analysis. *Journal of Managed Care & Specialty Pharmacy*, *24*(8), 778–794.

6.   Kleinsinger, F. (2018). The Unmet Challenge of Medication Nonadherence. *Permanente Journal, 22*, 18–033.

7.   McQuaid, E.L., & Wendy L. (2017). Cultural Issues in Medication Adherence: Disparities and Directions. *Journal of General Internal Medicine*, *33*(2), 200–206.

8.   Lam, W.Y., & Fresco, P. (2015). Medication Adherence Measures: An Overview. *BioMed Research International*, 217047–12.

9.   Kennedy, J., Tuleu, I., Mackay, K. (2008). Unfilled Prescriptions of Medicare Beneficiaries: Prevalence, Reasons, and Types of Medicines Prescribed. *Journal of Managed Care Pharmacy*, *14*(6), 453–560.

10.   Tanna, S., & Lawson, G. (2011). *Analytical Chemistry for Assessing Medication Adherence*. Elsevier.

11.   Chan, A.H.Y., Horne, R., Hankins, M., & Chisari, C. (2020). The Medication Adherence Report Scale: A measurement tool for eliciting patients' reports of nonadherence. *British Journal of Clinical Pharmacology, 86*(7), 1281–1288.

12.   Zhu, V.J, Tu, W., Rosenman, M.B, & Overhage, J.M. (2015). Nonadherence to Oral Antihyperglycemic Agents: Subsequent Hospitalization and Mortality among Patients with Type 2 Diabetes in Clinical Practice. *Studies in Health Technology and Informatics, 216*, 60–63.

13.   Desai, R., & Nayak, R. (2019). Effects of Medication Nonadherence and Comorbidity on Health Resource Utilization in Schizophrenia. *Journal of Managed Care & Specialty Pharmacy, 25*(1), 37–44,239.

14.   Whalen, U., & Sunil, K. (2009). Medication Adherence. In Rowe, B.H. (Ed.), *Evidence-Based Emergency Medicine* (pp. 49–57). John Wiley & Sons Ltd.

15.   Brown, M.T., & Bussell, J.K. (2011). Medication Adherence: WHO Cares?. *Mayo Clinic Proceedings, 86*(4), 304–314.

16.   Thiruchselvam, T., Naglie, G., Moineddin, R., Charles, J., Orlando, L., Jaglal, S., Snow, W., Tierney, M.C. (2012). Risk factors for medication nonadherence in older adults with cognitive impairment who live alone. *International Journal of Geriatric Psychiatry, 27*(12), 1275–1282.

17. U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. (2015). *Health Literacy Online: A guide to simplifying the user experience*. https://health.gov/healthliteracyonline/.

18. Julienne, K.K., Grzywacz, J.G, Arcury, T.A., Ip, E.H., Nguyen, H.T., Bell, R.A., Saldana, S., Quandt, S.A. (2011). Performance of Health Literacy Tests Among Older Adults with Diabetes. *Journal of General Internal Medicine, 27*(5), 534–540.

19. Fan, J.H., Lyons, S.A, Goodman, M.S, Blanchard, M.S, & Kaphingst, K.A. (2016). Relationship Between Health Literacy and Unintentional and Intentional Medication Nonadherence in Medically Underserved Patients With Type 2 Diabetes. *The Diabetes Educator, 42*(2), 199–208.

20. Ryan, C. (2010). *Language use in the United States: 2011*. U.S. Census Bureau. https://www.census.gov/prod/2013pubs/acs-22.pdf.

Principal Component Analysis for Clustering Stock Portfolios

Taylor Agarwal, Henk Quelle, Cooper Ryan

Abstract

Motivation: Due to the sheer size of the stock market and its dependencies on several factors, a catch-all method of determining stock performance continues to elude the public. Recent developments in machine learning have opened the door to new possibilities for a predictive algorithm. Principal Component Analysis (PCA) is one such tool that has allowed for the discovery of hidden interconnectivity in large data sets. We use PCA alongside k-means clustering to obtain groups of stocks with similar historical structure with the potential to assist in predictive stock management.

Results: In just over five seconds, our algorithm groups a hundred stocks from the New York Stock Exchange with mean correlation 0.2483 into fifty portfolios with mean correlation 0.4299. On average, the stocks in these fifty portfolios experienced price increases and decreases on the same day 65.34% of the time in the sample time frame of 4/17/2000 to 11/10/2017. The algorithm can be extended to encompass more stocks.

Implications: This algorithm provides a means to identify stocks with similar structure in both the short and long term. Stocks belonging to the same portfolio after clustering are shown to have positive and negative returns at the same time within the user defined periods of time.

Introduction

Principal Component Analysis, or PCA, is a feature detection and data reduction technique that has gained popularity in recent years for analyzing large data sets. PCA has an extensive history in numerical linear algebra dating back to 1901, when Karl Pearson devised the method before it was practically feasible on a computer [4]. Computational advancements since then has allowed PCA to become one of the most widely used data analysis techniques. PCA discovers the *principal components* of the stock data, which are eigenvectors that explain large portions of the variance in the data. The original data is then expressed in terms of these lower dimension features, revealing its hidden structure and cluster points.

From here, k-means clustering groups the lower dimensional stock data in portfolios with very high correlation between stocks and a high frequency of identical daily increases and decreases. The first mention of k-means clustering was by Hugo Steinhaus in 1956, but a formal algorithm was not presented until 1965 by Edward Forgy [1]. The apparent simplicity of k-means clustering conceals its ability to group complex data in meaningful ways. We will take advantage of these two powerful tools to form stock portfolios.

Since the founding of the New York Stock Exchange (NYSE) in 1817, investors have tried to analyze its behavior for financial gain. The NYSE consists of thousands of publicly traded companies ranging from small startups all the way up to massive corporations in the Fortune 500. Each stock is individually affected by a whole host of interwoven factors, including its employees, its management, its ability to sell goods or produce services, its competitors, and the health of the

market as a whole, to name a few. This complexity hides the fact that stocks have connections to each other that are not readily observable to the naked eye. We will reveal these connections.

Some of our portfolios are of no surprise, for instance it is easily imaginable that two health insurance provides like Anthem and UnitedHealth Group should experience similar market dynamics. However, we will see that there are some stocks that are heavily correlated with no discernible reason why, such as telecom company AT&T and fast food restaurant McDonald's.

Methodology

Using a bi-layered feed forward neural network, we identify portfolios of stocks with similar correlation and daily changes. The first layer is Principal Component Analysis, a data reduction technique used frequently on high dimensional data to reduce high dimensional data to a few key components. The second layer is k-means clustering, a labelling technique that is used to group data of similar structure. Linear algebra is used extensively in this paper, so we will discuss our notation.

Notation

X	Matrices, all entries in R
x	Vectors and columns of matrices, all in R
n	Scalar values and elements of vectors, all in R
$\lvert\lvert \cdot \rvert\rvert_2$	Euclidean Norm, $\lvert\lvert x \rvert\rvert_2 = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$
$\lvert \cdot \rvert$	Cardinality of a set

We additionally define a validity metric that will be used frequently throughout the paper which we call the Mean Percent of Days or MPD. The MPD gives the average of the percent of days each pair of stocks in a portfolio both increase or both decrease from their respective prices on the previous day. If we define the percent of days on which a pair of stocks $\boldsymbol{a}$ and $\boldsymbol{b}$ with m returns without dividends observations both increase or both decrease to be

$$p_{ab} = \frac{1}{m}\lvert\{i \mid a_i b_i > 0 \ for \ i = 1, \ldots, n\}\rvert$$

then

$$MPD = \frac{2}{n(n-1)}\sum_{i \in S}\left(\sum_{j \in S \setminus \{i\}} p_{ij}\right)$$

where S is a set of stock return vectors.

Principal Component Analysis

Our PCA algorithm follows that of Cadima Jolliffe [2]. The stock price data can be considered to be a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ where m is the number of *observations* and n is the number of *variables*.

$$\boldsymbol{A} = (\boldsymbol{x}_1 \ \boldsymbol{x}_2 \ldots \boldsymbol{x}_n) \ where \ \boldsymbol{x}_i \in R^{m \times 1} \ for \ i = 1, \ldots, n$$

The original data will consist of m observations of daily returns without dividends on n different stocks. Daily returns without dividends are calculated by

$$x_{ij} = \frac{p_j - p_{j-1}}{p_{j-1}}$$

where p_j is the closing price of the stock on day $j = 2, \dots, m + 1$. We can expect that $Rank(A) \leq n$ since this is a stock price dataset, and variations in stock prices are not perfectly linearly related to the prices of other stocks. The z-score normalized data z_i for each stock is used in place of the original data, which for a given stock x_i is found by

$$z_i = \frac{x_i - \mu_i}{\sigma_i} \text{ where } \mu_i = \frac{1}{m} \sum_{j=1}^{m} x_{ij}, \mu_i = \mu_i \cdot 1^{m \times 1} \text{ and } \sigma_i = \sqrt{\frac{1}{m} \sum_{j=1}^{m} (x_{ij} - \mu_i)^2}$$

where $1^{m \times 1}$ is a vector of size $m \times 1$ with a one in every entry. The value of μ_i for each x_i is stored in a vector $\mu \in \mathbb{R}^{1 \times n}$, which will be used later in Section 2.4. Z-score normalized data has a mean of 0 and a standard deviation of 1, making it easier to extract patterns in the data. We now have a normalized data matrix $\hat{A}$ with new variables $z \in \mathbb{R}^{m \times 1}$ as the columns and new observations $y \in \mathbb{R}^{1 \times n}$ as the rows.

$$\hat{A} = (z_1 \ z_2 \dots z_n) = (y_1 \ y_2 \dots y_m)^T$$

Next, we create a covariance matrix from the normalized data.

$$C = \frac{1}{n-1} \hat{A} \hat{A}^T = \begin{pmatrix} Var(y_1) & Cov(y_1, y_2) & \cdots & Cov(y_1, y_m) \\ Cov(y_2, y_1) & Var(y_2) & \cdots & Cov(y_2, y_m) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(y_m, y_1) & Cov(y_m, y_2) & \cdots & Var(y_m) \end{pmatrix}$$

We decompose this matrix into its eigenvector and eigenvalue pairs.

$$Cv = \lambda v$$

We find n or less orthonormal eigenvector-eigenvalue pairs that describe a new space in $\mathbb{R}^n$. These eigenvectors are referred to as *principal components*. The magnitude of the eigenvalue determines how much variance in the data is accounted for by that eigenvector. The p eigenvectors associated with the p largest eigenvalues that account for 99% of the variance in the data are retained. The selection of 99% variance retention is explained in the results section. Lastly, we project the columns of $\hat{A}$ onto these retained eigenvectors to form a new "weight space", via

$$W = V^T \hat{A}$$

where the columns of V are the p eigenvectors found in the previous step. This means $V \in \mathbb{R}^{m \times p}$

and $\boldsymbol{W} \in \mathbb{R}^{p \times n}$. Each column of $\widehat{\boldsymbol{A}}$ is transformed into this new space that amplifies the distinction between stocks based on their pair-wise correlations. This $\boldsymbol{W}$ matrix is fed forward into the second layer of the neural network, *k-means clustering*.

K-means Clustering

The goal of k-means clustering is to identify $k \in \mathbb{N}$ disjoint partitions in the vector space that group the data points such that each one belongs to only one cluster. In our case, the data points $\boldsymbol{w}$ are the columns of $\boldsymbol{W}$ and the k-means algorithm will partition n data points in the $\mathbb{R}^p$ vector space. The process begins by selecting k data points to be the starting "centroids". Then the following two steps are repeated until the stopping criteria is met:

1. The distance of each data point w to each centroid $\boldsymbol{c}_i$ for $i = 1, \ldots, k$ is computed via $||\boldsymbol{w} - \boldsymbol{c}_i||_2^2$, and each w is assigned to the cluster set S_i of its nearest centroid.
2. A new centroid is selected for each cluster by taking the mean of all vectors in that cluster, $\boldsymbol{c}_i = \frac{1}{|S_i|} \sum_{\boldsymbol{w}_j \in S_i} \boldsymbol{w}_j$

The stopping criteria for these iterations is met when

$$\sum_{i=1}^{k} \sum_{\boldsymbol{w}_j \in S_i} ||\boldsymbol{w}_j - \boldsymbol{c}_i||_2^2 \tag{1}$$

is minimized. This is commonly called the *intracluster distance*. Computationally, iterations are stopped when the intracluster distance changes by less than some threshold between iterations. Challenges that arise from this include determining k that forms meaningful clusters that do not overfit the data. Choosing $k = n$ would give each data point its own cluster, which while reducing the intracluster distance to zero would not provide much information about the similarity of the data points. When $k \ll n$ there is a risk that dissimilar data points will be grouped together. Finding the ideal k value between these two extremes requires careful consideration of both the intracluster distance and the problem being solved.

Forming Portfolios

Each $\boldsymbol{w}$ in the k clusters found in the previous section can be tied back to the stocks in the original data set by the transformation $\boldsymbol{x}_i = \sigma_i \boldsymbol{V} \boldsymbol{w}_i + \mu_i$. Stocks belonging to the same cluster are said to be a part of the same portfolio. Each stock in these portfolios has a very high average linear correlation with all other stocks in the same portfolio. These portfolios also have significantly high MPD. High correlation between two stocks over long periods of time implies that future prices will remain correlated as well. Take Bank of America and Wells Fargo (NYSE: BAC and WFC respectively) as an example.
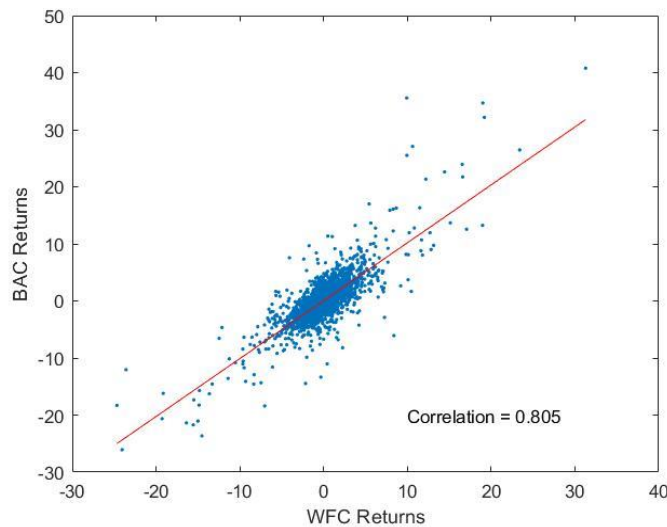
Figure 1: BAC vs WFC

From 2001 to 2017, these two multi-billion-dollar banking companies had stock returns with a linear correlation of 0.805. This is very nearly the perfect correlation of 1. A plot of the two stocks along with the least squares line of best fit is shown in Figure 1. It can be seen that the stock returns from both companies on the same days nearly match the linear fit. Therefore, the fluctuations of BAC's stock price serve as an excellent indicator for similar fluctuations in WFC's stock price, and vice-versa. This logic can be extended to the other stocks in the same portfolio as these two, as each had similarly high correlation to others in the portfolio. Therefore, each stock in the same portfolio has a very high probability to experience similar fluctuations in stock price as others in their portfolio and can therefore be used to predict the prices of those other stocks.

Results

We developed a feed-forward neural network that can identify stocks with very high correlations and similar daily movement. To do this we gathered 100 lucrative companies' stocks and put them through the methodology explained above. In our test sample of 100 stocks, the neural network formed stock groups that have a mean pair-wise correlation of 0.4299 and a MPD of approximately 65.3374% when the number of gathered clusters, k, was 50 and the percent of variance accounted for was 99%.
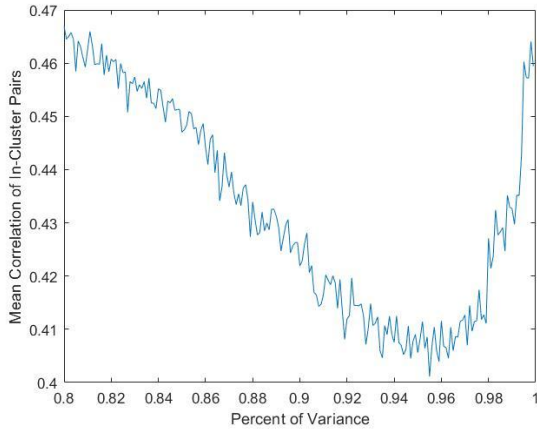
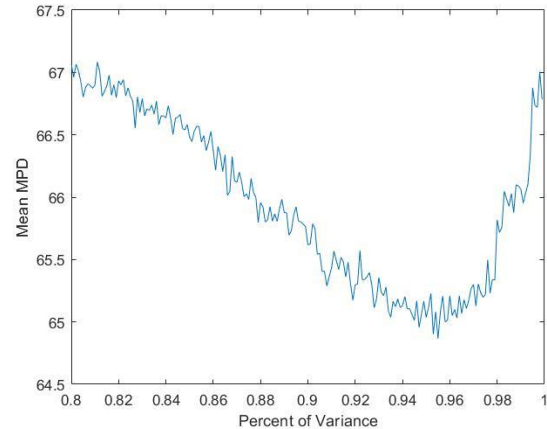Figure 2a: Mean Correlation of Clusters vs
Percent of Variance

Figure 2b: MPD vs Percent of Variance

Figure 2: Observations of Mean Correlation and MPD Averaged Over 100 Trials for Varied Percent of
Variance Explained. $k = 50$

It can be seen in Figure 2 that the mean correlation of in-cluster pairs and MPD both have local maxima at a variance retention of 99%. While retaining 100% of the variance would mean that correlation and MPD are maximized, it would mean that none of the dimensionality reduction from PCA was utilized. One of the goals of this project was to reduce the dimensionality of the stock market for use in k-means clustering, so we selected 99% variance retention for our analysis.
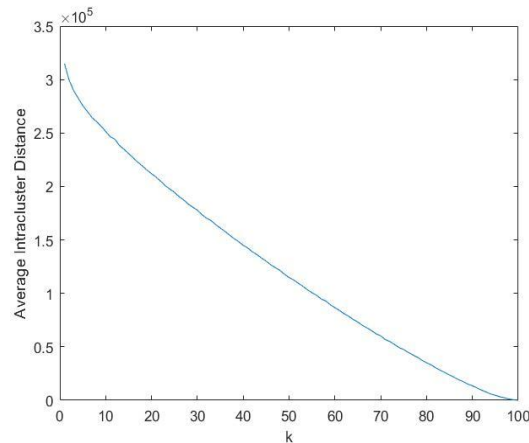

Figure 3: Average Intracluster Distance for Varied k in the 100 Stock Sample, Averaged Over
100 Trials

The value of k was determined using two analyses. Generally, k is selected to be at the point in Figure 3 where the intracluster distance begins to decrease at a decreasing rate. Here, the intracluster distance seems to decrease at a constant rate between $k = 10$ and $k = 90$, so it is difficult to infer the ideal k value from this plot alone. So we analyzed the average correlation and MPD for varied k to see if we could draw any insight from that. We found that larger numbers of clusters produced portfolios with very high correlation and high MPD but contained very few stocks, which does not yield much information to the observer. A graph of correlation vs k can be seen in Figure 4a and a graph of MPD vs k can be seen in Figure 4b. We noticed that correlation and MPD both seem to increase at a constant rate between $k = 10$ and $k = 90$, indicating that no

value of k would produce significantly better portfolios. We noticed that there are several local maxima in both plots, including one at $k = 50$. Taken in the context of the project, $k = 50$ would produce an average of 2 stocks per portfolio, which ensures that on average we can derive some insight about co-movement from every stock in the sample.
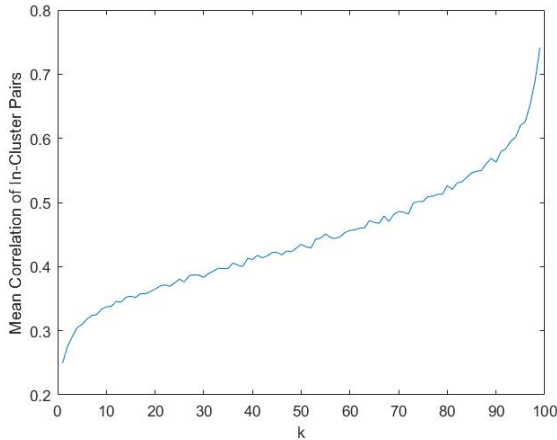

Figure 4a: Mean Correlation of Clusters vs k
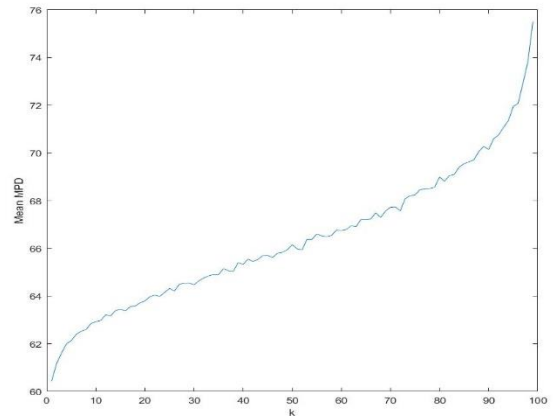
Figure 4b: MPD vs k

Figure 4: Observations of Mean Correlation and MPD Averaged Over 100 Trials for Varied k. Percent of Variance Explained = 99%

Figures 5 and 6 show a pair of stocks that were clustered into the same portfolio using our algorithm with the described values. Anthem Inc. and UnitedHealth Group Inc. (NYSE: ANTM and UNH, respectively) are a pair of lucrative health insurance corporations, so it is reasonable to assume that they would experience similar growth over time. Our algorithm paired these two together, and we do in fact see that the prices appear to be correlated over time, and their daily returns also appear to be heavily correlated.

Figure 5: ANTM and UNH Closing Prices

Figure 6: ANTM and UNH Returns

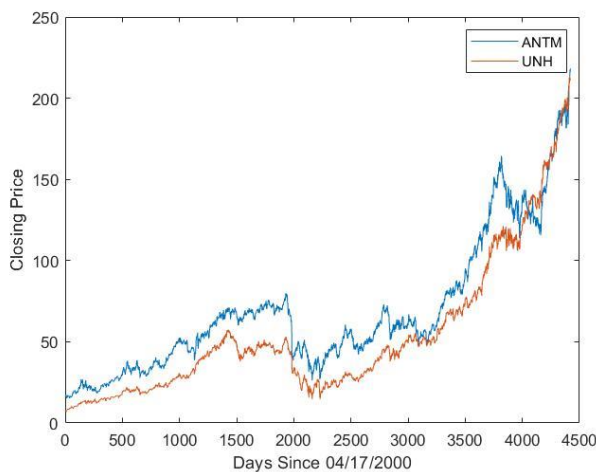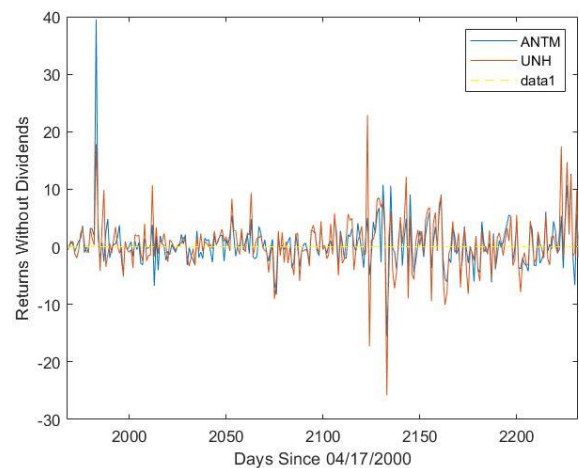Similarly, Figures 7 and 8 show a pair of stocks that were clustered into the same portfolio. However these stocks, AT&T and McDonald's (NYSE: T and MCD, respectively), do not belong to the same business sector of the market. Just as with ANTH and UNH, we see that these two stocks have correlated prices and returns, providing a unique insight into the relationship of these stocks in the stock market.
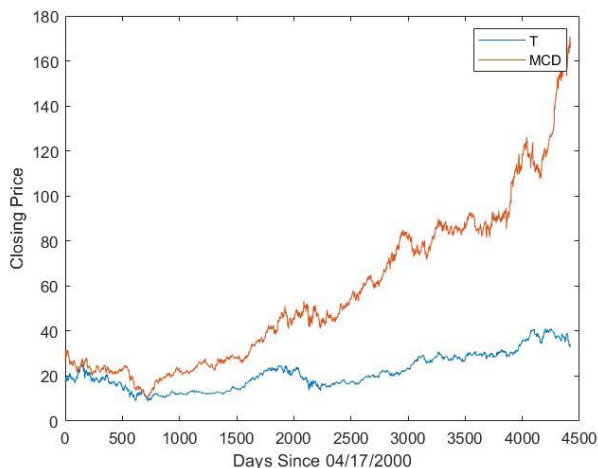
Figure 7: T and MCD Closing Prices
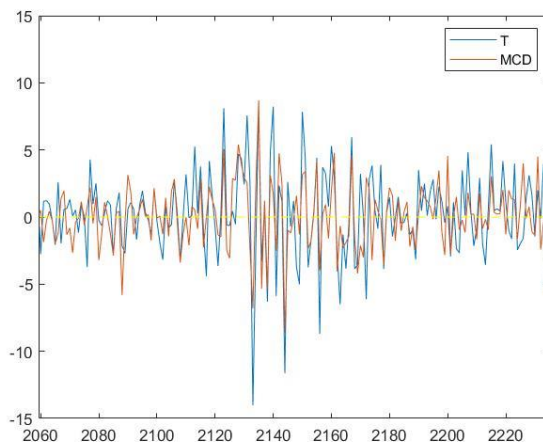
Figure 8: T and MCD Return

Limitations

For the feed-forward neural network to work, the data matrix A must be a complete matrix in $R^{m \times n}$ without any empty entries. Therefore, all stocks must have recorded prices for each day in the sample range. This means that stocks that open or close their public trading during that time period cannot be used for portfolio extraction. Furthermore, the time period on which the analysis is performed should be sufficiently large. Upwards of three years of data should provide actionable results, but longer time periods will be more accurate since linear correlation values are most significant for large numbers of data points. Additionally, a sufficient number of stocks should be used in the original data set. We used 100 stocks for the majority of our experimentation, and we do not recommend using much fewer than that.

Acknowledgements

Appendix

MATLAB Code for Clustering Stock Portfolios

```matlab
clear
rng(1)
%% User Inputs
% Place Path to Folder Here
testfiledir = 'Data 100 Stocks';
% Set minimum number of days that must be present in stock file
min_days = 0;
% Set the start date for analysis (# of days before last day in dataset)
start_date = 260*17;
% Set the end date for analysis (# of days before last day in dataset)
end_date = 260*0;
% Percent of variance to be explained
percent_explained = 0.99;
%% Read in Files
matfiles = dir(fullfile(testfiledir, '*.txt'));
nfiles = length(matfiles);
disp('Total Number of files:')
disp(nfiles)
j = 1;
for i = 1:nfiles
    fid = fullfile(testfiledir, matfiles(i).name);
    M = importdata(fid);
    if isempty(M)
        continue
    end
    if string(M.textdata(end,1)) == '2017-11-10' && size(M.textdata,1) >= min_days
        data{j} = M;
        file_names{j} = matfiles(i).name;
        j = j + 1;
    end
end
nfiles = j - 1;
disp(string(nfiles)+' Files Loaded')

%% Calculate returns and place into a matrix based on date
min_sz = size(data{i}.data, 1);
for i = 1:nfiles
    A = data{i};
    if size(A.data, 1) < min_sz
        min_sz = size(A.data, 1);
    end
end
returns_matrix = [];
if start_date < min_sz && start_date ~= 0
    min_sz = start_date;
end
for i = 1:nfiles
    A = data{i};
    ndays = size(A.data, 1);
    r = ((A.data(1:(ndays-1),4)-A.data(2:ndays,4))./A.data(2:ndays,4)).*100;
    sz = size(r, 1)+1;
    r =  r(sz-min_sz:end,:);
    returns_matrix = [returns_matrix, r];
end
date_diff = start_date - end_date;
returns_matrix = returns_matrix(1:date_diff,:);
%% Perform PCA
```

```matlab
A = returns_matrix;
nobservations = size(A,1);
nvariables = size(A,2);
mu = mean(A,1);
stdevs = std(A);
T = (A - mu)./stdevs;
C = cov(T');
[V,D] = eigs(C,min(nvariables,nobservations));
E = diag(D);
explained = E./sum(E);
tot_explained = 0;
j = 1;
V_red = [];
while tot_explained < percent_explained
    v_temp = V(:,j);
    v = v_temp./norm(v_temp);
    V_red = [V_red, v];
    tot_explained = tot_explained + explained(j);
    j = j + 1;
end

%% Find weight vector transformation
W = V_red'*T;

%% Calculate MPD and correlation for each pair
MPD = zeros(nvariables);
for i = 1:nvariables
    for j = 1:nvariables
        updown_1 = returns_matrix(:,i) > 0;
        updown_2 = returns_matrix(:,j) > 0;
        MPD(i,j) = sum((updown_1 - updown_2) == 0)/nobservations;
    end
end
R = corr(A,A);
%% Kmeans Clustering
k = ceil(nfiles/2);
[labels,~,SUMD] = kmeans(W', k);
k_correlations = [];
stock_names = {};
j = 1;
k_percent_accurate = [];
for i = 1:k
    cluster = find(labels == i)';
    if size(cluster,2) < 2
        continue
    end
    pairs = nchoosek(cluster,2);
    for row = 1:size(pairs,1)
        val = R(pairs(row,1),pairs(row,2));
        k_correlations = [k_correlations; val];
        val = MPD(pairs(row,1),pairs(row,2));
        k_percent_accurate = [k_percent_accurate; val];
    end
end
%% Print Results
k_indexes = {};
for i = 1:k
    cluster = find(labels == i)';
    k_indexes{i} = cluster;
    disp('Stocks in Portfolio '+string(i))
    if size(cluster,2) < 2
        name = split(file_names{cluster(1)}, ".");
        fprintf('%s\n', upper(name{1}))
```

```
        continue
    end
    for index = cluster
        name = split(file_names{index}, ".");
        fprintf('%s\n', upper(name{1}))
    end
end
[~,~,R_wo] = find(R - eye(nfiles));
R_Q = quantile(R_wo, [0.5, 0.75], 'all');
disp('Median and Third Quantile Correlation of All Stock Pairs')
disp(R_Q)
[~,~,P_wo] = find(MPD - eye(nfiles));
P_Q = quantile(P_wo, [0.5, 0.75], 'all')*100;
disp('Median and Third Quantile MPD of All Stock Pairs')
disp(P_Q)
disp('Mean Correlation of In-Cluster Pairs')
disp(mean(k_correlations))
disp('Mean MPD for In-Cluster Pairs')
disp(mean(k_percent_accurate)*100)
```

Sample Output for 100 Stock Sample

Portfolio 1	Portfolio 11	Portfolio 23	GT
GD	GIS	DDS	MET
LMT	PEP	Portfolio 24	SYY
Portfolio 2	PG	NOC	VZ
MU	Portfolio 12	Portfolio 25	WBA
Portfolio 3	MSFT	TJX	Portfolio 34
AA	Portfolio 13	Portfolio 26	HSY
Portfolio 4	CI	UNP	K
ADBE	Portfolio 14	Portfolio 27	KO
CSCO	TXN	AIG	Portfolio 35
IBM	Portfolio 15	Portfolio 28	NKE
INTC	ABC	ADM	Portfolio 36
Portfolio 5	CAH	Portfolio 29	DIS
BIIB	MCK	CAT	Portfolio 37
Portfolio 6	Portfolio 16		JNJ
ANTM	MMC		MRK
UNH	Portfolio 17	Portfolio 30	Portfolio 38
Portfolio 7	GS	COST	SYK
BRK-B	OMC	HD	Portfolio 39
CCL	Portfolio 18	TGT	AAPL
COF	KR	WMT	Portfolio 40
SCHW	Portfolio 19	Portfolio 31	PFE
Portfolio 8	BBY	BA	Portfolio 41
DLTR	Portfolio 20	Portfolio 32	XRX
Portfolio 9	CL	HON	Portfolio 42
F	CLX	Portfolio 33	ORCL
GE	Portfolio 21	ADP	Portfolio 43
Portfolio 10	SWK	AFL	COP
JCP	Portfolio 22	AXP	CVS
JWN	X	CMCSA	CVX

Portfolio 44	T	HPQ	Portfolio 49
DTE	XOM	Portfolio 47	DAL
ECL	Portfolio 45	AMZN	LUV
ITW	BBBY	EXC	Portfolio 50
MAR	KSS	FDX	BAC
MCD	LOW	UPS	C
PPG	ROST	Portfolio 48	JPM
SBUX	Portfolio 46	DE	WFC

References

[1] Forgy, Edward W. (1965). ``Cluster analysis of multivariate data: efficiency versus interpretability of classifications''. Biometrics. 21 (3): 768–769. JSTOR 2528559.

[2] Jolliffe IT, Cadima J. 2016 Principal component analysis: a review and recent developments. *Phil. Trans. R. Soc. A* 374: 20150202. http://dx.doi.org/10.1098/rsta.2015.0202

[3] Marjanovic, Boris 2017. Huge Stock Market Dataset. www.kaggle.com/borismarjanovic/price-volume-data-for-all-us-stocks-etfs

[4] Karl Pearson F.R.S. (1901) LIII. On lines and planes of closest fit to systems of points in space, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2:11, 559-572, DOI: 10.1080/14786440109462720

The Relationship Between Declensional Attributes and Relative Utterance Speed in the Continental Germanic Languages

Ivan S. Vanek

Abstract

This paper analyzes the different attributes of grammar that affect declension and what their physical impact is in relation to relative utterance speed in the continental Germanic languages. Relative utterance speed, in this study, is quantified by the number of syllables that a certain noun-article combination contains. Noun-article combinations in particular were chosen to be the objects of focus in this study because of their ability to be impacted directly by declension, leading to an observable change that can result from declensional modifications in a sentence. Discussion of the history, function, effects of declensional attributes, which are case, gender, number, and article presence, provide background for a study of Swedish and Norwegian comparing the number of syllables in words that have taken up specific combinations of these attributes. Based on prior research and the results of the study, there is not conclusive evidence to suggest a presence or an absence of effect that declensional complexity has on sentence length in the language group. It has been determined, however, that gender-based inflection of animate noun-article combinations does not have an effect on relative utterance speed in Norwegian, and that declensional complexity does not invoke a difference in relative utterance speed in animate noun-article combinations between Swedish and Norwegian.

Keywords[1]: Germanic, declension, inflection, case, gender, number, article.

Introduction

Declension is the inflection of nouns, pronouns, and adjectives in language. Doing this in the context of a sentence frequently results in morphological changes in the words of focus, typically done by the addition of an affix to them. Despite a trivial and somewhat irritating appearance to learners of a declension-heavy language, its use presents the existence of semantic aspects of words that would otherwise have not been realized. Put simply, the use of declension on nouns, pronouns, and adjectives gives them a meaning which is distinct from their non-inflected counterparts. Thus, one can question the effects that the use of declension has on the length of a sentence required to express an idea in language. Because the affixes used act as morphemes themselves, and as a replacement for prepositions or postpositions in certain cases, the relative time of utterance needed to express a particular idea in a declension-heavy language presumably changes in comparison to a less declension-heavy one.

Due to the variation in the amount of declension used between languages within it, the Germanic language family is an appropriate one to focus on for this study, particularly the group of most predominant Germanic languages that currently exist on the continent of Europe. These languages, which include English, German, Dutch, Danish, Norwegian, and Swedish, among others, will be referred to as the continental Germanic languages for the purpose of this study. The title of 'continental Germanic' was assigned to this select group of languages because they

---

[1] Definitions of linguistic keywords:
Declension: inflection of a noun, adjective, or pronoun based on class
Inflection: change made to a word to mark distinctions in case, gender, number, tense, person, mood, or voice
Case: inflectional form of a noun, adjective, or pronoun indicating its grammatical relation to other words
Gender: a grammatical subclass that determines agreement and selection of certain grammatical forms
Number: distinction of a word form to denote reference to one or more than one
Article: words or affixes that are used in conjunction with nouns to limit or give definiteness to its application

are among the most widely spoken Germanic languages that currently exist within the European continent. 'Continental Germanic' does not exclude non-contiguous regions of Europe, which allows for the inclusion of the insular Icelandic and Faroese languages in this study. In order to accurately analyze declension and compare it between languages in the continental Germanic group, a plethora of information related to inflection in its languages and the history behind it is needed before a study can be conducted. The topics of focus are organized into four groups: 1) how declension functions in the languages, 2) how declension developed and changed over time, 3) how grammatical gender interacts with declension, and finally, 4) the effect of declension on different parts of speech among languages within the group.

Function of Declension

The continental Germanic languages use declension in a variety of ways. Although their direct function is similar throughout the language family, the uses of declensional rules vary between each language. This leads to variation among the continental Germanic languages as a whole, and thus, each must be examined independently in order to fully grasp the nuances of each language in this regard.

Inflectional cases

A historical characteristic of the Germanic language family is the utilization of four inflectional cases: the nominative, genitive, accusative, and dative cases. Each case is meant to fulfill a specific function in showing how a noun is applied in a sentence. The existence of these four cases can still be seen in certain modern continental Germanic languages, such as Icelandic. Icelandic utilizes all four of the said inflectional cases, causing a more specific change in nouns, adjectives, and pronouns through the addition of a suffix based on their grammatical function

(Müller, 2005). The effect that the use of the four cases has on the noun at hand can allow a better understanding as to how the noun interacts with its target words. This conveys a more context-specific understood meaning of the noun to the language's speakers.

Although languages, such as Icelandic, have maintained the use of these four noun cases, the roles of inflectional cases in each Germanic language have changed over time, with certain languages experiencing the removal distinguishable characteristics of up to all four of said cases. A clear example of this reduction in the number of inflectional cases through a language's modern development is found in Swedish grammar. There is a striking difference between the number of cases used for the three aforementioned parts of speech seen when comparing Old Swedish and Modern Swedish. While Old Swedish functioned through the use of all four inflectional cases, the nouns and adjectives of Modern Swedish have been modified to function without any inflectional case at all. Alternatively, pronouns of Modern Swedish maintain the use of two inflectional cases, namely the nominative and genitive (Norde, 2002). To a similar effect, Danish has also undergone a reduction in the number of functional inflectional cases it has, shifting only to rely on the nominative and genitive cases for nouns and pronouns (Holmes & Lundskær-Nielsen, 2010). The variation in the number of cases only scratches the surface of the individualistic perspective with which each language must be viewed in this sense.

Irregularities in declension

Although there is a general set of guidelines that languages follow in order to convey ideas in a particular way, such as the addition of an -*s* suffix in English to denote a plural number, there are certain declensional anomalies that exist in each language which, again, must be examined and taken into account. A rather minimal example of this exists in English in the denoting of plural number for nouns. The "vast majority" of nouns in the English language

follow a rather formulaic pattern when denoting their plural forms. However, there also exists an "odd whimsical extension of irregular formation" found in the plural forms of certain words (Blevins, 2006, p. 512). The extension of the typical plural formation being referred to is what occurs in nouns that follow the regular pattern, such as *cat*, *book*, or *river*, as opposed to irregular-patterned nouns, such as *man*, *crisis*, or *ox*. The plurals for the normal pattern words are formulated quite simply, with just the addition of an /s/ phoneme to denote it, leading to the plural form of *cats* from *cat*, *books* from *book*, and *rivers* from *river*. However, in order to denote plurality in nouns that do not follow this rule, a more unique change must be made depending on the specific type of noun. Take the word *man* as an example, which has the plural form *men*. The change made to this specific word to denote plurality is the substitution of the vowel *a* for an *e*. Abnormalities like these are seen in a variety of other forms, such as the word *crises* from *crisis*, or *oxen* from *ox*. Even if a speaker of the English language who is not very proficient in the language were to use an incorrect plural form, such as by saying *mooses* as opposed to the correct plural *moose*, the statement would still be intelligible to most other English speakers. Although English has plenty of irregularities used when denoting plural number, they clearly are not extremely necessary to speakers in conversational or informal contexts. The productive grammatical rule in place in English of using an -*s* suffix to denote plural number is sufficient enough to render the irregular forms communicatively arbitrary, as the productive rule is so broadly standard in English that an assumption can be formed that plural number is being denoted when the suffix is used on a noun that does not follow the standard form.

The irregularities of declension in other Germanic languages in comparison to their counterparts, however, are far more impactful and can cause far more confusion than it is

possible with the case of English plurals. Within the continental Germanic languages, one that undoubtedly stands out the most is the case of Icelandic and its compound nouns. Compound nouns have the ability to form the meaning of a complex noun while existing as only one word. However, the case declension that must be taken into account when forming such nouns is not straightforward by any means. The individual nouns making up a larger compound one can each follow their own declensional pattern. This idea can be found in the word *einkabílastæði* (private parking spot), which is made up of the nouns *einka* (private), *bíla* (car), and *stæði* (space). Both the words *einka* and *bíla* exist in this word in the genitive case, while the word *stæði* exists in the nominative case. Alternatively, there exists another common form of nominal compounding where uninflected stems of compounded nouns are utilized, as opposed to their inflected genitive counterparts, as shown in the aforementioned example. One such occurrence is found in the Icelandic word *þjóðhagfræði* (macroeconomics), with *þjóð* (nation), *hag* (interest), and *fræði* (study) all existing in the nominative case (Harðarson, 2016). This shows the declensional variation that Icelandic nouns can have through this feature of nominal compounding, which is typically uncommon among the continental Germanic languages. As such, the irregularities of compound nouns in Icelandic, as well as other nominally-compounding Germanic languages, must be noted when understanding how its declensions as a whole function. Mistakes in this type of declension can cause far more confusion in conveying the meaning of a word, showing that declensional abnormalities as a whole should be kept in mind when comparing distinct languages within this group.

Changes in Declension

Living languages that are used by a population are bound to experience change in some form, ranging from inflectional cases to vocabulary. Changes can be brought on by a variety of factors that can be either internally or externally motivated.

Internally motivated modification

Internally motivated change to a language can be attributed to a few key elements: practicality and social influence. When a language is modified for the sake of practicality, the changes it undergoes are done to improve efficiency in the expression of ideas, or because certain facets of the language become deemed unnecessary by its speakers. This type of internally motivated change is visible when comparing late, or extinct, Germanic languages, from which the modern continental Germanic languages developed. When comparing Gothic and Old Norse languages to their older predecessors, such as Proto-Indo-European, it is evident that the preexisting necessary suffixes functioning as stem classifiers and inflectional endings had become separable from nouns (Norde, 2002). This example of declensional modification is not one that exhibited a reduction in the number of cases present in the languages, but rather a deflexion by means of reducing the necessary applications of declensional affixes. Because the change was widespread across multiple early Germanic languages, it is plausible that said modification was caused by the lack of importance for inflectional noun markings in the them.

Although practicality-based internal motivation can take place when causing inflectional change, social motivation is far more obvious and relevant to modern language, as it is more consciously caused by a language's speakers. Take the case of gendered inflection as an example. In multiple continental Germanic languages, such as English and Swedish, the use of gendered occupational and title-related suffixes has almost entirely disappeared. Distinctions

between words such as *actor* and *actress* have, in many instances, become viewed as "old fashioned or even clearly derogatory" (Ronneberger-Sibold, 2007, p. 205). This trend continues to this day, as modern German-speaking feminists have also criticized the use of said gendered inflection is their own language. Although some groups will deem it socially unacceptable to use gender-oriented titles as shown, others may feel that they are necessary to be clear in the presentation of ideas in their language, which is what their existence in German can most likely be attributed to. Considering the change that the other aforementioned languages underwent in their removal of extraneous gendered suffixes, it is likely that German will experience a similar adjustment in the future due to the cultural similarities shared by many Western populations.

It cannot be understated that although these different types of internal motivation factors are possible in promoting grammatical modification within a language or family of languages, namely the language group of focus in this research, it is difficult to identify any given one of them as the driving factor for any particular grammatical change, unlike the case of externally motivated modification, where specific sources can be identified as a change-eliciting factors due to commonalities shared between an influencing linguistic source and the language it has influenced. Because of the ambiguity that surrounds the discussion of internally motivated change as a whole, external motivating factors should be regarded with significantly more plausibility in causing modification.

Externally motivated modification

Alternatively, it is also possible for languages to be affected by outside influence, such as through interaction with foreign languages, cultures, and groups. This is evident, once again, when inspecting the changes that have taken place between extinct languages belonging to the Germanic family. The Scandinavian languages, which developed from Old Norse, were

"profoundly influenced by Middle Low German" due to the contact between the Germanic and Scandinavian people (Norde, 2002, p. 243). The loss in grammatical case that the Scandinavian languages experienced can be attributed to this contact because of the reduced elements of grammatical case which already existed in Middle Low German. Additionally, the influence that a language can experience as a result of foreign languages is also evident when analyzing modern languages, such as Faroese. In Faroese, for example, the typical plural number marker for the masculine-gendered nominative case is an *-ar* suffix. However, due to the increasing use of Danish in the Faroese region, the Danish plural marker *-s* has become popularized (Petersen, 2008). Despite being part of the same family, continental Germanic languages are able to impact and change one another, which has led to the inflectional changes that characterize their modern forms.

Interaction between Grammatical Gender and Declension

A characteristic of the continental Germanic languages that must be taken into account when reviewing their declensional features is the existence of grammatical gender within them. Declension interacts with grammatical gender and causes greater diversity in how nouns are affected.

Gender and declension as reciprocal determinants

Oftentimes, gender and declension can act as reciprocal determinants for one another in language. This allows either gender to determine what declension a word must have, or for declension to determine what the gender of a word must be (Kürschner & Nübling, 2011). In the Nynorsk dialect of Norwegian, for example, the use of indefinite articles on nouns with a singular number indicate the gender of a word. *Ein* (a) is used for masculine nouns, *ei* (a) is used

for feminine nouns, and *eit* (a) is used for neuter nouns (Enger, 2004). Therefore, given a phrase such as *ein bil* (a car), one could use the presence of the masculine declension of the indefinite article, *ein*, to determine that the noun *bil* is masculine without having directly known its gender beforehand. This is an example of declension determining the gender of a noun, but, as previously stated, the opposite can take place in language as well. The same is true in reverse, where the gender of the noun of focus can function as a determinant for declension. In knowing the gender of the noun *bil* as masculine, the implicitly resulting indefinite article that should be used is *ein*.

The use of both forms of determination is a visible characteristic of certain other continental Germanic languages, such as German. Take, for instance, the phrase *meine Ente* (my duck). As a feminine noun, the possessive pronoun that precedes it is required to take on an *e* suffix based on its grammatical gender (Martin, 2015). Because feminine possessive pronouns are the only ones that can take on this suffix, declension functions as a determinant for the gender of the noun *Ente*. Again, the same type of reciprocal determination can be seen in this case. Upon seeing the pairing of a possessive pronoun with an *e* suffix and a noun of any type, the automatic assumption can be made that the noun it is attached to has a feminine grammatical gender. As demonstrated, there is a high degree of interaction between declension and grammatical gender in the continental Germanic languages.

Implied gender without effect

Although there is a clear relationship between gender and declension in some continental Germanic languages, it is worth noting that this interaction does not take place in all of them. In fact, gender present within certain languages of the group is able to have no effect on a sentence in terms of declension. This concept is most predominant in modern English. Before the

modernization of Old English, nouns of the English language existed in one of three distinct genders: masculine, feminine, and neuter. However, after the language developed into the modern English, there is no longer a distinct grammatical gender assigned to a word, which is visible in the lack of gendered articles, or different forms of the articles *a* and *the* (Baron, 1971). Thus, it is not possible for variation to exist in the way that grammatical gender interacts with inflected articles.

However, the lack of effect that grammatical gender has is not limited to languages like English that do not use a defined gender system. This same phenomenon also takes place even in languages that continue to make use of the three aforementioned grammatical genders, such as in the case of German where plural number nouns in the masculine and neuter gender can both use the *-e* suffix when denoting its plurality (Kürschner & Nübling, 2011). Therefore, it is possible for the gender of a particular noun to either be masculine or neuter grammatical gender without changing any part of the sentence, as shown in Figure 1 seen below.

Figure 1:

Singular subject → Plural subject

Masculine: Der <u>Hund</u> geht zum Fluss. → Die <u>Hunde</u> gehen zum Fluss.

        (The <u>dog</u> is walking to the river. → The <u>dogs</u> are walking to the river.)

Neuter: Das <u>Pferd</u> geht zum Fluss. → Die <u>Pferde</u> gehen zum Fluss.

        (The <u>horse</u> is walking to the river. → The <u>horses</u> are walking to the river.)

The singular-subject sentences have been shown to clearly indicate that the subject nouns *Hund* and *Pferd* are of different grammatical gender, which is visible through their different articles placed before both words, with *Hund* and *Pferd* having a masculine and neuter gender,

respectively. As seen, the change between masculine gender and neuter gender for the subject words does not change the inflection they undergo when put in their plural forms. This means that if given a comparison between both of the subject nouns at hand in their plural forms, their genders are not distinguishable, displaying the possible absence of effect the gender can have on declension.

Effects on Parts of Speech

It is evident through the previously made comparisons between the continental Germanic languages that each language in the group has declensional features affecting nouns which are not always shared. The purpose of this section is to illustrate how particular parts of speech, namely nouns, pronouns, and adjectives, are affected by declension, as by nature they are more malleable than other parts of speech due to their ability to be influenced by case, gender, number, and article presence the most directly.

Denoting case

Typically, case is denoted in the continental Germanic languages by affixing a suffix to a noun. In languages that utilize a multiple-case system, or one that uses more than just the nominative case for nouns, the number of suffixes that are actually used to denote each case is less than the total number of cases that exist. In Faroese, for example, the declension that a noun takes on depends not only on case, but also which one of the twelve classes, or groups, the noun falls into. This creates forty-eight different possible declensional situations that can be encountered in nouns with singular number. However, despite there being so many different distinct declensional situations, there are only six different suffixes that can be added to singular nouns to represent these situations. In many of these declension possibilities, the noun may not

take on any suffix at all (Enger, 2013). Icelandic and Danish also follow a similar system to the Faroese one. In Icelandic, a rather small set of inflectional markers is used to denote a much greater number of cases (Müller, 2005). Danish is more of a unique case, in that although it does not contain all four Germanic declensional cases, it still maintains the use of fewer inflectional markings than its number of cases. The Danish genitive case makes use of only one marker to denote it, regardless of the gender of the noun (Holmes & Lundskær-Nielsen, 2010). Interestingly enough, the continental Germanic languages as a whole do not follow a one-to-one system for the declensional cases they carry, which refers to one unique affix is used to denote one type of declension. In addition, the languages that carry a higher amount of declensional cases do not make use of a one-to-one system by case either, they use a variety of suffixes for each case and even share suffixes between cases.

Simplification by removal of gender

Declension is also known to be simplified in languages where grammatical gender has been reduced. The resulting grammatical structure of declension becomes less complex, with the motivating factors for this simplification being any of those discussed previously. Grammatical gender of nouns can cause more complexity, which there is evidence of when comparing Old English and Modern English. Old English followed the three-gendered system, as many other Germanic languages do, as illustrated through the presence of agreement between nouns, adjectives, and pronouns. Contrarily, in Modern English, there is no longer a grammatical gender-based agreement between nouns and adjectives. To visualize this difference, anaphoric pronouns serve as an adequate example, where their role has been reduced in that gender is only reflected for the natural gender of animate nouns (Baron, 1971). The use of anaphoric pronouns and their role regarding natural gender and animacy is demonstrated in Figure 2.

Figure 2:

Animate Subject:

Masculine: A <u>man</u> was walking down the avenue, where a car passed <u>him</u>.

Feminine: A <u>woman</u> was walking down the avenue, where a car passed by <u>her</u>.

Neuter: A <u>child</u> was walking down the avenue, where a car passed by <u>it</u>.

Inanimate Subject:

Neuter: A <u>car</u> was parked on the avenue, where someone walked by <u>it</u>.

An anaphoric pronoun is used to refer back to an understood, specified noun. As illustrated in Figure 3, the gender of the anaphoric pronoun is only relevant when it refers to an animate noun that has a natural gender, where the pronoun used will either agree with the masculine, feminine, or neuter gender with *him*, *her*, and *it*, respectively. The absence of a natural gender for a given noun will result in it assuming a case identical to that of a neuter animate noun.

Simplification by the absence of gender can also be seen when specifically taking nouns into consideration. As previously mentioned, occupational nouns in German typically exist with both a masculine and feminine-gendered form, with the addition of a suffix used to denote the feminine-gendered form. However, many of the continental Germanic languages fail to maintain this same trend and have instead abandoned their use for the most part (Ronneberger-Sibold, 2007). There is a clear difference evident when comparing the complexity, in this respect, of languages with each type of system; the ones that do not make use of gendered occupational suffixes are able to convey the same idea using an occupational word in a shorter number of syllables. While other effects that case, gender, and number have on declension are not always visible, this is a very clear example of the absence of gender reducing the complexity of

declension. The extent to which gender is active in the language is arguably just as important as describing the functionality of its declensional system.

Declension and Syllabic Count

In order to most effectively visualize the difference that inflection makes in the length of sentences, and more generally, differences in relative utterance speed in the continental Germanic languages, I carry out an experiment which compares how declension functions in its usage through the analysis of examples of its use. To complete this, all attributes affecting declension must be taken into account and used in the comparison, which are case, gender, number, and article.

To effectively plan this experiment, I select two continental Germanic languages that are closely related in origin, but not so much so that they have the exact same declensional structure. The Swedish and Norwegian languages are chosen for this reason. While being closely related, Swedish uses fewer grammatical genders than Norwegian, as well as fewer cases. Neither language uses extensive noun compounding, meaning translation and identification of the inflected noun-article combinations are simplified.

To compare the effects of declension in each language, I count the number of syllables that a noun-article combination contains in both languages. There are thirty-two unique sentences created in which the noun of focus uses a combination of the four attributes. For example, a sentence created can have its noun of focus in the nominative case, neuter gender, plural number, and with a definite article attached.

The nouns that are used in the study are all animate ones, with the presence of a vowel or consonant ending being taken into account. In each particular gender, the nouns chosen have the

same number of syllables and the same type of letter, either vowel or consonant, with which they end. These measures are taken in order to maintain consistency in the starting form of each noun and allows any difference in the number of syllables between each language to be attributed to declension.

The Swedish nouns used are *pojke* (boy), a common-gender noun, and *barn* (child), a neuter-gender noun. The Norwegian nouns used are *løve* (lion), a masculine-gender noun, *jente* (girl), a feminine-gender noun, and *barn* (child), a neuter-gender noun. These nouns meet the criteria of having the same number of syllables within each gender, as well as ending with the same type of letter within each gender.

Although Swedish only has two genders, common and neuter, the comparison can still be made with Norwegian despite Norwegian having three: masculine, feminine, and neuter. Because the common gender in Swedish is derived from its former masculine/feminine roots, the common gender declensions would be comparable to both the masculine and feminine. To depict this similarity, both the masculine and feminine genders have been included for Norwegian.

The syllable counting process for Swedish translations are the same for both common and neuter gender, as there is only one phrase to count to serve as a comparison to Norwegian. However, because Norwegian has two genders, masculine and feminine, that must be compared to the Swedish common gender, a single value for the average syllable count had to be found in order to effectively make a comparison. Thus, the overall syllable count column of the data table for Norwegian, labelled $\mathbf{NO_{avg}=NO}$, both represents the average syllable count between Norwegian masculine and Norwegian feminine, as well as the direct syllable count from Norwegian neuter. Norwegian neuter phrases are counted for syllables using the same process as

those of Swedish. Once the number of syllables in the noun-article combination is counted for

each phrase, their difference is calculated by subtracting the number of syllables in the Swedish

translation from the number of syllables in the Norwegian translation.

Figure 3 first provides the source sentences as well as translations into both Swedish and

Norwegian, organized by the gender of the noun of focus. Translations were obtained using two

online software in an attempt to achieve as much accuracy as possible through comparing the

resulting translations yielded from both sources. The second table provided shows the data

which resulted from this study. The columns of the second table indicate the sentence translated

through its number corresponding to the first table, case, gender, number, and attached article of

the noun of focus, the number of syllables the translation has in Swedish, followed by those of

Norwegian in masculine, feminine, and average/direct syllable count, and finally the difference

in syllabic count between Norwegian and Swedish, respectively.

Figure 3:

Source Sentence	Swedish Common (SVC)	Norwegian Masculine (NOM)	Norwegian Feminine (NOF)	Swedish Neuter (SVN)	Norwegian Neuter (NON)
1. There is a boy(SVC)/ lion (NOM)/ girl (NOF).	Det finns en pojke.	Det er en løve.	Det er en jente.		
2. The house is a boy's/lion's/girl's.	Huset är en pojkes.	Huset er en løve.	Huset er en jentes.		
3. The deer struck a boy/lion/girl.	Hjorten slog en pojke.	Hjorten slo en løve.	Hjorten slo en jente.		
4. The man showed a boy/lion/girl the book.	Mannen visade en pojke boken.	Mannen viste en løve boken.	Mannen viste en jente boken.		
5. There is a child(SVN, NON).				Det finns ett barn.	Det er et barn.

6. The house is <u>a child's</u>.				Huset är <u>ett barns</u>.	Huset er <u>et barns</u>.
7. The deer struck <u>a child</u>.				Hjorten slog <u>ett barn</u>.	Hjorten slo <u>et barn</u>.
8. The man showed <u>a child</u> the book.				Mannen visade <u>ett barn</u> boken.	Mannen viste <u>et barn</u> boken.
9. There are <u>boys/lions/girls</u>.	Det finns <u>pojkar</u>.	Det er <u>løver</u>.	Det er <u>jenter</u>.		
10. The house is <u>boys'/lions'/girls'</u>.	Huset är <u>pojkars</u>.	Huset er <u>løvers</u>.	Huset er <u>jenters</u>.		
11. The deer struck <u>boys/lions/girls</u>.	Hjorten slog <u>pojkar</u>.	Hjorten slo <u>løver</u>.	Hjorten slo <u>jenter</u>.		
12. The man showed <u>boys/lions/girls</u> the book.	Mannen visade <u>pojkarna</u> boken.	Mannen viste <u>løvene</u> boken.	Mannen viste <u>jentene</u> boken.		
13. There are <u>children</u>.				Det finns <u>barn</u>.	Det er <u>barn</u>.
14. The house is <u>children's</u>.				Huset är <u>barns</u>.	Huset er <u>barnas</u>.
15. The deer struck <u>children</u>.				Hjorten slog <u>barn</u>.	Hjorten slo <u>barn</u>.
16. The man showed <u>children</u> the book.				Mannen visade <u>barnen</u> boken.	Mannen viste <u>barna</u> boken.
17. There is <u>the boy/lion/girl</u>.	Där är <u>pojken</u>.	Det er <u>løven</u>.	Der er <u>jenta</u>.		
18. The house is <u>the boy's/lion's/girl's</u>.	Huset är <u>pojkens</u>.	Huset er <u>løvens</u>.	Huset er <u>jentas</u>.		
19. The deer struck <u>the boy/lion/girl</u>.	Hjorten slog <u>pojken</u>.	Hjorten slo <u>løven</u>.	Hjorten slo <u>jenta</u>.		
20. The man showed <u>the boy/lion/girl</u> the book.	Mannen visade <u>pojken</u> boken.	Mannen viste <u>løven</u> boken.	Mannen viste <u>jenta</u> boken.		
21. There is <u>the child</u>.				Där är <u>barnet</u>.	Det er <u>barnet</u>.

22. The house is the child's.				Huset är barnets.	Huset er barnets.
23. The deer struck the child.				Hjorten slog barnet.	Hjorten slo barnet.
24. The man showed the child the book.				Mannen visade barnet boken.	Mannen viste barnet boken.
25. There are the boys/lions/girls.	Där är pojkarna.	Det er løvene.	Det er jentene.		
26. The house is the boys'/lions'/girls'.	Huset är pojkarnas.	Huset er løvenes.	Huset er jentenes.		
27. The deer struck the boys/lions/girls.	Hjorten slog pojkarna.	Hjorten slo løvene.	Hjorten slo jentene.		
28. The man showed the boys/lions/girls the book.	Mannen visade pojkarna boken.	Mannen viste løvene boken.	Mannen viste jentene boken.		
29. There are the children.				Där är barnen.	Det er barna.
30. The house is the children's.				Huset är barnens.	Huset er barnas.
31. The deer struck the children.				Hjorten slog barnen.	Hjorten slo barna.
32. The man showed the children the book.				Mannen visade barnen boken.	Mannen viste barna boken.

Sentence No.	Case	Gender	Number	Article	SV syl	NOM syl	NOF syl	$NO_{avg}=$ NO syl	(NO-SV)
1.	Nom	C/M/F	sing	indef	3	3	3	3	0
2.	Gen	C/M/F	sing	indef	3	3	3	3	0
3.	Acc	C/M/F	sing	indef	3	3	3	3	0
4.	Dat	C/M/F	sing	indef	3	3	3	3	0
5.	Nom	N	sing	indef	2			2	0
6.	Gen	N	sing	indef	2			2	0

#	Case	Gender	Number	Article					
7.	Acc	N	sing	indef	2			2	0
8.	Dat	N	sing	indef	2			2	0
9.	Nom	C/M/F	plu	indef	2	2	2	2	0
10.	Gen	C/M/F	plu	indef	2	2	2	2	0
11.	Acc	C/M/F	plu	indef	2	2	2	2	0
12.	Dat	C/M/F	plu	indef	3	3	3	3	0
13.	Nom	N	plu	indef	1			1	0
14.	Gen	N	plu	indef	1			2	1
15.	Acc	N	plu	indef	1			1	0
16.	Dat	N	plu	indef	2			2	0
17.	Nom	C/M/F	sing	def	2	2	2	2	0
18.	Gen	C/M/F	sing	def	2	2	2	2	0
19.	Acc	C/M/F	sing	def	2	2	2	2	0
20.	Dat	C/M/F	sing	def	2	2	2	2	0
21.	Nom	N	sing	def	2			2	0
22.	Gen	N	sing	def	2			2	0
23.	Acc	N	sing	def	2			2	0
24.	Dat	N	sing	def	2			2	0
25.	Nom	C/M/F	plu	def	3	3	3	3	0
26.	Gen	C/M/F	plu	def	3	3	3	3	0
27.	Acc	C/M/F	plu	def	3	3	3	3	0
28.	Dat	C/M/F	plu	def	3	3	3	3	0
29.	Nom	N	plu	def	2			2	0
30.	Gen	N	plu	def	2			2	0
31.	Acc	N	plu	def	2			2	0
32.	Dat	N	plu	def	2			2	0

*The colors in this diagram are used to denote the different changes that the noun can take on within each category (case, gender, number, and article type).

Abbreviation Key:

Case	Gender	Number	Article Type	Language
Nom - Nominative	C - Common	sing - Singular	indef - Indefinite	SV - Swedish
Gen - Genitive	M - Masculine	plu - Plural	def - Definite	NO - Norwegian
Acc - Accusative	F - Feminine			
Dat - Dative	N - Neuter			

Discussion

As seen by the data, there is only one point throughout the entirety of the experiment that shows a nonzero difference between the syllabic count of the noun-article combination between Swedish and Norwegian. However, there is a point to be made about that sentence, number 14, in particular as well as sentence 10 regarding the validity of their data.

Certain source sentences sound extremely awkward in that the noun and article would never be used in the context of a conversation or writing in English. This is because English entirely lacks a plural indefinite article, causing certain combinations of case, number, and article types to be impossible to express. This combination in particular is the genitive/plural/indefinite one. As a result, only the closest approximation to what this certain combination should be was made, which were those of sentences 10 and 14. The consequences of this are seen in the resulting translations for Swedish and Norwegian.

Having used online translation software as the means of obtaining translations from the source sentences, there is unfortunately room for error in using untranslatable source sentences, which is an issue that is seen particularly with sentences 10 and 14. Although the closest possible approximation was made for what the case/number/article combination should be in the

source sentence, the software is simply unable to translate it correctly due to the fact that there is no way to properly express it in English. In theory, the software used should be able to identify this combination correctly and consequently provide an accurate translation based on these criteria. However, because the source sentences are nevertheless properly inexpressible in English, it is not certain that translation of this type can take place with any strong degree of accuracy.

The consequences of this lack of expressibility are particularly noticeable in translating sentence 14. Although the phrase in sentence 10 seemed to translate consistently between both languages, the same is not true for sentence 14. The Swedish and Norwegian noun-article combinations take on strikingly different, inconsistent declensions, with that of Norwegian instead being the exact same as the one yielded in sentence 30. Therefore, it can be inferred that there is most likely an incorrect translation for Norwegian in sentence 14, meaning that the data it provided should not be regarded as worthy of consideration. Furthermore, because of the evident risk being taken in using logically inexpressible source sentences, data resulting from the translations yielded from sentence 10 should also be disregarded.

As seen by the usable data obtained, at no point is there a difference between the number of syllables that a specific Swedish or Norwegian declension makes. Throughout all usable examples, the number of syllables of the noun-article combination showed no difference when comparing the two languages. Although both languages see changes in syllable count by means of manipulating the four attributes, the changes they experience are consistent between them.

Conclusion

In analyzing the data that resulted from the study as well as prior research, conclusions can be drawn about the relationship that declension has to relative utterance speed within Swedish and Norwegian, as well as inferences regarding possible similar trends that may be present in other languages of the continental Germanic group.

Regarding solely the information derived from the study conducted, it is clear that there is no relationship between declensional complexity for nouns and phrase length within the languages compared. Despite Swedish having a declensional system far less complex than that of Norwegian, the number of syllables produced for any given declensional combination for the noun-article combinations was identical between the two languages in the viable phrases used for comparison. The data suggests that the presence of additional cases and grammatical genders present in Norwegian seem to have no effect on the declensions of nouns by means of causing extraneous addition of syllables.

Additionally, there is more to be said about the relationship between declensional complexity and grammatical gender based on these results. Even within Norwegian, there was no change in syllabic count that occurred based on gendered inflection whatsoever, such as in the case of changing between masculine and feminine gender. A similar inference regarding the language group as a whole can be drawn in considering preliminary research conducted as well. Referring back to Figure 1, the trend in German for the examples shown is consistent with the results of the study in comparing Norwegian, where two noun-article combinations that only differ in their grammatical gender have no difference in their syllabic count in the case of masculine and feminine. This suggests that it is a trait that can is consistently found in the continental Germanic languages. Thus, there is reason to believe that gender-based inflection

does not result in syllabic count differences, and more broadly, in sentence length differences, in the continental Germanic languages. Although there may be truth to this inference based on the results of the study, it however is not possible to accurately apply this to languages other than Swedish or Norwegian.

Inferences made regarding declensional cases as a whole in this language group should be expanded on in the future before being broadly applied to it, where further studies can provide a more conclusive assertion on the matter. Such studies may include research involving the use of inanimate nouns instead of solely animate ones, referencing different parts of speech that are inflected using declension, and even through the inclusion of more languages. There is, however, sufficient evidence to believe that gender-based inflectional complexity for animate noun-article combinations in Norwegian has no effect on syllabic count, and therefore, no effect on phrase length and therefore relative utterance speed. Pertaining to the study conducted in a broader sense, it has also been determined that declensional complexity does not invoke a difference in syllabic count and thus relative utterance speed in animate noun-article combinations between Swedish and Norwegian.

References

Baron, N. S. (1971). A reanalysis of english grammatical gender. *Lingua*, *(27)* pp. 113–140. doi: 10.1016/0024-3841(71)90082-9.

Blevins, J.P. (2006). English inflection and derivation. In B. Aarts & A. M. McMahon (Ed.), *The Handbook of English Linguistics (*pp. 507-536). Hoboken, NJ: Wiley-Blackwell.

Enger, H. (2004). On the relation between gender and declension: A diachronic perspective from Norwegian. *Studies in Language 28(1),* pp. 51-82. doi: 10.1075/sl.28.1.03eng

Google. Google Translate. https://translate.google.com/.

Harðarson, G. R. (2016). Peeling away the layers of the onion: On layers, inflection and domains in Icelandic compounds. *The Journal of Comparative Germanic Linguistics, 19(1),* pp. 1–47. doi:10.1007/s10828-016-9078-5.

Holmes, P. & Lundskær-Nielsen, T. (2010). Nouns. In T. Lundskær-Nielsen and P. Holmes (Eds.), *Danish: A comprehensive grammar (2nd ed.*, pp. 1-80). Milton, England: Taylor & Francis Group.

Kürschner, S., & Nübling, D. (2011). The interaction of gender and declension in Germanic languages. *Folia Linguistica, 45(2).* doi:10.1515/flin.2011.014.

Martin, H. (2015). Articles, simple past. In A. Ng (Ed.), *A Foundation Course in Reading German* (pp. 23-28). Madison, Wisconsin: University of Wisconsin – Madison.

Merriam-Webster. (2021). Article. In *Merriam-Webster*. Retrieved from: https://www.merriam-webster.com/dictionary/article.

Merriam-Webster. (2021). Case. In *Merriam-Webster*. Retrieved from: https://www.merriam-webster.com/dictionary/case.

Merriam-Webster. (2021). Declension. In *Merriam-Webster*. Retrieved from: https://www.merriam-webster.com/dictionary/declension.

Merriam-Webster. (2021). Gender. In *Merriam-Webster*. Retrieved from: https://www.merriam-webster.com/dictionary/gender.

Merriam-Webster. (2021). Inflection. In *Merriam-Webster*. Retrieved from: https://www.merriam-webster.com/dictionary/inflection.

Merriam-Webster. (2021). Number. In *Merriam-Webster*. Retrieved from: https://www.merriam-webster.com/dictionary/number.

Müller, G. (2005). Syncretism and iconicity in Icelandic noun declensions: A distributed morphology approach. *Yearbook of Morphology*, pp. 229–271. doi:10.1007/1-4020-2900-4_8.

Norde, M. (2002). The loss of lexical case in Swedish. *Grammatical relations in change*, pp. 241-272. doi: 10.1075/slcs.56.11nor.

Petersen, H. P. (2008). The Borrowing Scale in Danish and Faroese. *Fróðskaparrit 56,* pp. 97-115. Retrieved from https://www.researchgate.net/profile/Hjalmar_Petersen/publication/250940158_The_Borrowing_Scale_and_Danish_in_Faroese/links/02e7e51ee88d641e01000000.pdf.

Ronneberger-Sibold, E. (2007). Typologically motivated over- vs. underspecification of gender in Germanic languages. *Language Typology and Universals, 60(3),* pp. 205–218. doi:10.1524/stuf.2007.60.3.205.

Translate free and online. Translate free and online | TRANSLATOR.EU. https://www.translator.eu/english/.