HEC MONTRÉAL

The Covid-19 Vaccine Production Supply Chain - A Data Science Perspective

Par
Aïchata Souleymane Koné

Valérie Bélanger
HEC Montréal
Directrice de recherche

Thierry Warin
HEC Montréal
Codirecteur de recherche

Sciences de la gestion
(Spécialisation Global Supply Chain Management)

Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences en gestion
(M. Sc.)

August 2023

Résumé

La pandémie de Covid-19 a démontré l'importance des chaînes d'approvisionnement mondiales, en particulier la chaîne d'approvisionnement en vaccins de la Covid-19. La réponse à la pandémie a nécessité le développement, la production et la distribution rapides de vaccins, faisant de la chaîne d'approvisionnement du vaccin de la Covid-19 la plus importante de ces dernières années. Étant donné la rareté des recherches sur la chaîne d'approvisionnement en vaccins pandémiques et le risque d'autres pandémies, il est important d'en apprendre le plus possible sur la chaîne d'approvisionnement en vaccins Covid-19. Cet article vise à contribuer à la littérature en proposant trois approches quantitatives pour analyser la chaîne d'approvisionnement du vaccin Covid-19; 1) catégoriser la chaîne d'approvisionnement du vaccin de la Covid-19 en utilisant la classification géographique des entreprises internationales, 2) appliquer une analyse économétrique des données de panel pour identifier les facteurs pertinents de la chaîne d'approvisionnement qui influencent la distribution des vaccins, et 3) effectuer une analyse de réseau de la chaîne d'approvisionnement du vaccin de la covid-19 de Pfizer et Moderna pour identifier les principaux nœuds, liens et goulets d'étranglement dans la production de vaccins. Dans l'ensemble, la chaîne d'approvisionnement du vaccin Covid-19 est le point de départ de l'amélioration des stratégies futures pour des chaînes d'approvisionnement adaptables, équitables et résilientes.

Mots-clés : Covid-19, Chaîne d'approvisionnement du vaccin
Méthodes de recherche : Recherche quantitative

Abstract

The Covid-19 pandemic demonstrated the importance of global supply chains, particularly the Covid-19 vaccine supply chain. The response to the pandemic required the rapid development, production and distribution of vaccines, making the Covid-19 vaccine supply chain the most important in recent years. Given the scarce research on the pandemic vaccine supply chain and the risk of further pandemics, it is important to learn as much as possible about the Covid-19 vaccine supply chain. This article aims to contribute to the literature by proposing three quantitative approaches to analyze the Covid-19 vaccine supply chain. 1) categorize the Covid-19 VSC using international business geographic classification, 2) apply econometric panel data analysis to identify relevant supply chain factors influencing vaccine distribution, and 3) perform a network analysis of Pfizer and Moderna's Covid-19 vaccine supply chain to identify key nodes, links and bottlenecks in vaccine production. Overall, the Covid-19 vaccine supply chain is the starting point for improving future strategies for adaptable, equitable and resilient supply chains.

Keywords: Covid-19, Vaccine supply chain
Research methods: Quantitative Research

Table of Contents

List of Tables

List of Figures

List of abbreviations and acronyms

AR: Autoregressive
ARIMA: Autoregressive Integrated Moving Average
CDC: Centers for Disease Control and Prevention
FDI: Foreign Direct Investment
GMM: Generalized method of moments
GOI: Geocentric Orientation Index
II: Internationalization Index
IV: Instrumental Variable
MA: Moving average
MSC: Multiple supply chains
PAHO: Pan American Health Organization
R&D: Research and development
SNA: Social network analysis
TNI: Transnationality Index
UNCTAD: United Nations Conference on Trade and Development
VSC: Vaccine supply chain
WHO: World health organization

Acknowledgment

I would like to take this opportunity to express my heartfelt gratitude to all those who have supported and guided me throughout this journey of completing my thesis. Their unwavering encouragement, assistance, and understanding have been instrumental in making this work possible.

First and foremost, I extend my deepest appreciation to my esteemed thesis advisors, Valérie Bélanger & Thierry Warin, for their invaluable guidance, insightful feedback, and patient mentorship. Their expertise and dedication have significantly contributed to shaping this research and refining my academic abilities. I am truly grateful for the time and effort they have invested in me.

A special note of thanks goes to my friends and family, especially my father Souleymane Kone, my mother Kandji Sidibe and my grandmother Tenin Konate, who have been pillars of strength throughout this academic pursuit. Their unwavering belief in me, encouragement, and love have given me the motivation to overcome challenges and persevere. I also would like to thank my brother Lamine Souleymane Kone and my sister Mariame Souleymane Kone for always being present when I needed them and providing encouraging words during difficult times.

Additionally, I am grateful for the resources and facilities provided by HEC Montreal that facilitated the smooth execution of this study. Lastly, I want to acknowledge the influence of past researchers, scholars, and authors whose works have been essential in shaping the theoretical framework of this thesis.

Completing this thesis has been a challenging yet fulfilling experience, and I am grateful to all those who have played a part in its completion. Their contributions have enriched the quality of this work and have set me on a path of continuous learning and growth, particularly in times of pandemic which was one of the fears during the Covid-19 pandemic. By understanding the intricacies of the supply chain, policymakers can design strategies that better address the challenges of vaccine availability and accessibility. This ensures that vaccines reach every corner of the globe, contributing to the collective effort to achieve global herd immunity.

This project was based on the Covid-19 vaccine supply chain as of June 2021. Therefore, we suggest for future research to access the Covid-19 vaccine supply chain over time which will give insight into the evolution of a pandemic vaccine supply which has not been researched to this day.

Introduction

The 2020 COVID-19 pandemic brought the world to a halt, posing unprecedented challenges to global public health. In the race to combat the virus, the development and distribution of effective vaccines emerged as a crucial strategy. The successful deployment of COVID-19 vaccines depended not only on their development, but also on the complex network of the vaccine supply chain. This supply chain encompasses the complex process of manufacturing, distributing, and administering vaccines on a global scale. Understanding the COVID-19 vaccine supply chain is essential to ensure equitable access to vaccines and to meet the challenges posed by the pandemic.

The first phase of the Covid-19 vaccine supply chain is vaccine development and vaccine manufacturing. Numerous pharmaceutical companies, research institutes and biotech firms around the world have developed COVID-19 vaccines, each using distinct production processes. These processes typically involve large-scale cultivation of the components required for the vaccine, such as viral vectors or mRNA molecules. Once manufactured, the vaccines undergo stringent quality control measures to guarantee their safety and efficacy (clinical trials. Once the vaccines have been manufactured, the next critical step is their distribution. Given the urgency of the global immunization campaign, it is vital to ensure rapid and efficient vaccine delivery. The distribution process involves many stakeholders, including vaccine manufacturers, governments, international organizations, logistics service providers and healthcare facilities. Vaccines must be transported under controlled conditions, in compliance with specific temperature requirements, to preserve their efficacy. Cold chain logistics, involving refrigerated storage and transport, play a crucial role in maintaining the integrity of vaccines as they travel from production facilities to vaccination sites around the world.

To meet the worldwide demand for vaccines, various actors, such as pharmaceutical companies, health organizations, universities, and governments, worked collaboratively to set into motion the most important value chain of recent years. The Covid-19 vaccine supply chain (VSC was built amid a global pandemic and under very tight deadlines being one of the first VSC to be implemented under such conditions. With the risk of more and more pandemics, it is important to learn from the Covid-19 pandemic and analyze the decisions made concerning its supply chain. Thus, the thesis of this paper: What can we learn about the emerging vaccine supply chain in the wake of a global pandemic using data science? To answer this question, we develop three sub-question that will be presented in the form of articles as shown in *Figure 0-1*.

Article 1, titled “International Business Classification Approach to the Covid-19 Vaccine Production Supply Chain - A Data Science Perspective”, uses international business geographic classification to categorize Covid-19 vaccine supply chains. This first article recreates Rugman & Verbeke’s firms’ classification (2004 and applies it to six VCS. The categorization was conducted in two dimensions, one for the upstream geographic location of the supply chain and one for the downstream geographic distribution of the vaccine. The paper aimed to classify the six Covid-19 vaccine producers in terms of their internationalization and the associated risk of such strategies.

In article 2 of this project, titled “Econometric Approach to the Covid-19 Vaccine Production Supply Chain”, we want to identify which supply chain factors influence the distribution of two main Covid-19 vaccines, Pfizer and Moderna. To this effect, we used an econometric panel data analysis to extract relevant supply chain factors influencing vaccine distribution. This paper aims to look at vaccine distribution in countries from a supply chain perspective instead of indicators previously used, such as

demographic factors, economic factors and more. This would then give policy and decision makers a new way to see and manage vaccine distribution.

The last article is titled “Network Analysis Approach to the Covid-19 Vaccine Production Supply Chain - A Data Science Perspective”. The objective of this study is to conduct a network analysis of the Covid-19 vaccine supply chain of Pfizer and Moderna to identify the key nodes, links, and bottlenecks in the production of vaccines. This study will use a qualitative approach. The main purpose of this paper is to contribute by exploring innovative ways to analyze and improve the vaccine supply.

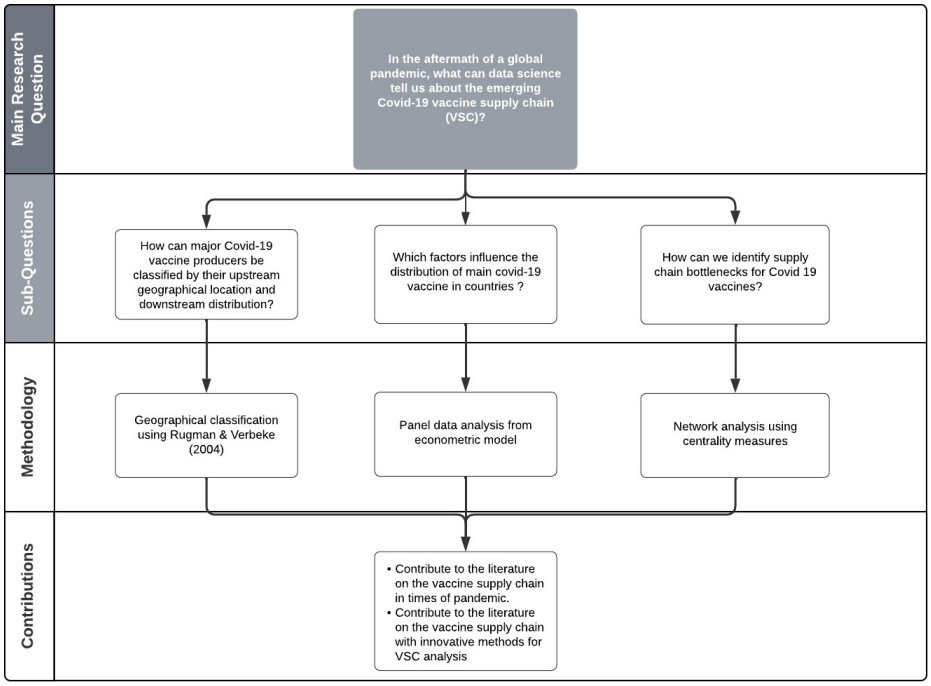


Figure 0-1: summarizes the research questions of this thesis.

The purpose of this project is to contribute to the literature on the vaccine supply chain. As the VSC literature in times of pandemic is lacking, we aim through this thesis to provide new direction to research on pandemic vaccine supply chains. Thus, opening the door for further research on the topic for future potential pandemics.

Chapter 1: International Business Classification Approach to the Covid-19 Vaccine Production Supply Chain - A Data Science Perspective

Koné Aïchata Souleymane, M.Sc. student, HEC
Montreal Supervision:
Bélanger Valérie, HEC Montreal
Warin Thierry, HEC Montreal

HEC Montréal

Abstract

The global response to the COVID -19 pandemic has required the rapid development, production, and distribution of vaccines on an unprecedented scale. Ensuring efficient and effective management of the COVID -19 vaccine supply chain is critical to successful vaccination campaigns worldwide. This paper proposes a novel approach to classifying the COVID -19 vaccine supply chain using internationalization classification techniques. This project aims to integrate the principles of international business classification, particularly geographic diversification, into the upstream and downstream dimensions of the Covid-19 vaccine supply chain by emulating the methodology of Rugman & Verbeke (2004.

1.1 Introduction

The first case of Covid-19 was reported on December 31, 2019, in the province of Hubei, China, and as of March 11, 2020, the World Health Organization (WHO declared the Covid-19 a global pandemic. Rapidly, the need for vaccines emerged and organizations around the world came together to develop a vaccine. With a vaccine developed in under a year, there was a rapid need to produce and deliver a large quantity of vaccine as possible. Thus, the emergence of the Covid-19 vaccine supply chain was one of the most important value chains during the pandemic.

However, vaccine development is highly regulated. Due to the unique nature of the vaccine supply chain (VSC, production planning and decision-making are under major time constraints combined with the thermal sensitivity of vaccines makes designing a VSC difficult.

A vaccine is the most effective defense against pathogen caused disease. Pathogens are organisms, such as viruses and bacteria, that cause health issues to the host it enters. Vaccines are designed to train the immune system in a controlled manner. Most vaccines contain inactive antigens of a virus that the immune system, mainly the white blood cell, whose role is to learn, then recognize and create antibodies to counter that pathogen. By doing so, the immune system is then ready for the eventuality of encountering that pathogen. The vaccine is generally administrated to healthy people in a preventive, before an outbreak (i.e., seasonal influenza), or reactive manner, during an outbreak (i.e. Covid-19). Vaccines are different from any other product; therefore, their supply chain also has some unique characteristics. Duijzer et al. (2018) identified the following characteristics that are unique to the VSC:

- **High uncertainty**: The vaccine supply chain has uncertainty in both the availability of supply and demand forecasting, which makes vaccine production difficult to plan.

- **Misalignment of objectives**: There is a misalignment between the producer and the buyer. Indeed, vaccine manufacturers are for profit organisms while vaccine buyers are non-profit organizations, such as the government and the WHO. This misalignment creates divergence in objective.
- **Decentralized decision-making**: With multiple suppliers involved in the VSC the decision-making process is decentralized along the chain. This makes coordination a challenge in the chain.
- **Difference between buyers and users**: The buyer of vaccine are public health organizations and governments, while the end user of vaccine is the general population. This difference makes vaccination strategies complex, as end users refuse to be vaccinated. This makes it harder to estimate demand.
- **Complex political decisions concerning allocation**: Decisions regarding the first people to receive the vaccine and the number of doses allocated per country are the most important decision once the vaccine is produced that governments and public health organizations must manage.
- **Importance of deciding and acting in time**: Vaccine production as well as its distribution are under time constraints. For some vaccines, such as influenza, it is possible to start production before the dominant strand is identified. This help to produce more vaccine however the risk is that they are not sure if the vaccine produced is effective that the strand of that season.

Figure 1-1 represents the flow of the vaccine Supply chain from development to the end user. The vaccine supply chain is divided into three main phases: vaccine origination and development, vaccine manufacturing, and delivery *(*Brown & Bollyky, 2021*)*. All VSC start with vaccine development. This phase consists of two main processes. First, a phase of research and development (R&D) which is necessary to create a vaccine. Once a formulation has been completed, it goes through multiple clinical trials, which cannot be bypassed. Each country has their vaccine regulations; therefore, one clinical trial can be approved in one country and not in another.

Vaccines that pass all the trial moves to the second phase after being licensed, while the others are sent back for further development. During the vaccine manufacturing phase, the licensed product is sent to be produced in large quantities (substance production process), then the vaccine is packaged in glass vials, sealed off and ready for delivery (fill and finish process). Both stages of vaccine manufacturing require different inputs. The substance production requires capital equipment, raw and single-use materials, and other pharmaceutical ingredients. The capital equipment includes bioreactors, pumps, and filtration units, whereas the raw and single-use materials include bioreactor bags, cellular materials, and filters. Other pharmaceutical ingredients are adjuvants, lipids, preservatives, and excipients. The fill and finishing process requires vial-filling equipment, glass vials, stoppers, and refrigeration. The last phase of the VSC is the delivery/distribution which requires inputs such as needles, syringes, antiseptic wipes, and diluents.

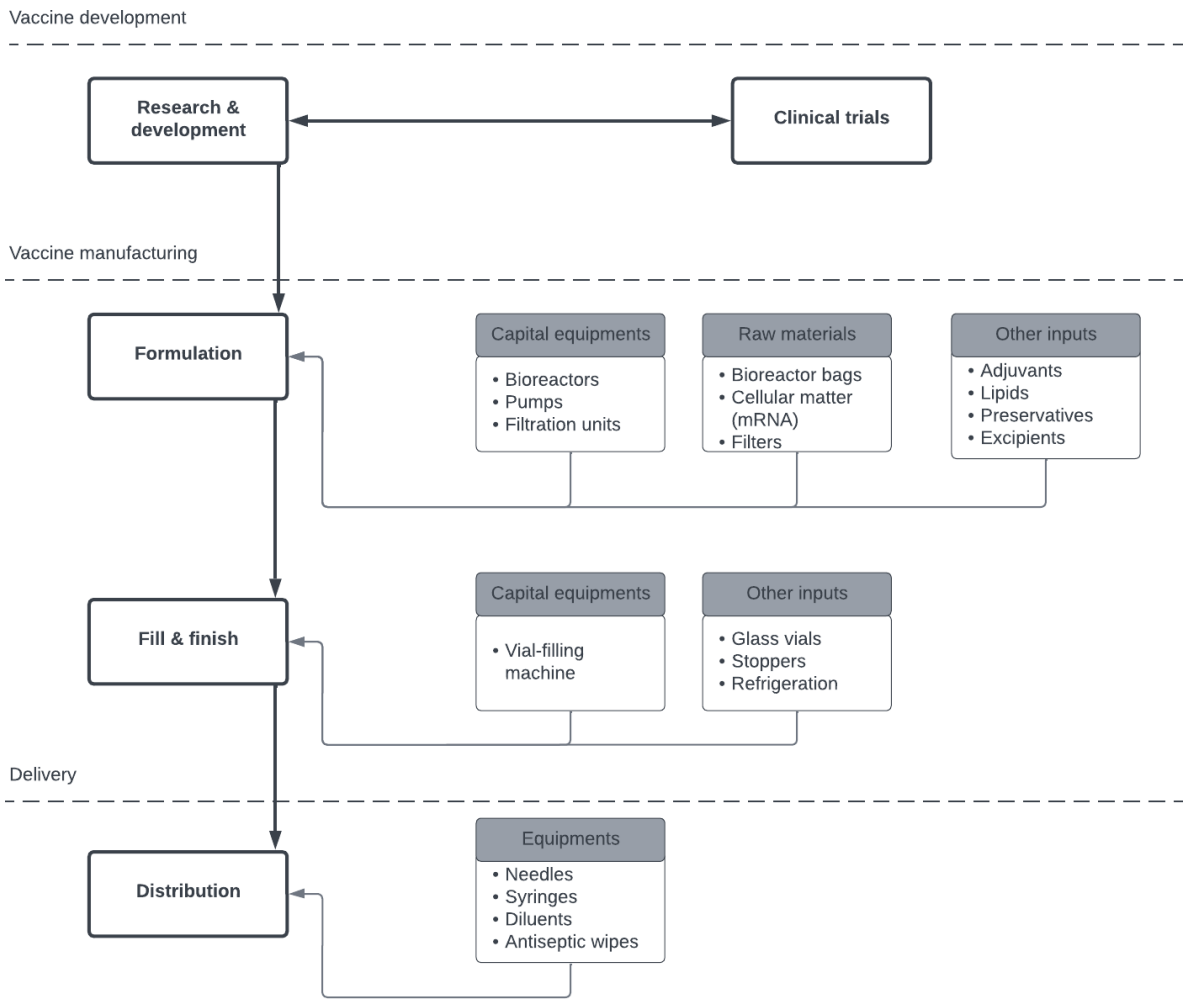


Figure 1-1: Vaccine Supply Chain Flow

As mentioned previously, vaccine development takes time. Thus, making this stage the most time-consuming step of the VSC, most vaccine takes years before they are approved to be on the market. The Covid-19 vaccine being developed in under a year, made the public question the efficacy and safety of the vaccine. However, the short period of development is not necessarily an indication of shortcuts in development. Indeed, global resources, through collaboration, were allocated to develop a vaccine, during the early stage of the pandemic. This was the main reason why the Covid-19 vaccine was among one of the fastest vaccines to be developed. The global effort included universities, pharmaceutical companies, worldwide organizations, and governments, each contributing in their own ways.

In this collective effort, universities played a major role with their human and capital resources, in the development stage. They were able to map the genome of the virus in a timely manner which is a crucial step when developing a vaccine. They also provided additional support during the human clinical trials. With their experience in the development and manufacturing of pharmaceutical products, pharmaceutical companies such as Pfizer, were crucial in providing manufacturing support but also financial support in some cases. Worldwide organization, such as the WHO, provided their expertise in epidemiology and vaccine regulation. During the development stage, the WHO created and provided a

Covid-19 vaccine blueprint, which is the requirements the vaccine had to follow (WHO, 2020). Furthermore, through their epidemiology expertise and their experience with previous epidemic management, they proposed a vaccination strategy to flatten the epidemic curves. This is as important as the vaccine itself if not more, poor vaccination strategies fail to mitigate the spread of a viral disease. The last major player in the Covid-19 vaccine development was the government. Indeed, governments provided mainly financial support to both universities and pharmaceutical companies to aid them in moving closer to a marketable vaccine. This may be due to the urgency of needing a vaccine to protect their population.

As a result, of this collective effort, 200 vaccine candidates were proposed by December 2020 with only 52 approved for human clinical trials (WHO,2021). To get a vaccine, new and old approaches were used, making in total 3 types of Covid-19 vaccine. The whole-microbe approach uses the whole SARS virus either active or not as the base of the vaccine. The sub-unit approach, on the other hand, uses a specific part of the virus that is unique to its species, i.e. protein or sugars. In the case of the Covid-19 vaccine, some vaccines opted for the sub-unit approach using a protein based. The last and newer approach is the genetic approach. This vaccine uses isolated genes that code for specific proteins for the immune system to recognize.

When designing a supply chain for the Covid-19 vaccine, many pharmaceutical companies opted for multiple supply chains (MSC) to ramp up their production capacity. MSC occurs when a company creates separate supply chains to produce a product. This strategy can also be explained by the fear of vaccine nationalism, as many feared at the beginning of the pandemic. Additionally, the use of multiple chains for different regions provided a safety net by providing manufacturing support in the case where one supply chain is not able to meet its required production. However, having multiple supply chains is both financially and capital intensive, therefore many pharmaceutical companies decided to have partnerships. While partnerships helped to lessen the production load, they also provided some with access to new technology. This was the case with BioNTech and Pfizer. Other companies, such as Moderna, did not form any partnership and decided to outsource part of their production to other companies. This makes it possible for them to keep their intellectual propriety of vaccine formulation.

With the risk of new pandemics, the need to create, produce and deliver a vaccine becomes one of our priorities to limit its impact. In this project, we aim to classify vaccine producers in times of a global pandemic, where having a global supply chain would be necessary to fulfill the growing demand. Looking at their upstream and downstream may give us a good understanding of vaccine producers' approach and structure in times of crisis. This project will combine international business (IB) classification techniques, mainly geographical diversification, to the Covid-19 vaccine producers' supply chain during the early stage of the pandemic. By applying an IB technique to the vaccine supply chain, we seek to provide insights into the distinct characteristics and strategies employed by major vaccine manufacturers.

1.2 Literature Review

The efficient and effective distribution of vaccines plays a crucial role in ensuring widespread immunization and preventing the spread of infectious diseases. The vaccine supply chain encompasses various stages, from vaccine production to delivery and administration. Research has been done about the vaccine supply chain. Such research includes the paper on the literature review of the vaccine supply chain by Duijzer et al. (2018). In their paper, they have highlighted several key points in the VSC research field. The vaccine supply chain is a complex multiple stage chain and requires efficient and effective management to ensure that vaccines reach those who need them most, especially in low- and middle-income countries (Duijzer et al., 2018). Duijzer et al. (2018) conclude that collaboration and coordination among the VSC's stakeholders are key to achieving a successful VSC. In their paper, the authors also have identified key challenges unique to VSC. The vaccine supply chain's main challenges include maintaining vaccine potency and stability during transportation and storage, ensuring equitable distribution, and addressing supply chain disruptions due to emergencies or outbreaks (Duijzer et al., 2018).

The vaccine supply chain faces numerous challenges that can impact its efficiency and effectiveness. Research has identified several key areas of concern. Cold chain management is a critical aspect of supply chain operations, particularly in industries such as pharmaceuticals, food, and vaccines, where temperature control is essential. This literature review examines the key challenges, strategies, and technologies associated with cold chain management. In the pharmaceutical industry, cold chain management is essential for preserving the potency and effectiveness of temperature-sensitive drugs and biologics. According to Ong et al. (2021), temperature excursions during storage and transportation can lead to drug degradation and compromise patient safety. Hence, robust cold chain management practices are vital to maintaining pharmaceutical product quality. Monitoring and maintaining the required temperature range throughout the supply chain is critical. However, achieving accurate and real-time temperature monitoring can be challenging, especially during transportation. Yavuz et al. (2021) note that temperature monitoring devices, such as data loggers and IoT sensors, can help ensure temperature compliance and enable proactive interventions. Additionally, developing and maintaining the necessary infrastructure for cold chain operations is crucial. Adequate storage facilities, refrigeration units, and transportation systems are essential to maintain the required temperature conditions. According to Tseng et al. (2020), optimizing infrastructure design, layout, and capacity can enhance the efficiency and effectiveness of cold chain operations.

Efficient distribution networks are crucial for reaching remote areas and ensuring timely vaccine delivery. Perea et al. (2020) highlight the need for optimized routing and scheduling algorithms to minimize transportation costs and improve delivery speed. Additionally, effective inventory management and demand forecasting are essential to avoid stockouts and wastage. Furthermore, achieving equitable vaccine distribution is a significant challenge, particularly in low-income countries. Ahmed et al. (2022) emphasize the importance of addressing disparities in access, availability, and affordability of vaccines across different regions and populations. Strengthening supply chain infrastructure and implementing targeted interventions can help improve vaccine equity.

To overcome the challenges associated with the vaccine supply chain, researchers have proposed various strategies including collaboration, technology adoption and data analytics. Effective collaboration among stakeholders, including manufacturers, governments, healthcare providers, and NGOs, is crucial for optimizing the vaccine supply chain. Li et al. (2021) argue that partnerships can help enhance coordination, information sharing, and resource allocation, leading to improved supply chain performance. Leveraging technology can significantly enhance vaccine supply chain operations.

For instance, blockchain technology can enhance traceability and transparency, ensuring the authenticity of vaccine shipments (Praveen et al., 2021). Similarly, the Internet of Things (IoT) devices and sensors enable real-time temperature monitoring and quality assurance (Musa et al., 2022). The adoption of these technologies can enhance supply chain visibility and reduce risks. Additionally, advanced data analytics techniques and predictive modelling can improve decision-making within the vaccine supply chain. Shen et al. (2020) demonstrate that data-driven forecasting models can aid in accurate demand prediction and inventory optimization. Furthermore, real-time data analytics can help identify supply chain bottlenecks and enable proactive interventions.

The Covid-19 pandemic has highlighted the need for resilient vaccine supply chains. Resilience refers to the ability to withstand disruptions and recover quickly. The literature has explored strategies to enhance the resilience of vaccine supply chains. Building redundancy and diversification into the supply chain can mitigate risks associated with single-source dependencies and transportation disruptions. Kumar et al. (2021) argue that maintaining multiple manufacturing facilities, transportation routes, and suppliers can enhance resilience and minimize the impact of disruptions. Developing comprehensive contingency plans and conducting risk assessments are critical for identifying vulnerabilities and potential disruptions. According to Danilov et al. (2021), scenario-based planning and risk assessment can help supply chain managers make informed decisions and prepare for contingencies effectively. Researchers such as Meng et al. (2020) emphasize the importance of conducting risk assessments at different stages of the cold chain, including transportation, storage, and handling, to proactively identify and mitigate risks to product integrity.

Vaccine Nationalism & Vaccine Clubs

The global response to the COVID-19 pandemic has unveiled a concerning trend known as "vaccine nationalism". Vaccine nationalism is characterized by countries prioritizing their populations at the expense of others (Ngo, 2021). This phenomenon is not a novel concept but has been exacerbated by the urgency of the COVID-19 pandemic (Gupta, 2020). Historically, nations have sought to secure essential medical supplies during health crises, but the scale and speed of vaccine nationalism during the current pandemic have raised significant concerns (Thompson, 2021). Indeed, vaccine nationalism perpetuates global inequity by allowing high-income countries to secure the lion's share of vaccine doses, leaving low- and middle-income nations at a disadvantage (Polack et al., 2020). This brings us back to one of the characteristics of VSC. As mentioned previously, vaccine producers are for-profit organizations thus giving an advantage to high-income countries, who have the means to bide a higher price for vaccines than low- and middle-income, thus leaving them to share and salvage what vaccine is left.

Furthermore, By prioritizing their populations, vaccine-nationalist countries risk prolonging the global pandemic as the virus continues to circulate unchecked in other regions (Sridhar et al., 2021). Indeed, the WHO and epidemiologists have pushed forward the necessity of vaccination equity to reach global herd immunity, thus containing the spread of the virus (WHO, 2020).

Finally, according to Fidler (2021), competition for vaccines has led to diplomatic tensions, export bans and trade disputes, straining international relations. Indeed, countries such as the United States, the United Kingdom, the European Union, Russia and China have been criticized and accused of hampering the global effort against the virus by hoarding vaccines. The most criticized country has been the United States. Indeed, the "America First" approach has come under intense scrutiny, with the US initially securing large doses of vaccine for its population (Bollyky, 2021). Additionally, Canada was also

challenged and associated with vaccine nationalism. As of 2020, Canada had enough vaccines to vaccinate its population 9 times (Mullard, 2020)

In addition to the concept of vaccine nationalism, the notion vaccine club has also received a lot of attention during the Covid-19 pandemic. Unlike vaccine nationalism, the vaccine is a cooperative alliance among countries aimed at improving equitable access to vaccines and these agreements can take various forms, including regional alliances or multilateral initiatives (Smith, 2021). In the paper titled "The Covid-19 Vaccine Production Club: Will Value Chains Temper Nationalism?", the vaccine club refers to vaccine producers and vaccine ingredient producers (Evenett et al., 2021). Based on these definitions, researchers have concluded the following implications. Vaccine clubs offer mechanisms to pool resources, expertise, and vaccine supplies, ensuring that low- and middle-income countries receive their fair share (Bollyky & Bown, 2021). Additionally, collaborative efforts within vaccine clubs can expedite the research, development, and production of vaccines, making them more readily available (Fidler, 2020). Lastly, vaccine clubs are a favourable environment for fostering international cooperation, vaccine clubs can reduce conflicts arising from vaccine shortages and promote global political stability (World Bank, 2021). On the other hand, Pisani-ferry et al. (2021) highlighted the exclusivity of vaccine clubs, which implies that participating countries secure preferential access to vaccines. Thus, potentially undermining the global goal of equitable vaccine distribution. Moreover, Effective governance and coordination among member countries pose challenges, as diverse interests and priorities must be reconciled (Gostin, 2020). Another critique of vaccine clubs is geopolitical tensions, this is particularly true between major powers if not structured carefully (Phelan et al., 2020).

Internationalization classification

Measuring the internationalization of a firm is crucial for understanding its level of involvement and success in global markets. Scholars have developed various techniques to assess and quantify the internationalization process. Which is the best strategy for a firm? Regionalization or globalization. This question has been extensively discussed in previous and recent academic papers as well as their implication on firms' performance (Berrill, 2015). The debate divided the International Business literature into different schools of thought regarding the strategy the largest firms are using. Indeed, some academics, as demonstrated in Yip (2002) and Govindarajan & Gupta (2008) papers, strongly argue that global strategy is primordial for MNEs to compete among the largest firm, like, Ghemanwat (2001), argues the form theory by demonstrating that there is a positive relationship between geographical and/or cultural distance and the cost of internationalization, therefore concluding that the largest firms follow a semi-global strategy. More authors have taken part in this debate, such as Doremus et al. (1998) who concluded that large MNEs are national with a regional focus rather than global. Lastly, using the concept of Triad, Rugman & Verbeke (2003, 2004, 2007, 2008) concluded that the largest MNEs are national.

Asmusen et al. (2007) discuss in their paper that the already covered ways of measuring a firm's internationalization were not the actual measuring the degree of globalization. They identify three main indicators commonly used to measure the degree of internationalization firms based on IB literature. According to the authors, internationalization has been measured using an index of asset distribution; (2) using an index of spread measure; (3) using an index of psychology.

- Index of asset distribution: This type of measure uses the distribution of certain assets, such as, between the home country and other countries.

- Index of spread measure: This type of measure uses a firm's spread of function, assets and/or employees across cultures and/or countries.
- Index of psychology: This type of measure is focused on how employees, mainly management, are culturally diverse.

Furthermore, they pointed out that some indices use a dichotomous approach, using one of the three measurement indicators, while others used two or all indicators making it a composite index (multi-type indices). They found that when using the dichotomous approach sales is the most used (Rugman & Vebeke,2004), in addition to its being the simplest internationalization indices. Other variables had been used, such as shareholder (Hassel et al. 2003) or value added (UNCTAD, 2004). In their paper, Asmusen et al. (2007) identify two limitations to the dichotomous approach. The first one is that using home vs foreign become void when dealing with cross-country comparison. The second limitation is that this approach does not consider the spread of foreign activities across different countries. While one-type indices give the bigger picture, it gives one facet of the answer. for that reason, several authors introduced the use of multi-type indices (Letto-Gillies, 1998; Germann et al., 1999; Sullivan, 1994). As mentioned previously, the multi-type indices use two or all the measure indicators. In their paper, Hassel et al. (2000) combined asset distribution with the spread measure. Similar to the one-type indices, the validity of multi-type indices' results has been questioned due to their complexity (Ramaswamy et al., 1996; Fisch/Oesterle, 2003).

In their paper, Qian et al. (2008) discussed how most international forms are regional based rather than global. They classify regional diversification into two groups: low to moderate diversification and moderate to high diversification. Through their research, they came to that conclusion after evaluating the impact of the performance of different degrees of regional diversification on a sample of 189 firms in the US. This can be explained by the fact that expanding activities within and out of the region comes at a cost, and that said cost varies from region to region. Overall, the authors believed that firms in the developed country would benefit more, in terms of infrastructure and economic development, by diversifying those activities in a moderate number of developed country regions and limiting that global expansion to a limited number of developing countries (Qian et al. 2008).

As mentioned above, measuring the internationalization of a firm has been a recurrent topic in the IB field and several techniques have been used to it. One widely recognized approach to measuring internationalization is through the analysis of Foreign Direct Investment (FDI) activities. Dunning (1988) proposed the eclectic paradigm, which emphasizes ownership, location, and internalization advantages as drivers of FDI. FDI reflects a firm's investment in foreign countries, such as establishing subsidiaries or acquiring ownership stakes in foreign companies. FDI data from sources like UNCTAD can offer insights into a firm's internationalization efforts.

Export intensity is another technique used to measure the internationalization of firms. It quantifies the proportion of a firm's total sales revenue generated from exports. By dividing export sales by total sales and multiplying by 100, export intensity indicates a firm's reliance on international markets. Financial statements and trade databases can be utilized to calculate and analyze export intensity.

Another method used to measure internationalization is by examining the growth rate of a firm's international sales, which provides insights into its expansion in foreign markets. This measure of international sales growth reflects the degree of a firm's internationalization over a specific period. The formula for calculating international sales growth involves subtracting the previous year's international sales from the current year's international sales, dividing by the previous year's international sales, and multiplying by 100. Contractor et al. (2007) highlight the importance of understanding the relationship

between international expansion and firm performance, which can be further assessed through international sales growth.

Lastly, geographical diversification is another crucial technique for measuring the internationalization of firms. It quantifies the number of countries in which a firm operates, reflecting its geographical spread and market diversification. A higher number of countries indicates a greater level of internationalization. This information can be extracted from company reports, databases, and market research sources.

Furthermore, several internationalization indexes have been developed to provide comprehensive measures of a firm's internationalization. For instance, the Transnationality Index (TNI), Internationalization Index (II), and Geocentric Orientation Index (GOI) consider factors such as foreign assets, foreign sales, and foreign employment to calculate a score representing the firm's internationalization level. Rugman and Verbeke (2004) emphasize the significance of regional and global strategies in the context of multinational enterprises, which can be further examined using these internationalization indexes.

Rugman & Verbeke classification

Rugman & Verbeke (2004) proposed a new way of classifying MNS based on previous international business literature using geographical diversification. The triad power concept is one of the main key concepts, they based their paper on. This concept, first introduced in 1985 by Kenichi Ohmae, divides the world into three main geographic locations: the US, the EU and Japan. The geographic location was chosen by Ohmae due to the similarities. According to Ohmae, the US, EU and Japan, all have low macroeconomic growth, similar technological infrastructure, presence of large firms in most industries, protectionist pressure and homogenize demand. In their paper, Rugman & Verbeke kept the concept of the triad but expanded the region proposed by Ohmae to include more countries. Thus, creating a new triad constituted of North America, Europe, and Asia, represented by *Figure 1-2*.

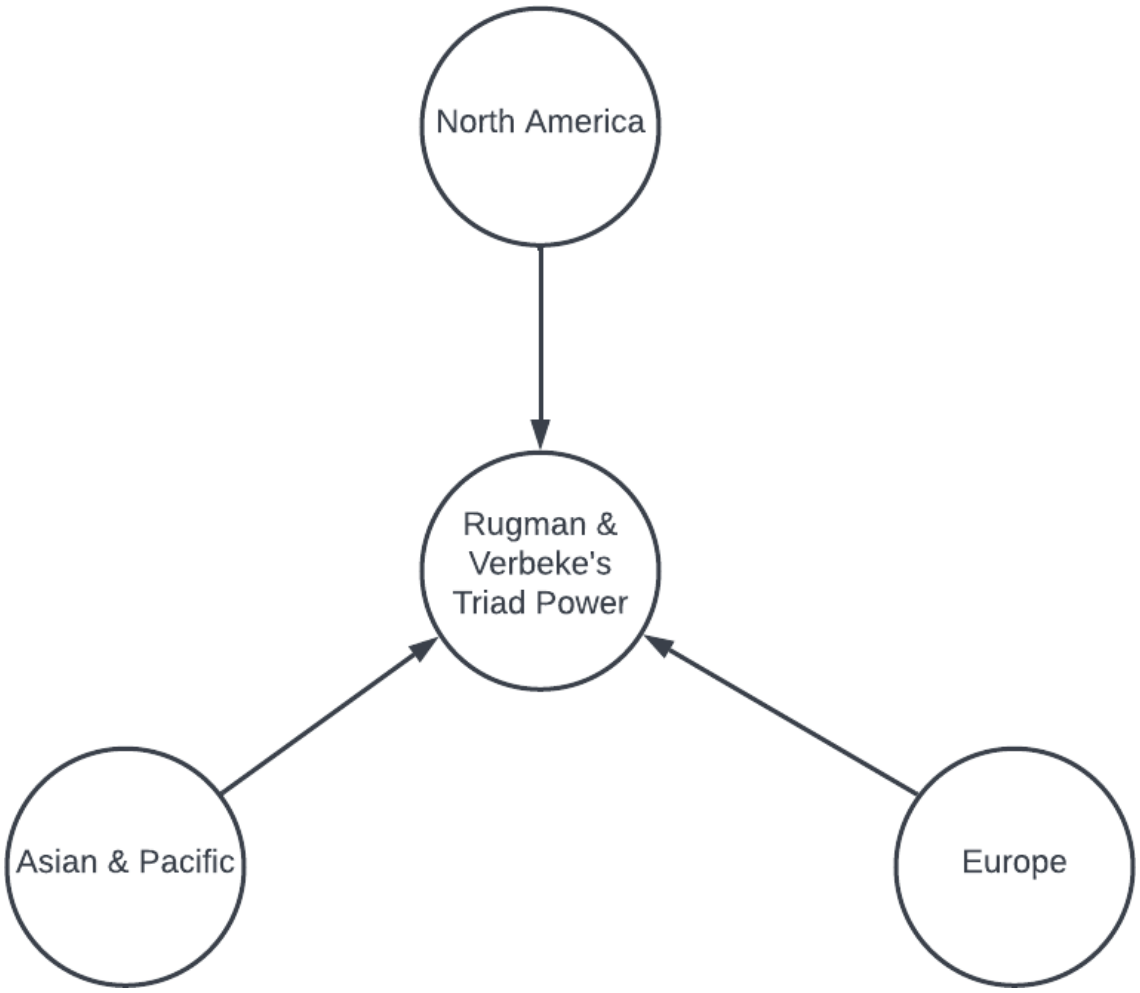


Figure 1-2: Rugman & Verbeke's Triad

Once the geographic location had been set, they were left with choosing a dimension on which to base their classification. For their project, Rugman & Verbeke decided on the sales of the 500 largest MNEs. Their choice was based on the fact that sale is a good indicator of market success.

In their paper, Rugman & Verbeke choose a sample of 500 companies, representing the 500 largest MNEs in the world. However, out of the 500 companies, only 380 provided data sales data per geographic segment. Out of the 380 MNEs, used in their paper, they were able to classify them as whether they are home region-oriented, bi-regional, host region-oriented or global companies. Home region-oriented companies are characterized by firms having 50% or more of their sale in the home region. Bi-regional MNEs are companies with 20% or more of their sale in two regions of the triad, but no less than 50% in one of them. Host region-oriented MNEs have 50% or more of their sales in a triad region other than their home region. Global MNEs have 20% or more of their sales in all parts of the triad, but 50% or less in one specific region. Based on their definition, Rugman and Verbeke found that out of the initial sample of 500 companies, 320 firms were judged as being home region-oriented, 25 were bi-regional, 11 were host region-oriented and only 9 were global firms.

From their finding, they were able to assess that very few companies are global, according to their criteria, despite popular belief. They also concluded that balancing sales across the triad is not crucial for all MNEs.

Risk and internationalization

As companies expand their operations, understanding the risks associated with internationalization classification becomes crucial. Therefore, extensive literature explores the risks linked to different internationalization classifications (home region-oriented, bi-regional, host region-oriented and global), shedding light on the potential challenges and considerations for organizations engaging in global expansion.

One of the key risks associated with a home-oriented company is market saturation. As Brouthers and Nakos (2004) suggest, relying solely on the domestic market can limit growth opportunities for companies, particularly in mature markets. The authors argue that market saturation can hinder a company's ability to expand and achieve sustained growth. Home-oriented companies are also more susceptible to economic downturns within their home country, as highlighted by Xu and Shenkar (2002). The authors assert that companies with a limited international presence face higher risks during economic crises and recessions. This vulnerability stems from the lack of diversification into international markets, leaving the company heavily reliant on the economic conditions of a single market. Additionally, operating only in the home market can expose companies to a competitive disadvantage compared to their global counterparts, as noted by Rugman and Verbeke (2004). The authors argue that companies that have successfully internationalized often benefit from economies of scale, access to a broader customer base, and enhanced brand recognition. A home-oriented company may struggle to compete effectively on a global scale, facing intensified competition and potentially losing market share. Lastly, a home-oriented company's success can be heavily dependent on local factors and is exposed to political risks. As highlighted by Contractor, Kundu, and Hsu (2003), reliance on a single market exposes the company to political and regulatory risks that could significantly impact its operations. These risks include changes in trade policies, shifts in government regulations, or geopolitical tensions that could disrupt supply chains or impede market access.

Among the risks associated with a bi-regionally oriented company, market concentration was found to be one of the main ones. As indicated by Klier and Rubenstein (2008), reliance on a limited number of regions can lead to vulnerability when those regions face economic or political instability. The authors argue that diversifying into multiple regions can help mitigate this risk by spreading exposure to regional market fluctuations. Being bi-regional oriented also exposes companies to the risks associated with specific regions, including economic, political, and regulatory factors. According to Delios and Henisz (2003), dependence on regional factors can result in heightened exposure to country-specific risks such as changes in government policies, economic downturns, or legal and regulatory uncertainties. Companies with a more global footprint can diversify their risk exposure by operating in multiple regions, reducing dependence on any one specific market. This strategy can result in a lack of in-depth knowledge and expertise in markets outside the selected regions. As highlighted by Cantwell and Mudambi (2005), expanding into global markets allows companies to gain valuable insights into diverse customer preferences, market dynamics, and technological advancements. This broader knowledge base facilitates adaptation, innovation, and the ability to respond effectively to changing global trends.

A significant risk associated with a host region-oriented company is the limited market reach and growth potential it may experience. As noted by Mudambi and Navarra (2004), companies that concentrate their operations solely within a host region can miss out on opportunities available in other regions.

This limited market access restricts their ability to tap into new customer bases, emerging markets, and potential growth opportunities, potentially leading to stagnation or slower growth rates. Similar to home region-oriented companies, a host region-oriented company is highly dependent on the economic conditions of the specific region in which it operates. As highlighted by Anderson and Gatignon (1986), relying solely on a single host region exposes the company to risks associated with local economic downturns, industry-specific challenges, or changes in consumer spending patterns. Fluctuations in the host region's economic climate can significantly impact the company's financial performance and stability. Operating in a single host region exposes companies to regulatory and political risks unique to that specific region. As discussed by Meyer and Nguyen (2005), host region-oriented companies face challenges related to compliance with local laws, regulations, and government policies. Changes in regulatory frameworks, shifts in political landscapes, or geopolitical tensions can disrupt business operations, hinder market access, and increase uncertainty for the company.

Becoming a global company exposes organizations to the complexities and uncertainties of diverse markets. As noted by Ghemawat (2001), operating in multiple markets entails navigating diverse cultural, economic, and regulatory landscapes. This complexity increases the risks associated with market volatility, shifts in consumer preferences, and changes in political and legal environments. Global companies must develop strategies to manage and adapt to the inherent uncertainties of operating across multiple markets. Additionally, operating as a global company involves managing complex and interconnected supply chains that span multiple countries and regions. As discussed by Choi, Dooley, and Rungtusanatham (2001), global companies face risks associated with supply chain disruptions, including natural disasters, political instability, transportation disruptions, and supplier failures. The interconnected nature of global supply chains amplifies the potential impact of such disruptions, emphasizing the need for robust supply chain risk management practices.

Ghemawat (2001) found that global companies are exposed to the complexities and uncertainties of their diverse markets and currency and financial risks arising from fluctuations in exchange rates, interest rates, and global economic conditions (Buckley et al., 2006). Ghemawat (2001) noted that operating in multiple markets entails navigating diverse cultural, economic, and regulatory landscapes, thus increasing the risks associated with market volatility, shifts in consumer preferences, and changes in political and legal environments. Global companies must develop strategies to manage and adapt to the inherent uncertainties of operating across multiple markets. In their paper, Buckley, Clegg, and Tan (2006) conclude that conducting business across borders necessitates dealing with foreign currencies impacting profitability, cash flow, and financial performance. Effective risk management strategies, such as hedging currency exposures or diversifying financing sources, are essential for mitigating these risks.

Hypotheses

From the literature review, we were able to formulate three hypotheses:

- **H1**: Certain regions of the triad will show a higher concentration of facilities.
- **H2**: Formulation facilities distribution will lead to a classification of home region-oriented using Rugman & Verbeke's classification.
- **H3**: Downstream dimension will show that Covid-19 vaccine producers are depending on certain regions of the triad.

1.3 Methodology

This study uses a quantitative approach using R to analyze the Covid-19 vaccine supply chain as of June 2021 using international business internationalization classification. We decided to adopt the geographical diversification techniques to classify the Covid-19 VSC, by recreating Rugman & Verbeke's classification model on the production of vaccines in the case of a global pandemic. We chose this method as it was the best fit with the data, we had on the vaccine supply chain. Additionally, this classification method is compatible with publicly available data, (e.g., address, country) which makes up the majority of our database, without undermining the result of the classification.

The variable used in Rugman & Verbeke's original model has been adjusted to fit our project. Instead of just using sales, we are using two different dimensions, an upstream and downstream dimension, to assess the globalization of vaccine suppliers. The upstream dimension, we used in this project is the presence of the supply chain in each region of the triad. As our downstream dimension, we used the number of doses of each vaccine administrated. The issue of downstream vs. upstream globalization was discussed in Rugman & Verbeke's paper.

In this project, we are measuring and categorizing the globalization level of Covid-19 vaccine producers. The following producer and their supply chain were used: Pfizer, Moderna, Johnson&Johnson, Novavax, AstraZeneca and CureVac. Lists of companies directly involved in the vaccine production as of June 2021 were collected and six data frames were created, one for each producer. Pfizer, Moderna, Johnson&Johnson, AstraZeneca, Novavax & CureVac have 28, 16, 13, 24, 20 & 9 companies actively involved in their supply chain. All the data used for this part of the project was collected from the article by Bown and Bollyky on the Covid-19 vaccine titled "How Covid-19 vaccine supply chains emerged amid a Pandemic" (2021). The article provided the companies involved in the supply chain of the six Covid-19 vaccine producers mentioned above.

To create our datasets, 7 variables were identified as presented in *Table 1-1*. Each database was processed and observations/companies with missing information were removed. Variables, such as latitude and longitude, were created by converting the address of a company into its geographic coordinates. As for the stage of production variable, a company can be in the substance manufacturing stage, the lipids production or adjuvant production stage and the fill and finish stage. This information was available in Bown and Bollyky's article.

Table 1-1: Variable description Pfizer dataset, the Moderna dataset and the Covid-19 vaccine producer dataset

Variable name	Description
Company names	The name of the company.
Country	The country in with the company is located.
Address	The address of the company involved in the vaccine production.

Variable name	Description
Stage of production	The stage of production of the vaccine at which the company is.
Function	The specific function of the company within the chain.
Longitude	The longitude coordination of the company.
Latitude	The latitude coordination of the company.

As per Rugman & Verbeke's paper, we need to define the Triad region that will be used in this project. We decided to follow the same triad as in Rugman & Verbeke's paper. Thus, the regions used in this paper are North America, Europe, and Asia/Pacific. With the region of the triad defined, we can proceed with the classification of our supply chains.

As mentioned previously, two dimensions were used in this project, the dose administrated ratio and the active presence ratio, downstream and upstream dimensions. Starting with the upstream, we started by taking the ratio of companies in each region of the triad, this was done by counting the number of companies in a region and then dividing it by the total number of companies in the supply chain. Once the ratio was calculated, we then created a table with the results and applied Rugman & Verbeke's (2004) criteria of classification to our ratios. Additionally, we wanted to calculate the proportion of each stage of production for our six vaccine producers. This was also calculated using the ratio of formulation companies, the ratio of lipids production companies and the fill & finish companies on the total number of active companies in the chain. Once we had the percentage of formulation manufacturers, lipids producers and fill & finish companies, we want to evaluate whether the stages of production were evenly distributed in the triad or if they were bound to any specific region. As a result, we were able to extract the three additional tables presented in the result section of this paper. This process was conducted six times for all the vaccine producers mentioned in the paper.

Lastly, "sales" quantity was deemed as an important variable during the development of the project since it would be considered a measurement of success for the covid-19 vaccine supply chain. However, as sales information for the six vaccine producers is not publicly available, the number of vaccine doses administered has been used as an approximation. This data will be used as an estimate of the quantity sold by each company during a given period and in given countries. We collected the number of Pfizer's, Moderna's, Johnson&Johnson's, AstraZeneca's & Novavax's Covid-19 vaccines administrated in 34 countries for one year between January 2021 and January 2022 using "Covid vaccine doses by manufacturer" by Our World in Data(2023). Unfortunately, information related to the doses of CureVac administrated was not available. As a result, its downstream classification will be omitted from this section part of the project. With this information, we can find the ratio of vaccine doses administrated per region based on the total number of doses administrated across the 34 countries by counting and then dividing the number of administrated doses in each region by the total. This was done for all six Covid-19 vaccine producers.

Once the six VCS were categorized, we created six maps to visualize the six supply chains based using the geographical coordinates derived from each company's address. Each map was created in R using the leaflet package. To create the maps, we first needed to upload our six datasets into R. Once the data have been loaded, we were then able to create the maps. The leaflet package already provides blank world maps that we will use as the base of our six maps, therefore all that is left to be done is to add a marker at the location of each company present in the chain. Thus, we only require two variables, the longitude, and the latitude coordinate of each company to create our maps. This visualization step was

important as it reinforce the fitness of Rugman & Verbeke's model by providing us with a visual confirmation of the geographic distribution of the six supply chains.

1.4 Results

Starting with the upstream dimension, *Table 1-2* shows the supply chain distribution of the Covid-19 vaccine producers. Pfizer has 32% of its supply chain in North America, 61% in Europe, 4% in Asia-Pacific and the remaining 3% are outside of the Triad. Therefore, based on Rugman & Verbeke's classification, is both home region-oriented and host region-oriented. Indeed, Pfizer has two headquarters one in the US and the other in Germany, due to their partnership with a German company, BioNTech. This makes it possible for Pfizer to have two different classifications depending on the headquarters used for the classification. When using the German headquarter, Pfizer would be classified as home region-oriented, while when taking the US headquarter it would be host region-oriented. Moderna has 50% of its supply chain in North America, 44% in Europe and 6% in Asia-Pacific. With its headquarter in the US, Moderna is then classified as a home region-oriented Covid-19 vaccine producer, with 50% of its chain located in North America. Johnson&Johnson has 38% of its supply chain in North America, 38% in Europe and 15% in Asia-Pacific and the remaining 9% are outside of the Triad. Similar to Moderna, Johnson&Johnson's headquarter is in the US. This headquarters location classifies it as a bi-regional, North American and European, vaccine producer. AstraZeneca has 17% of its supply chain in North America, 42% in Europe and 33% in Asia-Pacific and the remaining 8% are outside of the Triad. With its headquarter located in the UK, AstraZeneca is classified as bi-regional, alongside Johnson&Johnson, with 42% of the chain in Europe and 33% in Asia-Pacific. Novavax has 45% of its supply chain in North America, 40% in Europe and 15% in Asia-Pacific. Based on its headquarter located in the US, Novavax is classified as bi-regional, with 45% of the chain in North America and 40 in Europe. CureVac has 0% of its supply chain in North America, 100% in Europe and 0% in Asia-Pacific. With its headquarter located in Germany, CureVac is then classified as home region-oriented since 100% of its chain is in Europe. Make CureVac, the only producer in our project to have its whole chain in one region of the triad.

Table 1-2: Covid-19 vaccine producers' presence in the Triad region & function distribution

Brand	% in North America	% in Europe	% in Asia & Pacific	% in formulation	% in lipid	% in fill and finish
Pfizer	32%	61%	4%	36%	39%	25%
Moderna	50%	44%	6%	38%	19%	44%
Johnson&Johnson	38%	38%	15%	38%	0%	62%
AstraZeneca	17%	42%	33%	54%	0%	46%
Novavax	45%	40%	15%	50%	25%	25%
CureVac	0%	100%	0%	78%	0%	22%

Table 1-3 is the distribution of the formulation facilities across the triad. Pfizer has 30% of its vaccine formulation facilities in North America, 60% in Europe and 10% in Asia-Pacific. As mentioned above, Pfizer has two classifications based on the headquarters location we look at. Thus, when looking from the US headquarter point of view, Pfizer is classified as host region-oriented, while from German headquarter Pfizer is then home region-oriented. Moderna has 50% of its vaccine formulation facilities in North America, 50% in Europe and 0% in Asia-Pacific. With this distribution, Moderna has a bi-

regional strategy for its vaccine formulation. Johnson&Johnson has 40% of its vaccine formulation facilities in North America, 20% in Europe and 40% in Asia-Pacific. Based on Rugman & Verbeke's paper, Johnson&Johnson would be considered as global. AstraZeneca has 15% of its vaccine formulation facilities in North America, 38% in Europe, 38% in Asia-Pacific and the remaining percentages are located outside of the Triad. Thus, classifying AstraZeneca as having a bi-regional for its formulation facilities. Novavax has 40% of its vaccine formulation facilities in North America, 30% in Europe and 30% in Asia-Pacific. Thus, classifying Novavax as having a global strategy for its formulation facilities. CureVac has 0% of its vaccine formulation facilities in North America, 100% in Europe and 0% in Asia-Pacific. By having all its formulation facilities in Europe, CureVac is then classified as home region-oriented.

Table 1-3: Distribution of formulation facilities across Triad per Covid-19 vaccine producers

Brand	Formulation/North	Formulation/Europe	Formulation/Asia
Pfizer	30%	60%	10%
Moderna	50%	50%	0%
Johnson&Johnson	40%	20%	40%
AstraZeneca*	15%	38%	38%
Novavax	40%	30%	30%
CureVac	0%	100%	0%

* *Other region = 8%*

For the lipids/adjuvant facilities distribution, *Table 1-4* shows that Pfizer has 55% of its vaccine lipids/adjuvant facilities in North America, 45% in Europe and 0% in Asia-Pacific. Therefore, Pfizer is both home region-oriented, and host region-oriented given its two headquarters. Moderna has 33% of its vaccine lipids/adjuvant facilities in North America, 67% in Europe and 0% in Asia-Pacific. It would then be classified as host region-oriented for its lipids production. Novavax has 60% of its vaccine lipids/adjuvant facilities in North America, 40% in Europe and 0% in Asia-Pacific. It would then be classified as home region-oriented for its lipids production. From the data collected neither Johnson&Johnson nor AstraZeneca nor CureVac seem to have any lipids/adjuvant input as part of their supply chain.

Table 1-4: Distribution of lipids/adjuvant facilities across Triad per Covid-19 vaccine producers

Brand	Lipid/North	Lipid/Europe	Lipid/Asia
Pfizer	55%	45%	0%
Moderna	33%	67%	0%
Johnson&Johnson	n/a	n/a	n/a
AstraZeneca	n/a	n/a	n/a
Novavax	60%	40%	0%

Brand	Lipid/North	Lipid/Europe	Lipid/Asia
CureVac	n/a	n/a	n/a

Lastly, *Table 1-5* is the distribution fill and finish facilities in the triad. Pfizer has 0% of its fill & finish facilities in North America, 86% in Europe, 0% in Asia-Pacific and the remaining percentages are located outside of the Triad. in this case, Pfizer is still home region oriented and host region-oriented. Moderna has 57% of its fill & finish facilities in North America, 29% in Europe, and 14% in Asia-Pacific. Given these percentages, Moderna would be classified as home region-oriented. Johnson&Johnson has 37.5% of its fill & finish facilities in North America, 50% in Europe, 0% in Asia-Pacific and the remaining percentages are located outside of the Triad. Given these percentages, Johnson&Johnson would be classified as home region-oriented. AstraZeneca has 18% of its fill & finish facilities in North America, 45% in Europe, 27% in Asia-Pacific and the remaining percentages are located outside of the Triad. With these percentages, AstraZeneca would be classified as bi-regional. Novavax has 40% of its fill & finish facilities in North America, 60% in Europe and 0% in Asia-Pacific. Given these percentages, Novavax would be classified as host region-oriented. CureVac has 0% of its fill & finish facilities in North America, 100% in Europe and 0% in Asia-Pacific. Given these percentages, CureVac would be classified as home region-oriented.

Table 1-5: Distribution of fill and finish facilities across Triad per Covid-19 vaccine producers

Brand	Fill/North	Fill/Europe	Fill/Asia
Pfizer*	0%	86%	0%
Moderna	57%	29%	14%
Johnson&Johnson**	37.5%	50%	0%
AstraZeneca***	18%	45%	27%
Novavax	40%	60%	0%
CureVac	0%	100%	0%

** Other region = 14%*

*** Other region = 13%*

**** Other region = 9%*

Now looking at the downstream dimension. *Table 1-6* is the distribution of doses of the Covid-19 vaccine in the triad region. We can see that Pfizer has 23% of its dose administrated in North America, 41% in Europe, and 5% in Asia-Pacific and the remaining percentages are located outside of the Triad. In this case, Pfizer is bi-regional. Moderna has 52% of its dose administrated in North America, 27% in Europe, 4% in Asia-Pacific and the remaining percentages are located outside of the Triad. Given these percentages, Moderna would be classified as home region-oriented for its downstream distribution. Johnson&Johnson has 37% of its dose administrated in North America, 35% in Europe, and 5% in Asia-Pacific and the remaining percentages are located outside of the Triad. Given this allocation structure of the dose administrated, Johnson&Johnson would be classified as bi-regional. AstraZeneca has 0.25% of its dose administrated in North America, 44% in Europe, 13% in Asia-Pacific and the remaining percentages are located outside of the Triad. Given this percentage, it is hard to allocate AstraZeneca to

a specific group for the number of doses administrated. Indeed, when following Rugman & Verbeke's classification grid, we can see that AstraZeneca would be just at the edge of two groups without fulfilling all their requirement. Novavax has 0.003% of its dose administrated in North America, 27% in Europe, and 48% in Asia-Pacific percentages are located outside of the Triad. Given this allocation of the dose administrated, Novavax would be classified as bi-regional.

Table 1-6: Distribution of vaccine dose administrated across triad per Covid-19 vaccine producers.

Brand	% North America	% Europe	% Asia-Pacific
Pfizer	23%	41%	5%
Moderna	52%	27%	4%
Johnson&Johnson	37%	35%	5%
AstraZeneca	0.257%	44%	13%
Novavax	0.003%	27%	48%

When taking a closer look are the disposition of our six supply chains. *Figure 1-3* shows the location of Pfizer's Covid-19 vaccine supply chain active companies. In this map, we are able to see that most of Pfizer's supply chain is located in Europe with 17 facilities located there. The second region with the most companies is the US with 8 companies. On the Pfizer map, it is also possible to see that one company is located in South Africa, one in China and one in Canada. Those companies are mostly used for manufacturing support. Indeed, the facilities in South Africa and China are used to assist at the fill and finish level of the supply chain. This heavy reliance on the facilities in Europe and the US leaves Pfizer's supply chain vulnerable to disruption if there were to be risk associated with one or both regions.

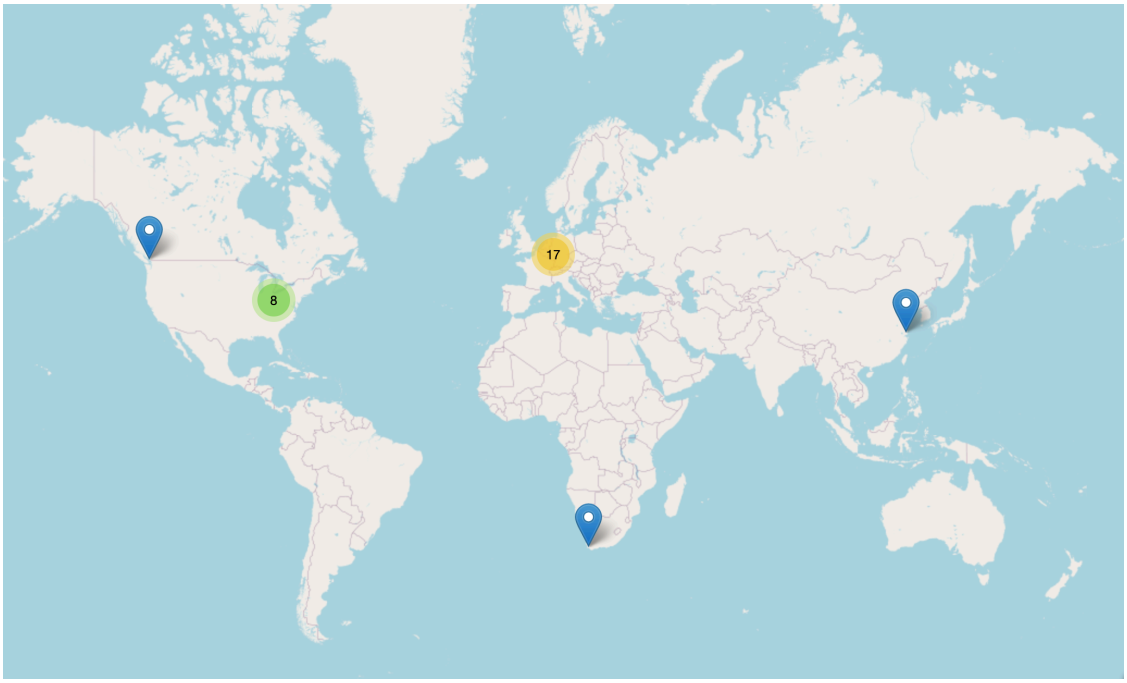


Figure 1-3: Cartography of Pfizer's supply chain

When comparing Moderna's map to Pfizer, we can that there is a difference between both supply chains in terms of their size. Indeed, Moderna's supply chain has a smaller number of companies operating in its supply chain in comparison to Pfizer. We believe this is due to Pfizer being an older company than Moderna, therefore giving a possible advantage to Pfizer over Moderna. Like Pfizer, we can see most of Moderna's supply chain is located in the US with 8 facilities (*Figure 1-4*). The second region with the most companies of Moderna's Covid-19 VSC is Europe with 7 companies. It is also possible to see that one company is in South Korea. This company is mostly used for manufacturing support, more specifically to assist at the fill and finish level of the supply chain. Heavy reliance on the facilities in Europe and the US, leaves Moderna's supply chain vulnerable to disruption if there were to have risk associated with one or both regions like Pfizer's supply chain.

Figure 1-4: Cartography of Moderna's supply chain

Figure 1-5 depicts Johnson&Johnson's supply chain, and we can see that it is dispersed mostly in two regions, Europe, and the US with 5 locations each. Additionally, there are two companies located in India and one located in South Africa. Like in Pfizer and Moderna's case, those facilities served as manufacturing supports. As we can see Johnson&Johnson is also highly dependent on their North American and European location. This leaves Johnson&Johnson vulnerable in the eventuality of disruption in those locations.

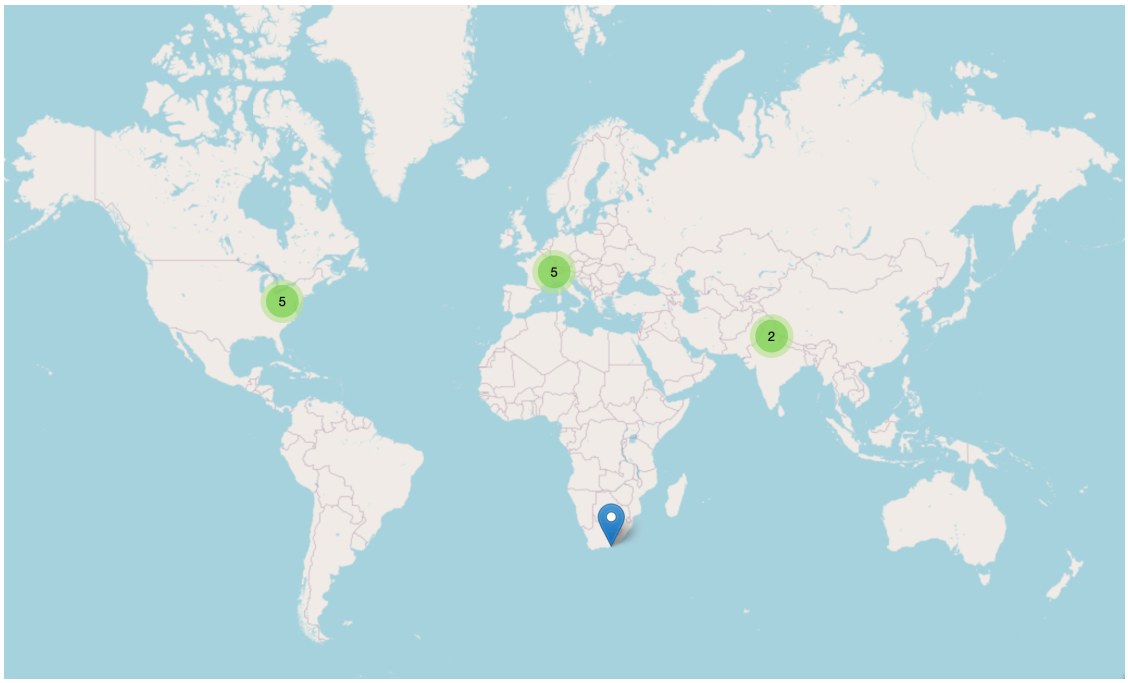


Figure 1-5: Cartography of Johnson&Johnson's supply chain

AstraZeneca has in total 24 locations to produce their Covid-19 vaccines. The majority of AstraZeneca's SC is in Europe with 10 companies (*Figure 1-6*). There remaining companies are in Japan (4), Thailand (2), Australia (2), Brazil (1), Argentina (1), Mexico (1) and the US (3). Those companies help in the manufacturing of the vaccine. As most of their supply chain is in Europe with little presence in other regions, AstraZeneca is also highly vulnerable to the eventuality of disruption in Europe.

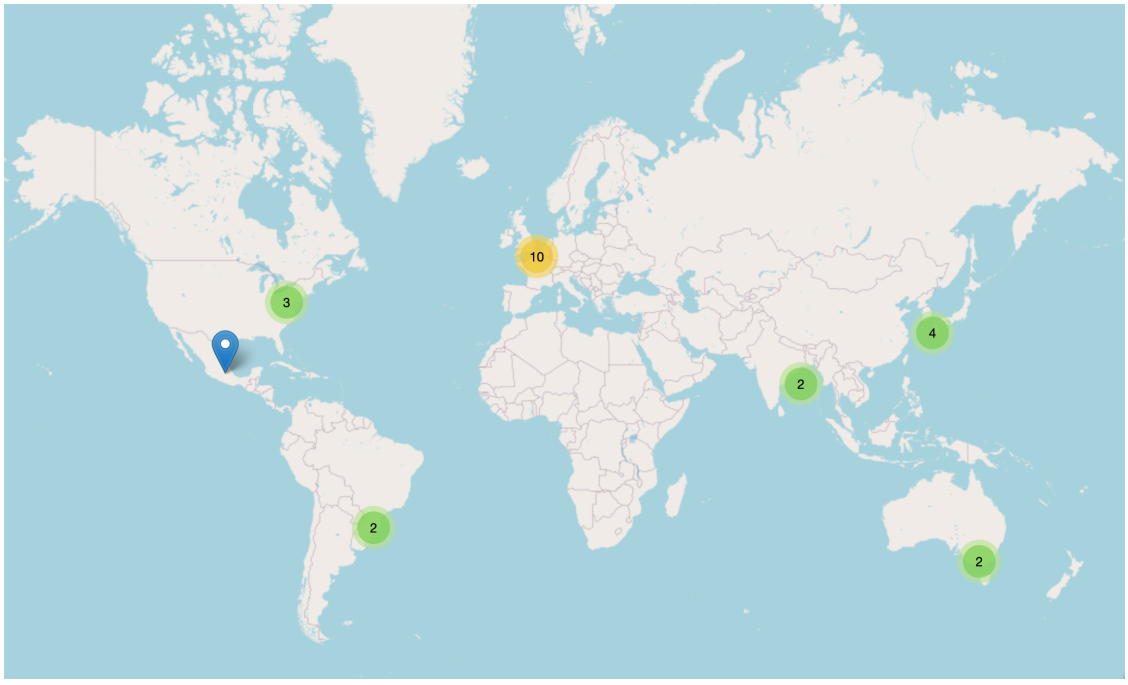


Figure 1-6: Cartography of AstraZeneca's supply chain

Figure 1-7 shows the location of Novavax's Covid-19 vaccine supply chain. In this map, we can see that Novavax has 9 facilities located in the US and 8 located in Europe. The map also shows that one company is in India, one in Japan and one in South Korea. All three are mostly used to increase the manufacturing capacity for the formulation stage. With their heavy reliance on the facilities in Europe and the US, Novavax is vulnerable to disruption if there were to have risk associated with one or both regions.

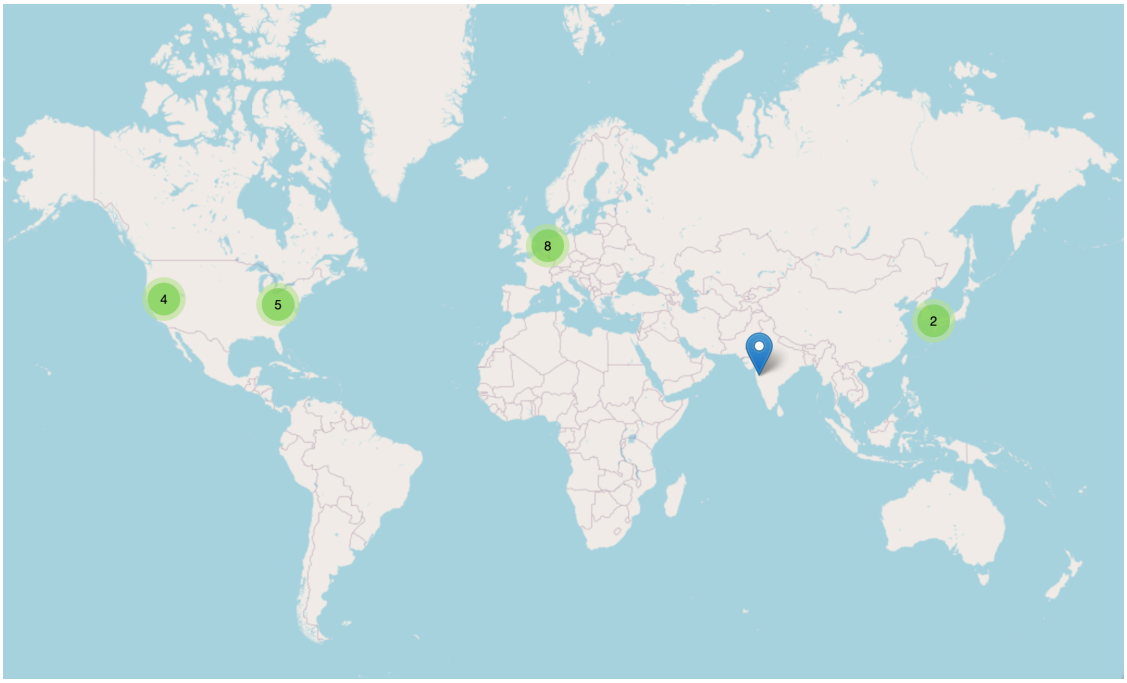


Figure 1-7: Cartography of Novavax's supply chain

Figure 1-8 shows the location of CureVac's Covid-19 vaccine supply chain. In this map, we can see that CureVac only has 9 facilities in their supply chain, and they are all located in Europe. By focusing its supply chain in only one region, CureVac is highly vulnerable to disruption, compared to the other vaccine producers, if there were to have risk associated with that region.

Figure 1-8: Cartography of CureVac's supply chain

Table 1-7 summarizes the result of classifying the Covid-19 vaccine producers. Most producers were classified as bi-regional. While Pfizer, AstraZeneca and CureVac upstream classifications remain the same, home region/host region-oriented and bi-regional respectively, the other producers' upstream classifications vary. As presented in the table below, Moderna's supply chain and fill & finish facilities are classified to be home region-oriented, its formulation is bi-regional, and its lipids classification is host region-oriented. Johnson&Johnson has three different classifications for its upstream dimensions, bi-regional for its whole SC, global for its formulation classification and home region-oriented for its fill classification. Like Johnson&Johnson, Novavax has also three upstream classifications. Novavax's supply chain is bi-regional with a global formulation classification and a home region-oriented lipids and fill classification. For their downstream classification, most of the classifiable producers are bi-regional except Moderna which has a home region-oriented classification.

Table 1-7: Classification of the six main Covid-19 vaccine producers

Brand	SC classification	Formulation classification	Lipids classification	Fill classification	Distribution classification
Pfizer	Home/host region-oriented	Home/host region-oriented	Home/host region-oriented	Home/host region-oriented	Bi-regional
Moderna	Home region-oriented	Bi-regional	Host region-oriented	Home region-oriented	Home region-oriented
Johnson&Johnson	Bi-regional	Global	n/a	Home region-oriented	Bi-regional
AstraZeneca	Bi-regional	Bi-regional	n/a	Bi-regional	unclassifiable
Novavax	Bi-regional	Global	Home region-oriented	Host region-oriented	Bi-regional
CureVac	Home region-oriented	Home region-oriented	n/a	Home region-oriented	n/a

1.5 Discussion

The first case of Covid-19 was reported on December 31, 2019, in China's Hubei province, and on March 11, 2020, the World Health Organization (WHO declared Covid-19 a global pandemic. The need for a vaccine was soon felt, and organizations around the world came together to develop a vaccine. With a vaccine developed in less than a year, there was a rapid need to produce and distribute as many vaccines as possible. Thus, emerged the Covid-19 vaccine supply chain, which was one of the most important value chains during the pandemic. To respond to the rapid demand for the Covid-19 vaccine, pharmaceutical companies came up with different strategies to increase their production capacity, i.e., expanding their operation either through partnerships, outsourcing and/or relocation production to existing vaccine manufacturing facilities.

The literature on the vaccine supply chain in time of pandemic is a subject that has been under-research, therefore limiting the amount of information how been acquired. This project aims to integrate the principles of international business classification, particularly geographical diversification, into the supply chain of Covid-19 vaccine producers. By implementing this approach to the supply chain, our objective was to gain a deeper understanding of the unique attributes and strategies employed by prominent vaccine manufacturers during the initial phase of the pandemic. Using Rugman & Verbeke's triad classification, we were able to categorize six Covid-19 vaccine producers based both on their production geographical location (upstream and their vaccine distribution (downstream. Thus, contributing to expanding the literature on the vaccine supply chain more specifically in a global pandemic context. With the literature, we formulated three hypotheses; H1 certain regions of the triad will show a higher concentration of facilities, H2 formulation facilities distribution will lead to a classification of home region-oriented using Rugman & Verbeke's classification and H3 downstream dimension will show that Covid-19 vaccine producers are depending on a certain region of the triad.

As mentioned previously, this project uses a quantitative method using R to answer the thesis stated above. Data regarding the geographic location of Pfizer, Moderna, Johnson&Johnson, AstraZeneca, Novavax and CureVac Covid-19 VSC and the doses distributed across North America, Europe, and Asia/Pacific. We use this data to recreate Rugman & Verbeke's classification using the ratio of companies and doses in each region of the triad. This would then give us the upstream and downstream classification of the six vaccine producers.

All the Covid-19 vaccine producers, mentioned in this paper, opted for either a home region-oriented or bi-regional, focused between the EU and North America, for their upstream strategy regardless of whether they did a partnership, used their existing facilities, or outsourced all or part of their production. This coincides the Rugman & Verbeke's findings, where the majority of MNEs, 320, were home oriented followed by 22 companies that were classified bi-regional. Our results also further support Qian et al. (2008) paper on regional diversification and firm performance. As shown by the results in both papers, firms in developed countries benefit from the available resources (e.g., facilities, governments and/or universities support) within developed countries to maximize their performance. This may explain why most of the companies in this paper were concentrated in the two "developed" regions with very little presence in another region.

From the result, we see that the participation in the Covid-19 VSC was not a global effort but an EU/north American effort as our results show a higher concentration of companies. This observation confirms our first hypothesis, as H1 stated that certain regions would have a higher concentration of companies than others. Most of the effort/production/supply chains were concentrated in those two

regions, which limits the scalability of production. Furthermore, certain vaccine producers located their supply close to their customers, meaning that looking at the upstream distribution of a producer we can identify with the region are their biggest consumers. Taking Pfizer as an example, this producer has most of its supply chain in North America and Europe, which also coincide with the region where Pfizer administrated most of their vaccine (23% and 41% respectively). This can be attributed to the fact that vaccine producers are for-profit organizations therefore their desire to be close to big customers. This proximity is a good thing under normal circumstances as it gives possible to respond to demand in a timely manner, however, this may create issues in the global effort to manage the spread of the virus in a pandemic, making some regions receive leftovers and have last pick which undermine the collective effort.

Similar to Rugman and Verbeke's finding, we arrived at the conclusion that very few companies are global. In terms of upstream, 3 out of the 6 vaccine producers are bi-regional oriented for their supply chain structure, 2 are home region-oriented and 1 is both home region-oriented and host region-oriented. Additionally, when looking in terms of the stage of production. We can see that for the formulation, 1 out of the 6 is home oriented based, 2 are bi-regional, 2 are global and 1 is both home region-oriented and host region-oriented. This observation contradicts our second hypothesis as only one producer formulation distribution was classified as home region oriented. In terms of lipid/adjuvant production, 1 is home region-oriented, 1 is host region-oriented and 1 is classified as both. In terms of downstream, 3 of the Covid-19 producers would be categorized as having a bi-regional distribution (Pfizer, Johnson&Johnson and Novavax), and 1 is home region-oriented (Moderna). For the two-remaining producer, AstraZeneca and CureVac, it was not possible to categorize their distribution with our data. Based on our findings, most producers are bi-regional for both upstream and downstream. As none of the Covid-19 vaccine producers had a downstream classification of global, our results support our last hypothesis. As mentioned in the literature review section, bi-regional companies as susceptible to risk associated with market concentration, dependency on regional factors, and limited knowledge.

Overall, this project has shown the benefits of applying a new methodology to analyze the vaccine supply chain. Applying geographical classification to categorize the Covid-19 vaccine supply chain can provide data-driven decision-making. Geographical classification provides a framework for collecting and analyzing data on vaccine production, distribution, and uptake within specific regions. This data can be used to identify trends, optimize allocation strategies, and make data-driven decisions. By understanding the geographic distribution of vaccine doses, vaccination rates, and population coverage, policymakers and organizations can make informed choices to optimize the allocation and utilization of vaccines. Furthermore, by categorizing the vaccine supply chain based on different regions, risks such as natural disasters, geopolitical tensions, or trade disruptions can be assessed more effectively. Companies and organizations can develop contingency plans, diversify supply sources, and establish alternative routes or storage facilities to mitigate potential disruptions. This enhances the overall resilience and robustness of the vaccine supply chain.

As with any project, ours has its limitations. Starting with the data collected, the datasets related to the supply chains were created based on information from June 2021, thus, making our classification outdated. Therefore, more recent data would give a better and more representative classification of the current situation of the Covid-19 vaccine producer structure. Furthermore, despite being able to find good results using their method of classification, we were able to assess two flaws in Rugman & Verbeke's classification. The first is related to the Pfizer case. Indeed, when a company has two headquarters, which was the case in Pfizer & BioNTech partnership, this makes the classification more complex and gives two contradicting results. The second flaw is related to the classification delimitation. It is indeed possible to fall at the edge of two classifications without being in either one.

This was the case when analyzing the downstream dimension for AstraZeneca. Future directions to this project would be to consider comparing different classifications and/or expand the scope to different vaccines, thus comparing pandemic VSC to non-pandemic VSC.

References

Ahmed, F., Ahmed, N., Pissarides, C., & Stiglitz, J. (2022). Why Inequality Could Spread Covid-19. The Lancet, 395(10230), 809-810.

Bollyky, T. J., & Bown, C. P. (2021). The Tragedy of Vaccine Nationalism: Only Cooperation Can End the Pandemic. Foreign Affairs.

Bollyky, T. J., Gostin, L. O., Hamburg, M. A., Piot, P., & Redlener, I. E. (2021). National Security and Global Health: Safeguarding Global Health in the 21st Century. Columbia University Press.

Brewer, N. T., Chapman, G. B., Rothman, A. J., Leask, J., & Kempe, A. (2017). Increasing Vaccination: Putting Psychological Science Into Action. Psychological Science in the Public Interest, 18(3), 149-207.

Contractor, F. J., Kumar, V., & Kundu, S. K. (2007). Nature of the relationship between international expansion and performance: The case of emerging market firms. Journal of World Business, 42(4), 401-417.

Dembek, Z. F., Smith, C. B., Wagner, M., & Dausey, D. J. (2021). Vaccine nationalism and the role of international organizations in the COVID-19 pandemic response. Global Public Health, 16(5), 685-699.

Dunning, J. H. (1988). The eclectic paradigm of international production: A restatement and some possible extensions. Journal of International Business Studies, 19(1), 1-31.

Evenett, S. J., Hoekman, B., Rocha, N., & Ruta, M. (2021). The COVID-19 vaccine production club: Will value chains temper nationalism? *Policy Research Working Papers*. https://doi.org/10.1596/1813-9450-9565

Fidler, D. P. (2021). COVID-19: The Politics of Vaccination. Foreign Affairs

Ghebreyesus, T. A., Moeti, M., Vaz, R. G., Bhatnagar, N., & Lal, S. (2021). Covid-19 Vaccination: Pathway to Equitable Global Access. World Health Organization.

Gostin, L. O. (2020). International Pandemic Response: Key Legal Issues. JAMA.

Gupta, S. (2020). COVID-19 Vaccine Supply Chains: Implications for Global Health Governance. BMJ Global Health.

Krause, P. R., Fleming, T. R., Longini, I. M., & Peto, R. (2021). Covid-19 vaccine trials should seek worthwhile efficacy. The Lancet, 397(10284), 1607-1609.

MacDonald, N. E., SAGE Working Group on Vaccine Hesitancy. (2015). Vaccine Hesitancy: Definition, Scope and Determinants. Vaccine, 33(34), 4161-4164.

Mullard, A. (2020, November 30). *How covid vaccines are being divvied up around the world*. Nature News. https://www.nature.com/articles/d41586-020-03370-6

Ngo, A. (2021). Global Vaccine Nationalism: A Threat to Collective Security. Harvard Kennedy School.

Oliver, S. E., Gargano, J. W., Marin, M., Wallace, M., Curran, K. G., Chamberland, M., … & Dooling, K. (2021). The Advisory Committee on Immunization Practices' Interim Recommendation for Use of Pfizer-BioNTech Covid-19 Vaccine in Adolescents Aged 12-15 Years—United States, May 2021. Morbidity and Mortality Weekly Report, 70(20), 749-752.

Ozawa, S., Stack, M. L., Publicover, N. G., & Walker, D. G. (2017). The economics of vaccine supply chains in developing countries. Vaccine, 35(17), 2279-2285.

Phelan, A. L., & Oluyomi, A. (2021). Covid-19 Vaccine Equity in the United States: A Rapid Review. Journal of Racial and Ethnic Health Disparities, 1-10.

Phelan, A. L., et al. (2020). What SARS-CoV-2 Reveals About the Global Health Security Agenda. The Lancet.

Pisani-Ferry, J., et al. (2021). A Proposal for the COVID-19 Vaccine Supply Chain. Project Syndicate.

Polack, F. P., et al. (2020). Safety and Efficacy of the BNT162b2 mRNA COVID-19 Vaccine. New England Journal of Medicine.

Riaz, M.M.A., Ahmad, U., Mohan, A. *et al.* Global impact of vaccine nationalism during COVID-19 pandemic. *Trop Med Health* **49**, 101 (2021). https://doi.org/10.1186/s41182-021-00394-0

Rugman, A. M., & Verbeke, A. (2004). A perspective on regional and global strategies of multinational enterprises. Journal of International Business Studies, 35(1), 3-18.

Smith, J. C., & Sivaramakrishnan, K. (2020). Covid-19 and vaccine logistics: Challenges and opportunities. Journal of Global Health, 10(2).

Smith, K. F. (2021). Global vaccine alliances and the role of vaccine diplomacy in addressing global challenges. The Lancet, 397(10294), 1034-1042.

Smith, M. J. (2020). "The Politics of a Covid-19 Vaccine." The New England Journal of Medicine, 383(22), e129.

Sridhar, D., Gurdasani, D., & Borelli, G. (2020). Will vaccine nationalism hinder efforts to control the COVID-19 pandemic? The Lancet, 395(10236), 1339-1341.

Thompson, M. (2021). Vaccine Nationalism: The Game of Strategic Ethics. Journal of Bioethical Inquiry.

Vazquez-Prokopec, G. M., Mercado, M. I. G., & Dumonteil, E. (2021). The Challenge of Vector-Borne Diseases in Underserved Regions: A Scoping Review. Frontiers in Public Health, 9, 613540.

Verma, S., Gokhale, M., & Bhatia, R. (2021). Global vaccine supply chains: Current status, challenges and opportunities. Vaccine, 39(7), 1016-1027.

World Bank. (2021). Global Economic Prospects. World Bank Group

World Health Organization. (2020). Guidelines on the management of vaccine supply chains during the Covid-19 pandemic. Geneva: World Health Organization.

Yadav, P., & Lyngdoh, T. I. (2020). Covid-19 vaccine supply chain: Challenges and strategies. Indian Journal of Public Health, 64(Supplement), S152-S156.

Zhang, Y., Li, Y., & Lu, J. (2021). From emergency to sustainability: Building resilient vaccine supply chains against Covid-19 and future pandemics. Globalization and Health, 17(1), 1-10.

Chapter 2: Econometric Approach to the Covid-19 Vaccine Production Supply Chain

Koné Aïchata Souleymane, M.Sc. student, HEC
Montreal Supervision:
Bélanger Valérie, HEC Montreal
Warin Thierry, HEC Montreal

HEC Montréal

Abstract

The COVID -19 pandemic has posed significant challenges to global health systems and requires rapid development, production, and distribution of effective vaccines. Ensuring a robust and efficient vaccine supply chain is essential to combat the spread of the virus. This study explores the use of an econometric model, panel data analysis, to understand and optimize the Covid-19 vaccine supply chain. Panel data analysis is a valuable method for analyzing the Covid-19 vaccine supply chain. It integrates data from multiple sources to understand the complex dynamics associated with vaccine production, distribution, and delivery. By quantifying the impact of factors such as production capacity, distribution networks, transportation infrastructure, regulatory frameworks, and demand patterns, panel data analysis provides information that can inform decision making, resource allocation, logistics planning, and risk management to improve the efficiency and responsiveness of the vaccine supply chain and ensure equitable and timely distribution of vaccines.

2.1 Introduction

The Covid-19 pandemic has presented a global health crisis, requiring the rapid development, production, and distribution of vaccines on a massive scale. Efficient and equitable allocation of vaccines is crucial to counter the spread of the virus and mitigating its impact on public health and economies. Econometric models offer a valuable approach to analyzing and optimizing the complex dynamics of the vaccine supply chain. These models leverage economic theory and statistical techniques to quantify the relationships between key variables such as production capacity, logistics, demand dynamics, and policy interventions. By providing quantitative insights into the supply chain, econometric models can inform decision-making processes, enhance distribution strategies, and address potential bottlenecks. This research aims to explore the application of econometric models in analyzing and optimizing the Covid-19 vaccine supply chain, contributing evidence-based strategies for effective vaccine allocation and distribution.

The objective of this project is to analyze the factors influencing the Covid-19 vaccine supply chain of the two main vaccine producers (Pfizer and Moderna), develop econometric models that capture the relationships between key variables in the vaccine supply chain and quantify the impact of various factors on vaccine availability and distribution in a country, such as present vaccine production facilities, policies regarding the approval of the vaccine and more.

The application of econometric models in the context of the Covid-19 vaccine supply chain can offer valuable insights into various aspects of the distribution process. For instance, econometric models can be used to estimate the impact of different factors on vaccine availability, such as production capacity expansions or transportation infrastructure improvements. These models can also help identify potential bottlenecks in the supply chain, allowing policymakers and stakeholders to address critical issues and ensure a smooth and efficient distribution process. By quantifying the relationships between production capacity, logistics, and demand dynamics, econometric models can provide valuable information for decision-making related to resource allocation, vaccine distribution strategies, and policy interventions.

Furthermore, econometric models can assist in forecasting vaccine demand based on various variables, including population demographics, epidemiological trends, and vaccine efficacy data. This forecasting capability can be essential for planning and resource allocation at local, regional, and national levels, enabling proactive measures to meet future demand and optimize vaccine distribution. Moreover, econometric models can evaluate the effectiveness of different policy interventions, such as export restrictions, priority allocation strategies, or public-private partnerships, providing insights into their potential impacts on the vaccine supply chain.

By leveraging econometric models, policymakers, public health officials, and vaccine manufacturers can make informed decisions to enhance the efficiency and equity of the Covid-19 vaccine supply chain. These evidence-based strategies are crucial for achieving global vaccination goals and minimizing the socio-economic impact of the pandemic.

2.2 Literature Review

The equitable distribution of vaccines is one of the challenges in the vaccine supply chain and a critical aspect of public health, particularly in times of global pandemics. In recent years, the world has witnessed the emergence of numerous studies focusing on vaccine distribution equity. This literature review aims to summarize and analyze the key findings and trends from relevant research articles, highlighting the importance of addressing factors influencing vaccine distribution.

Several studies have highlighted the presence of significant factors affecting vaccine distribution across different populations and regions. Factors such as socioeconomic status, race, ethnicity, and geographical location have been identified as key determinants of vaccine access and uptake (Hotez et al., 2021; Delamater et al., 2021). The literature suggests that marginalized and underserved communities face higher barriers to vaccine distribution, leading to increased vulnerability to infectious diseases. Additionally, numerous barriers contribute to the inequitable distribution of vaccines. Lack of healthcare infrastructure, limited vaccine supply, inadequate funding, and logistical challenges are commonly identified barriers (Hotez et al., 2021; Mehrotra et al., 2021). Language and cultural barriers, vaccine hesitancy, and misinformation also play a significant role in exacerbating disparities in vaccine access (Leask et al., 2018; Karafillakis et al., 2019).

The literature offers various strategies and interventions to address vaccine distribution disparities. Targeted vaccination campaigns, mobile clinics, and community-based outreach programs have proven effective in reaching underserved populations (Delamater et al., 2021; Asch et al., 2021). Improving vaccine supply chains, enhancing healthcare infrastructure, and implementing effective communication campaigns are crucial steps in achieving equitable distribution (Hotez et al., 2021; O'Leary et al., 2021). Additionally, policy interventions such as vaccine mandates and subsidies have been proposed to ensure equitable access to vaccines (Mehrotra et al., 2021; Phelan et al., 2020). The literature also emphasizes the importance of global collaboration and solidarity in addressing vaccine distribution equity.

International organizations, such as the World Health Organization (WHO) and COVAX, have been instrumental in facilitating equitable access to vaccines (Bollyky et al., 2021). Lessons learned from previous pandemics, such as the H1N1 influenza and Ebola outbreaks, have informed current strategies for improving vaccine distribution equity (O'Leary et al., 2021; Moon et al., 2020).

Econometric Models: Foundations and Applications

Econometric models play a crucial role in empirical economic analysis, enabling researchers to quantify and examine the relationships between economic variables. This literature review aims to provide an overview of key econometric models developed over the years and their applications in various economic domains. The review begins with a discussion of foundation models and then explores recent advancements in econometric modelling techniques and provides the contributions of different studies and researchers.

Foundation models enable economists to quantify the effects of various factors on economic outcomes, estimate causal relationships, and make predictions. Two key econometric models that have shaped the field are simple linear regression and multiple regression. Simple linear regression, introduced by Galton (1886), establishes a linear association between a dependent variable and a single independent variable. Multiple regression, pioneered by Pearson (1895), extends this model by incorporating multiple independent variables to explain the variation in the dependent variable. These models have played a pivotal role in econometric analysis and continue to serve as the building blocks for more sophisticated econometric techniques.

Simple Linear Regression Simple linear regression is a fundamental econometric model that establishes a linear relationship between a dependent variable and a single independent variable. This model, first introduced by Sir Francis Galton (1886) has since been widely used in econometric analysis. The model can be expressed as:

$$Y = \alpha + \beta X + \epsilon,$$

where Y represents the dependent variable, X represents the independent variable, α and β denote the intercept and slope coefficients, respectively, and ϵ represents the error term. Simple linear regression enables researchers to estimate the causal impact of the independent variable on the dependent variable by quantifying the magnitude and direction of the relationship.

Multiple Regression Multiple regression expands upon the simple linear regression model by incorporating multiple independent variables to explain the variation in a dependent variable. Karl Pearson's work in 1895 laid the foundation for multiple regression, which has since become a cornerstone of econometric analysis (Pearson, 1895). The multiple regression model can be expressed as:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \epsilon,$$

where X_1, X_2, …, X_n are the independent variables, β_1, β_2, …, β_n are the corresponding coefficients, and α and ε represent the intercept and error term, respectively. Multiple regression allows researchers to simultaneously consider the effects of multiple independent variables on the dependent variable, providing a more comprehensive analysis of the relationships between economic variables.

The significance of these foundation econometric models lies in their ability to estimate the parameters of interest, assess statistical significance, and make predictions based on empirical data. By employing

these models, researchers can quantitatively analyze economic phenomena, test hypotheses, and inform policy decisions.

Advancements in Econometric Modeling Techniques

Econometric modelling techniques have continuously evolved to address the complexities and challenges encountered in empirical economic analysis. This section highlights key advancements in econometric modelling and their contributions to the field.

Time Series Analysis Time series analysis focuses on modelling and forecasting variables observed over time. Autoregressive Integrated Moving Average (ARIMA) models have been widely employed in this context. Box and Jenkins (1970) introduced the ARIMA model, which incorporates autoregressive (AR), moving average (MA), and differencing components to capture the temporal dynamics of the data (Box & Jenkins, 1970). This approach has been instrumental in analyzing and predicting economic time series, enabling researchers to uncover patterns, seasonality, and trends in data. The advantages of ARIMA models lie in their ability to handle non-stationary data and provide accurate short-term forecasts. However, ARIMA models may struggle with complex nonlinear relationships and may not capture long-term trends effectively (Box & Jenkins, 1970).

Panel Data Analysis Panel data analysis deals with data collected over multiple individuals or entities over time. Fixed Effects (FE) and Random Effects (RE) models have been pivotal in this area. Mundlak (1978) introduced the random effects estimator, which effectively captures unobserved heterogeneity and time-varying effects in panel data (Mundlak, 1978). These models allow researchers to control for individual-specific characteristics, test for individual heterogeneity, and estimate the effects of time-varying variables, enhancing the understanding of complex economic phenomena. FE models control for individual-specific characteristics, while RE models capture time-varying effects. Panel data analysis includes the ability to control for unobserved heterogeneity, estimate dynamic effects, and improve efficiency. However, these models assume specific assumptions, such as no individual-specific omitted variables in FE models, which can be challenging to verify and may lead to biased estimates (Mundlak, 1978).

Instrumental Variable (IV) Regression Instrumental Variable (IV) Regression addresses endogeneity issues and helps establish causal relationships between variables. The instrumental variable approach was initially developed by Wright in 1928 (Wright, 1928) and expanded upon by Frisch in 1936 (Frisch, 1936). IV regression uses instrumental variables that are correlated with the endogenous variables of interest but unrelated to the dependent variable. The two-stage least squares (2SLS) method is commonly employed in IV regression. This technique has been crucial in overcoming endogeneity biases and providing unbiased estimates of causal effects, particularly in studies involving observational data. The advantages of IV regression lie in its potential to provide consistent estimates of causal effects when endogeneity is present. However, IV regression relies on the assumption that the instrumental variables are valid and fulfill certain criteria, which can be difficult to satisfy in practice. Additionally, IV regression may suffer from weak instrument problems and can introduce additional sources of bias (Wright, 1928; Frisch, 1936).

Machine Learning and Big Data The integration of machine learning techniques and the availability of big data have revolutionized econometric modelling. Machine learning algorithms, such as neural networks, random forests, and support vector machines, offer powerful tools for predictive modelling and pattern recognition. These approaches can handle large and complex datasets and capture intricate nonlinear relationships. The advantages of machine learning lie in their flexibility, adaptability, and

ability to uncover complex patterns in data. However, machine learning techniques may lack interpretability, making it challenging to provide a clear economic interpretation of the results. Furthermore, they may be prone to overfitting when the data is noisy or when the model complexity is not properly controlled (Bai et al., 2018).

Hypotheses

From the literature review, we were able to formulate two hypotheses:

- **H1**: The presence of the vaccine producers' supply chain in a country affects the number of vaccine dose administrated in that country.
- **H2**: A higher number of pharmaceutical companies present in a country has a positive impact on the number of doses administrated.

2.3 Methodology

This study uses the R language to develop an econometric model for the Covid-19 vaccine supply chain. The data on Pfizer and Moderna Covid-19 vaccine and its supply chain will be collected from various sources creating a total of three datasets. We decided to collect data for the last 3 years from December 2020 to March 2023. These data sets were created using information on the number of doses of vaccine administrated by Pfizer, Moderna and their total using "Covid vaccine doses by manufacturer" by Our World in Data(2023), whether the vaccine was approved from the Covid-19 Vaccine tracker. We also required data on the number of Pfizer and/or Moderna supply chains present in the country, the number of formulation facilities, the number of lipids producing facilities, and the number of fill and finish facilities, which was acquired and derived from the article of Bown and Bollyky titled "How Covid-19 vaccine supply chains emerged in the midst of a pandemic" (2021). The last data we need was the number of the top 1000 pharmaceutical companies present in the country from the Orbis database.

We then created three panel dataframes which includes the variables as shown in *Table 2-1*. We chose to create our dataset in a panel format as we had information on the number of doses administrated throughout a period of 3 years for 43 countries.

Table 2-1: Variable description Pfizer, Moderna and the Covid-19 vaccine producer panel datasets

Variable	Description
id	Country id
Country	Country the vaccine was administrated
year	Year of administration
dose_total	Sum of the Pfizer dose and Moderna dose administrated
reg_moderna	binary, 1 if Moderna vaccine is approved in the country, and 0 is not
number_cov_chain_moderna	Number of Moderna supply chain present in the country
reg_pfizer	binary, 1 if Pfizer vaccine is approved in the country, and 0 is not
number_cov_chain_pfizer	the number of Pfizer supply chain present in the country
number_formulation	Number of formulation facilities in the country
number_lipids	Number of lipids production facilities in the country
number_fill	Number of fill and finish facilities in the country
number_top_1000	Number of top 1000 facilities in the country

rug_ver_pfizer	binary, 1 if the country is in Pfizer's VSC classification region (based on Rugman & Verbeke classification), and 0 if not
rug_ver_moderna	binary, 1 if the country is in Moderna's VCS classification region (based on Rugman & Verbeke classification), and 0 if not
continent	Continent the county is in

Model

Once our datasets are created, we can formulate a linear regression model for the number of vaccine doses administered. We decided to use the simple linear regression model despite having panel data because our initial panel data was over a short period of time to be affected by time effect. Despite this decision, we still decided to perform a robustness test. To create our linear models, we first needed to create dummy variables for our three datasets using the id variable. This resulted in 35 additional variables one for each of the 35 countries of the datasets. For the linear regression, we decided to use the number of doses administrated as our dependable variable and all the variables presented in *Table 2-1* in addition to the dummy variable generated as independent variables. This process was repeated for all three datasets, thus creating three simple linear regressions for each dataset. Model 1 examines the influence of the Pfizer and Moderna supply chain variable on distribution, through the sum of vaccine doses from both producers across 43 countries. Models 2 and 3 examine the influence of supply chain variables on distribution focusing solely on Moderna for model 2 and Pfizer for model 3. The equation of the three models is shown in *Table 2-2*.

Table 2-2: Linear equation of Model 1,2 & 3

Model	Regression
Model 1 - Total Pfizer & Moderna	$dose_{(total)} = \alpha + \beta_1 reg_{(pfizer)} + \beta_2 reg_{(moderna)} + \beta_3 number_{(cov\ chain\ pfizer)} + \beta_4 number_{(cove\ chain\ moderna)} + \beta_5 number_{(formulation)} + \beta_6 number_{(lipids)} + \beta_7 number_{(fill)} + \beta_8 number_{(top\ 1000)} + \beta_9 rugver_{(pfizer)} + \beta_{10} rugver_{(moderna)} + \beta_{11} id_{(1)} + \cdots + \beta_{46} id_{(35)} + \varepsilon$
Model 2 - Moderna	$dose_{(pfizer)} = \alpha + \beta_1 reg_{(pfizer)} + \beta_2 number_{(cov\ chain\ pfizer)} + \beta_3 number_{(formulation)} + \beta_4 number_{(lipids)} + \beta_5 number_{(fill)} + \beta_6 number_{(top\ 1000)} + \beta_7 id_{(1)} + \cdots + \beta_{42} id_{(35)} + \varepsilon$
Model 3 - Pfizer	$dose_{(moderna)} = \alpha + \beta_1 reg_{(moderna)} + \beta_2 number_{(cove\ chain\ moderna)} + \beta_3 number_{(formulation)} + \beta_4 number_{(lipids)} + \beta_5 number_{(fill)} + \beta_6 number_{(top\ 1000)} + \beta_7 id_{(1)} + \cdots + \beta_{42} id_{(35)} + \varepsilon$

Test of Robustness

Once our panel data created, the next step is to perform a validity test of our data. We looked at the heterogeneity in our three datasets using the number of doses administrated as dependente variable and the country as an independent variable. Once the datasets were validated, we tested our dataset for fixed and random effects using the Hausman test. This step is important as it dictates the nature of the model we need to use. Considering the results of the Hausman test, all three datasets had a p-value lower than 5% therefore we need a fixed-effects model. With this result, the question is now to know

whether the fixed effects are cross-sectional, or time-based. Indeed, when running a fixed effects model, the question then arises as to whether we should test for time-fixed effects.

In addition to testing for time-fixed effect, we also need to check if we have an unbalanced panel (missing values. We ran a Breusch-Pagan test for unbalanced panels, which also confirmed the need for time-fixed effects.

Now that we know for certain that our data required a fixed effect model, we need to check for contemporaneous correlation (or cross-sectional dependence. Based on the following two tests: (1 Breusch-Pagan and (2 Pesaran, we did not find cross-sectional dependence. We also tested for serial correlation using the Breusch-Godfrey/Wooldridge test, which did not highlight the presence of serial correlation in idiosyncratic errors. To go a little further, we investigated the presence or not of a unit root (i.e., non-stationary. If the series is non-stationary, then the generalized method of moments (GMM estimators would be perfect candidates.

Here, there is no unit root, but this does not exclude GMM estimators anyway. Indeed, we need to check for the potential presence of heteroskedasticity. Based on the Breusch-Pagan test, we highlight the presence of heteroskedasticity. To correct for heteroskedasticity, we need to promote an estimation technique with a robust covariance matrix. To summarize all these tests, we do not have serial and cross-sectional dependence, with a non-stationary and homoscedastic unbalanced dataset.

For the results of the tests, we were able to proceed with our models. We decided to use both Beck and Katz model and Arellano model to identify the key variable influencing the doses of the Covid-19 vaccine administrated in a country. Model 1 looks at the influence of the supply chain variable of Pfizer and Moderna on the distribution, through the sum of doses of vaccine of the two producers in all 43 countries. While Models 2 and 3 look at the influence of supply chain variables on the distribution, the number of doses administered in all 43 countries focused solely on Moderna for Model 2 and Pfizer for Model 3. Once we created the model for all 43 countries, we wanted to model for different geographic regions and compare their result. Thus, five additional models were derived from Model 1. Model 1.1 test for the European region with 31 countries, Model 1.2 test for the North American region with two countries, Model 1.3 test for the South American region with five countries, Model 1.4 test for the Asian region with four countries and Model 1.5 test for the African region with one country. The same geographic separation was done for Model 2 and Model 3, creating a total of 15 models using both Beck and Katz model and Arellano model.

2.4 Results

In the following section, the results of our models will be discussed and analyzed. The results are presented in three parts: Model 1, Model 2 and Model 3. For each model, a description of the results will be provided, with additional tables and graphs.

Model 1 - Total Pfizer & Moderna

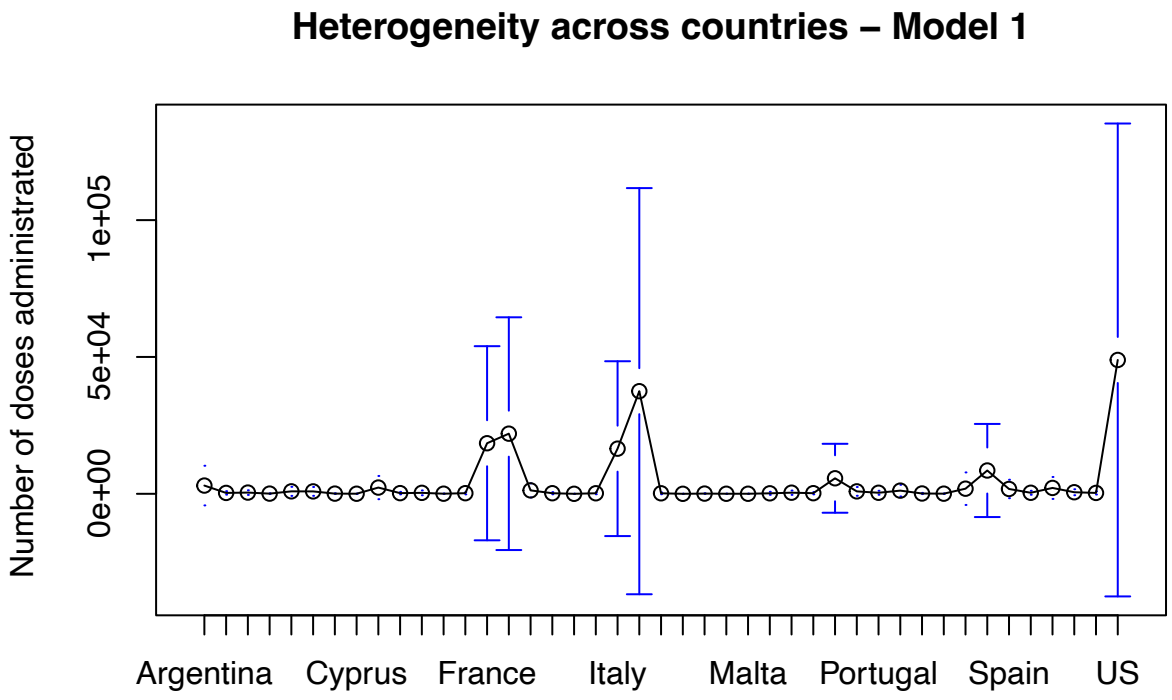
Table 2-3 is the statistical summary of our first dataset, which has both information on Pfizer and Moderna supply chain and the total of their doses administrated from December 2020 to March 2023, representing 172 observations or 34 countries. The average total number of doses administrated (dose_total) is 4,155.6 with a standard deviation of 14,922.2. The average number of Covid-19 vaccine formulation facilities per country is 0.9, the average number of Covid-19 vaccine lipids prod-

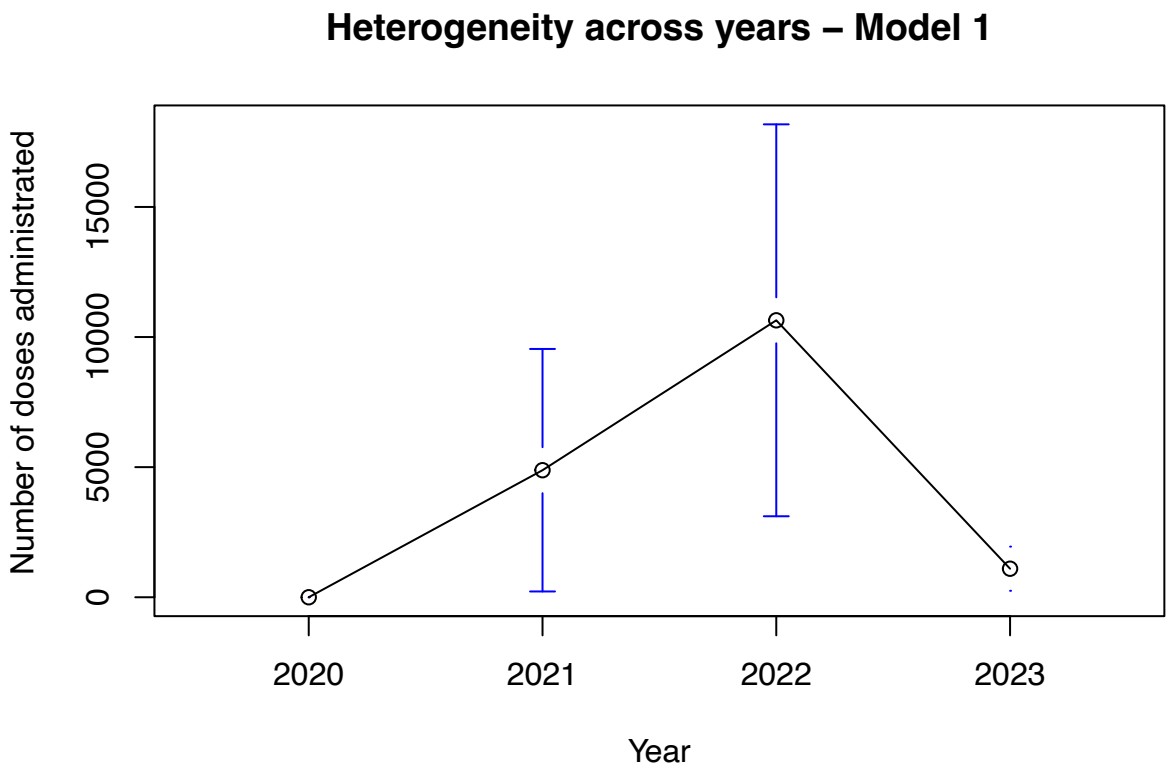
uction is 0.3 and the average number of Covid-19 fill & finish facilities is 0.7 with a standard deviation of 2.3, 1.1 and 2 respectively. Furthermore, the average number of Moderna Covid-19 SC facilities per country is 0.53 while the average is 0.5 for Pfizer.

Table 2-3: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
id	172	22.0	12.4	1	43
year	172	2,021.5	1.1	2,020	2,023
dose_total	172	4,155.6	14,922.2	0.0	105,000.0
reg_moderna	172	0.7	0.5	0	1
number_cov_chain_moderna	172	0.3	1.2	0	8
reg_pfizer	172	0.9	0.2	0	1
number_cov_chain_pfizer	172	0.5	1.8	0	10
number_formulation	172	0.9	2.3	0	13
number_lipids	172	0.3	1.1	0	7
number_fill	172	0.7	2.0	0	12
number_top_1000	172	10.9	25.9	0	163
rug_ver_pfizer	172	0.8	0.4	0	1
rug_ver_moderna	172	0.05	0.2	0	1

All the commands and algorithms are coded in R 4.2.2




Heterogeneity across years – Model 1
Number of doses administrated
0
5000
10000
15000
2020
2021
2022
2023
Year

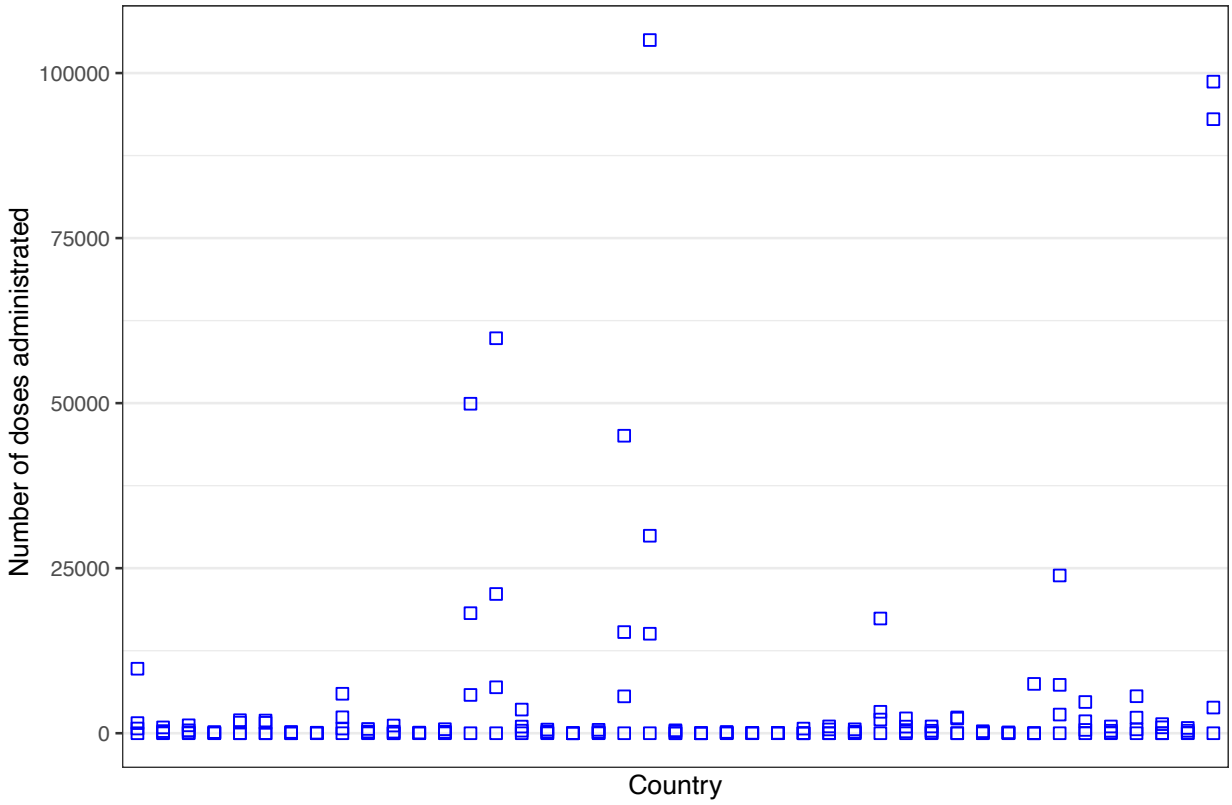

Number of doses administrated
100000
75000
50000
25000
0
Country

Table 2-4 shows the regression results of Model 1. We can see that only one variable is significant. The id_21 is significant at the 1 percent level for Model 1. Additionally, this variable has a positive impact on the number of doses administrated. *Table 2-5* is the result of the test of robustness for Model 1. We can see that in Beck & Katz model only two of our 10 variable was significant. The number of fill & finish facilities in the country is significant at the 10 percent level and the variable of Pfizer Rugman's classification (rug_ver_pfizer) is significant at the 5 percent level. Additionally, the presence of the fill & finish facility has a positive influence on the number of doses administrated while the rug_ver_pfizer variable has a negative influence. On the other hand, Arellano & Bond's model found only one significant variable. According to *Table 2-5*, the lag of the number of Moderna SC facilities present in the country is significant at the 1 percent level and has a negative impact on the number of doses administrated.

Table 2-4: Linear Regression Results – Model 1

	Number of Vaccine dose administrated
reg_pfizer	6,375,642,231.00 (−3,023,044,179.00, 15,774,328,640.00)
reg_moderna	615,204,446.00 (−4,092,858,272.00, 5,323,267,164.00)
number_cov_chain_pfizer	6,548,827,772.00 (−3,518,162,457.00, 16,615,818,000.00)
number_cov_chain_moderna	3,768,211,504.00 (−2,168,253,356.00, 9,704,676,365.00)
number_formulation	−5,855,759,578.00 (−14,066,920,135.00, 2,355,400,979.00)
number_lipids	305,551,317.00 (−13,695,870,500.00, 14,306,973,134.00)
number_fill	2,258,931,190.00 (−4,470,875,442.00, 8,988,737,822.00)
number_top_1000	253,942,964.00 (−320,555,582.00, 828,441,510.00)
rug_ver_pfizer	−3,410,476,735.00 (−15,495,245,795.00, 8,674,292,326.00)
rug_ver_moderna	−17,506,598,179.00 (−96,295,700,969.00, 61,282,504,611.00)
id_1	9,362,363,440.00 (−9,659,205,121.00, 28,383,932,000.00)
id_2	1,755,851,131.00 (−12,332,016,311.00, 15,843,718,572.00)
id_3	1,085,589,820.00 (−22,954,674,506.00, 25,125,854,146.00)
id_4	3,199,042,425.00 (−11,739,046,445.00, 18,137,131,294.00)
id_5	23,245,751,362.00 (−51,916,906,694.00, 98,408,409,418.00)
id_6	1,322,251,129.00 (−13,461,225,801.00, 16,105,728,059.00)
id_7	3,268,471,230.00 (−11,660,701,868.00, 18,197,644,328.00)
id_8	3,733,077,656.00 (−11,174,474,475.00, 18,640,629,787.00)
id_9	11,582,272,246.00 (−4,983,340,305.00, 28,147,884,797.00)
id_10	2,427,708,020.00 (−12,806,843,926.00, 17,662,259,966.00)
id_11	2,446,071,374.00 (−12,105,358,128.00, 16,997,500,876.00)
id_12	3,740,541,231.00 (−11,167,010,900.00, 18,648,093,362.00)
id_13	2,911,569,698.00 (−12,127,236,667.00, 17,950,376,062.00)
id_14	−10,266,991,965.00 (−43,747,749,923.00, 23,213,765,993.00)
id_15	1,658,003,377.00 (−33,898,580,593.00, 37,214,587,347.00)
id_16	2,577,666,546.00 (−11,494,007,055.00, 16,649,340,146.00)
id_17	3,155,642,722.00 (−11,817,429,927.00, 18,128,715,371.00)
id_18	3,552,868,766.00 (−11,317,038,725.00, 18,422,776,258.00)
id_19	−938,631,553.00 (−21,642,353,006.00, 19,765,089,899.00)
id_20	2,217,578,050.00 (−22,458,495,296.00, 26,893,651,395.00)
id_21	40,641,134,386.00*** (21,730,533,221.00, 59,551,735,552.00)
id_22	3,885,503,036.00 (−11,022,049,096.00, 18,793,055,167.00)
id_23	3,706,602,640.00 (−11,200,949,491.00, 18,614,154,771.00)
id_24	3,778,091,621.00 (−11,129,460,510.00, 18,685,643,752.00)
id_25	3,725,379,148.00 (−11,182,172,983.00, 18,632,931,279.00)
id_26	3,724,731,533.00 (−11,182,820,598.00, 18,632,283,664.00)
id_27	2,064,638,112.00 (−12,333,536,632.00, 16,462,812,856.00)
id_28	19,177,707,927.00 (−9,746,142,937.00, 48,101,558,792.00)
id_29	3,672,378,019.00 (−11,234,922,588.00, 18,579,678,626.00)
id_30	7,561,414,297.00 (−6,836,760,447.00, 21,959,589,040.00)
id_31	3,062,665,064.00 (−12,171,886,882.00, 18,297,217,010.00)
id_32	3,828,144,329.00 (−11,079,156,278.00, 18,735,444,936.00)
id_33	4,859,657,471.00 (−10,047,894,660.00, 19,767,209,602.00)
id_34	3,821,463,455.00 (−11,086,088,676.00, 18,729,015,586.00)
id_35	3,245,812,772.00 (−11,683,360,326.00, 18,174,985,870.00)
Constant	−7,132,099,789.00 (−17,857,998,089.00, 3,593,798,511.00)
Time Fixed Effects	No

Notes:

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.
Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

Table 2-5: Robustness Regression Results – Model 1

	Number of Vaccine dose administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_pfizer	7,538.64 (−4,621.22, 19,698.50)	
reg_moderna	2,641.06 (−1,463.91, 6,746.04)	
number_cov_chain_pfizer	−2,083.85 (−9,021.78, 4,854.08)	
number_cov_chain_moderna	−1,788.50 (−12,192.79, 8,615.80)	
number_formulation	1,289.29 (−1,205.77, 3,784.36)	
number_lipids	1,557.31 (−5,026.57, 8,141.19)	
number_fill	6,149.27* (23.26, 12,275.28)	
number_top_1000	−58.63 (−484.47, 367.20)	
rug_ver_pfizer	−6,242.70** (−11,296.20, −1,189.20)	
rug_ver_moderna	−1,608.40 (−13,924.58, 10,707.77)	
lag(reg_pfizer, 1)		2,111.17 (−866.95, 5,089.29)
lag(reg_moderna, 1)		1,221.80 (−276.24, 2,719.85)
lag(number_cov_chain_pfizer, 1)		−5,611.23 (−20,013.35, 8,790.90)
lag(number_cov_chain_moderna, 1)		−14,811.30*** (−25,745.25, −3,877.36)
lag(number_formulation, 1)		3,692.92 (−2,314.08, 9,699.92)
lag(number_lipids, 1)		−7,800.16 (−27,244.65, 11,644.33)
lag(number_fill, 1)		11,543.29 (−3,058.98, 26,145.56)
lag(number_top_1000, 1)		237.43 (−62.18, 537.04)
Constant	−3,780.17 (−14,105.89, 6,545.54)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.
Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

When modelling Moderna and Pfizer vaccine distribution in 31 European countries, Beck & Katz's model found that the number of fill & finish facilities in the country is significant at the 5 percent level with a positive impact on the number of vaccines administrated. Arellano & Bond's results state that 4 variables are significant. The lag of Pfizer vaccine approval in the country and the lag of having a Pfizer supply chain facility in the country are both significant at the 1 percent level. The lag of Pfizer vaccine approval has a negative impact while the lag of having a Pfizer supply chain facility has a positive impact. Additionally, the lag of having Moderna supply chain facilities and the lag of having a formulation facility in the country is significant at the 5 percent level, with a positive and negative influence respectively. The results for the North American region of Moderna and Pfizer vaccine distribution show that for both Beck & Katz and Arellano & Bond models, none of the variables were significant. The result of modelling Moderna and Pfizer vaccine distribution in South American countries shows that for both models, Beck & Katz, and Arellano & Bond, none of the variables were significant. The result of model 1.4, representing the modelling of Moderna and Pfizer vaccine distribution in 4 Asian countries shows that Beck & Katz model found no significant variable. On the other hand, Arellano & Bond's model found that the lag of Pfizer vaccine approval in the country, the lag of having Moderna supply chain facilities in the country and the lag of having fill & finish facilities in the country are both significant at the 1 percent level. Lastly, the result of the African region shows that Beck & Katz found no significant variable. However, Arellano & Bond found that the lag of Pfizer and Moderna vaccine approval is significant at the 1 percent level *(Appendix A to E)*.

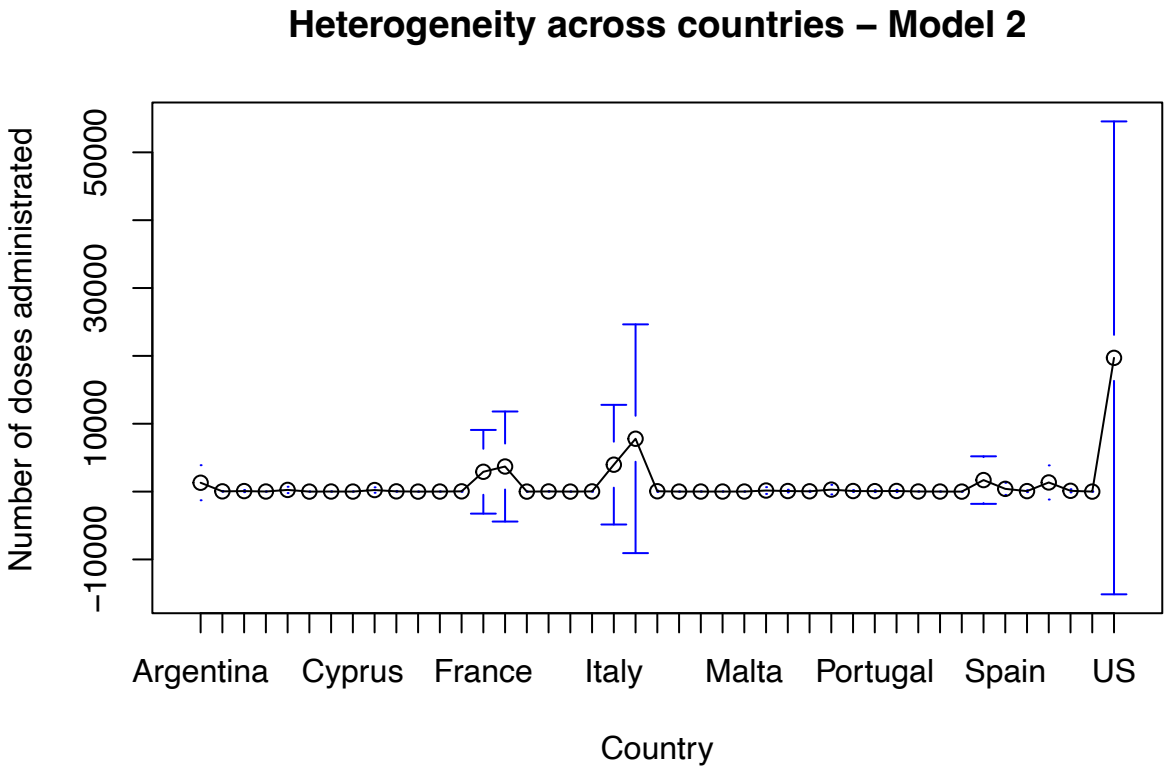
Model 2 - Moderna

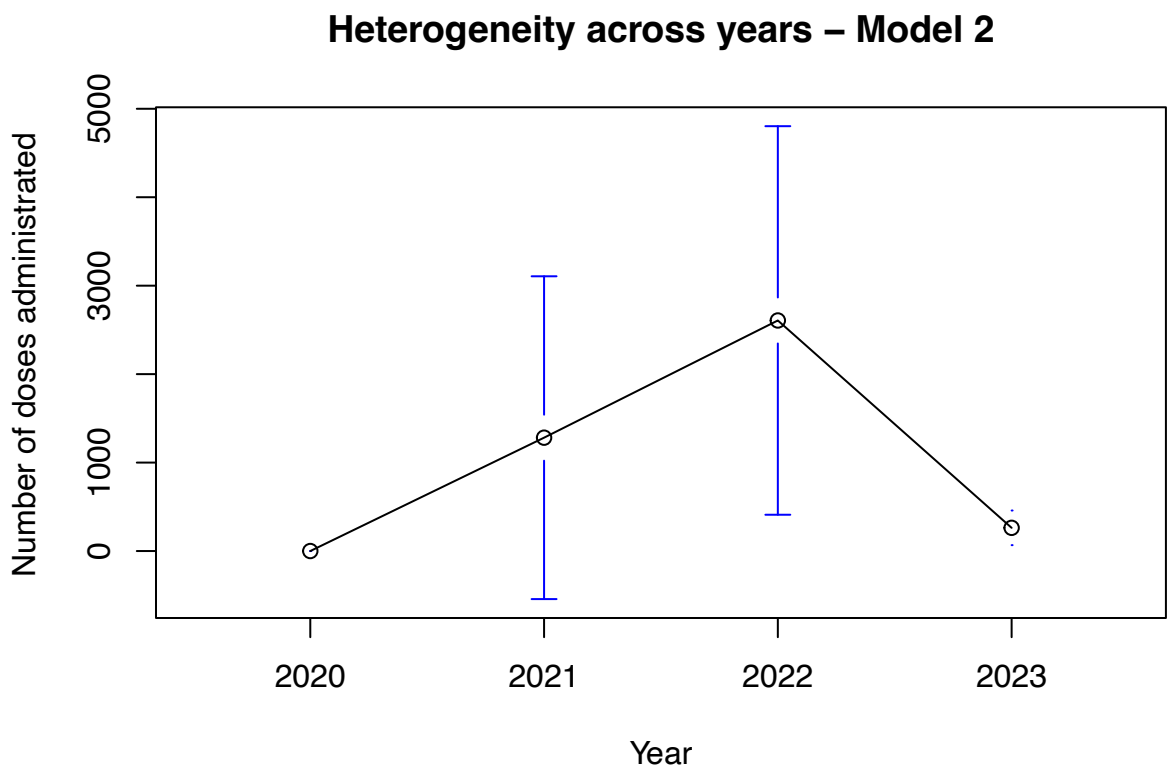
Table 2-6 is the statistical summary of our second dataset, which is for Moderna's supply chain and the number of Moderna vaccine doses administrated from December 2020 to March 2023, representing 172 observations or 34 countries. The average total number of doses administrated (dose_moderna is 1,037.6 with a standard deviation of 4,722.2. The average number of Covid-19 vaccine formulation facilities per country is 0.9, the average number of Covid-19 vaccine lipids production is 0.3 and the average number of Covid-19 fill & finish facilities is 0.7 with a standard deviation of 2.3, 1.1 and 2 respectively. Furthermore, the average number of Moderna Covid-19 SC facilities per country is 0.3 with a standard deviation of 1.2.

Table 2-6: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
id	172	22.0	12.4	1	43
year	172	2,021.5	1.1	2,020	2,023
reg_moderna	172	0.7	0.5	0	1
dose_moderna	172	1,037.6	4,722.2	0.0	38,837.9
number_cov_chain_moderna	172	0.3	1.2	0	8
number_formulation	172	0.9	2.3	0	13
number_lipids	172	0.3	1.1	0	7
number_fill	172	0.7	2.0	0	12
number_top_1000	172	10.8	25.9	0	163
rug_ver_moderna	172	0.05	0.2	0	1

All the commands and algorithms are coded in R 4.2.2




Heterogeneity across years – Model 2
Number of doses administrated
0
1000
3000
5000
2020
2021
2022
2023
Year

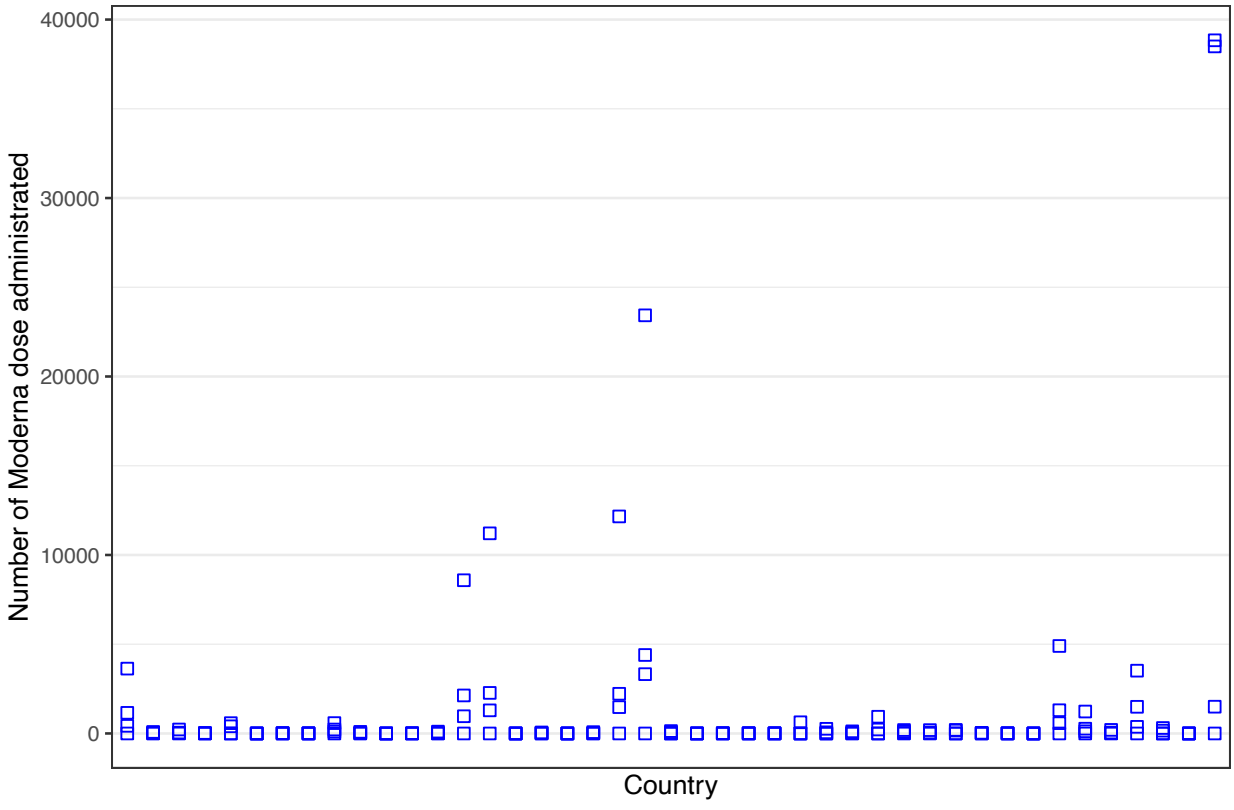

Number of Moderna dose administrated
40000
30000
20000
10000
0
Country

Table 2-7 shows the linear regression results of Model 2. We can see that 7 variables are significant at different levels. The id_21 is the only variable that is significant at the 1 percent level and has a positive impact on the dependent variable. The number_lipids, the id_1 and the constant variable are all significant at the 10 percent level. Both the number_lipids and the id_1 have a positive impact on the number of doses administrated and the constant variable has a negative impact on the dependent variable. The number_cov_chain_moderna, the number_fill and the id_14 are significant at the 5 percent level and the first two have a positive impact on the Covid-19 vaccine administrated while the id_14 variable has a negative impact. *Table 2-8* shows the regression results of the test of robustness for Model 2. We can that in Beck & Katz model found that none of our 6 variables was significant. However, all the variable has a positive influence on the number of Moderna vaccine administrated in a country except the variable about the number of the top 1000 pharmaceutical company in the country which has a negative impact. On the other hand, Arellano & Bond's model found 4 significant variables at different levels. The lag(number_lipids) and the lag(number_fill) are both significant at the 10 percent level, while the first has a negative impact on the Covid-19 vaccine administrated and the second has a positive. Additionally, the lag(reg_moderna,1) is significant at the 5 percent level with a positive influence and the lag(number_cov_chain_moderna,1) is significant at the 1 percent level with a negative impact on the dose administrated.

Table 2-7: Linear Regression Results – Model 2

	Number of Vaccine dose administrated
reg_moderna	46,496,773.00 (−1,313,453,305.00, 1,406,446,851.00)
number_cov_chain_moderna	1,556,998,936.00** (58,801,213.00, 3,055,196,660.00)
number_formulation	−1,481,404,003.00 (−3,356,804,584.00, 393,996,578.00)
number_lipids	2,050,589,141.00* (−90,016,405.00, 4,191,194,687.00)
number_fill	1,838,251,066.00** (403,672,227.00, 3,272,829,905.00)
number_top_1000	−21,868,650.00 (−99,002,515.00, 55,265,214.00)
id_1	4,473,805,338.00* (−106,782,018.00, 9,054,392,693.00)
id_2	1,369,452,874.00 (−2,631,391,445.00, 5,370,297,194.00)
id_3	4,269,860,922.00 (−1,372,377,918.00, 9,912,099,762.00)
id_4	1,663,787,315.00 (−2,088,652,877.00, 5,416,227,506.00)
id_5	1,165,911,434.00 (−2,928,270,876.00, 5,260,093,744.00)
id_6	1,643,456,217.00 (−2,131,461,258.00, 5,418,373,692.00)
id_7	1,683,885,165.00 (−2,063,009,318.00, 5,430,779,648.00)
id_8	1,635,404,579.00 (−2,124,156,004.00, 5,394,965,162.00)
id_9	3,362,104,975.00 (−1,017,687,014.00, 7,741,896,965.00)
id_10	1,796,429,230.00 (−1,944,139,501.00, 5,536,997,961.00)
id_11	1,688,573,253.00 (−2,195,638,901.00, 5,572,785,407.00)
id_12	1,636,503,291.00 (−2,123,057,293.00, 5,396,063,874.00)
id_13	1,754,031,894.00 (−1,986,521,249.00, 5,494,585,037.00)
id_14	−8,102,823,985.00** (−15,980,026,108.00, −225,621,862.00)
id_15	5,077,502,265.00 (−3,838,240,616.00, 13,993,245,145.00)
id_16	1,754,179,204.00 (−2,105,214,761.00, 5,613,573,168.00)
id_17	1,719,305,101.00 (−2,023,625,356.00, 5,462,235,558.00)
id_18	1,620,296,540.00 (−2,154,590,563.00, 5,395,183,643.00)
id_19	3,140,362,690.00 (−889,470,264.00, 7,170,195,645.00)
id_20	1,131,665,406.00 (−4,571,385,347.00, 6,834,716,158.00)
id_21	9,699,518,901.00*** (5,161,212,174.00, 14,237,825,628.00)
id_22	1,681,528,809.00 (−2,078,031,774.00, 5,441,089,393.00)
id_23	1,632,565,247.00 (−2,126,995,337.00, 5,392,125,830.00)
id_24	1,638,677,265.00 (−2,120,883,318.00, 5,398,237,849.00)
id_25	1,637,099,567.00 (−2,122,461,016.00, 5,396,660,150.00)
id_26	1,636,057,715.00 (−2,123,502,869.00, 5,395,618,298.00)
id_27	1,788,538,088.00 (−1,971,022,495.00, 5,548,098,672.00)
id_28	6,450,078,299.00 (−1,574,247,507.00, 14,474,404,104.00)
id_29	1,696,085,192.00 (−2,056,355,000.00, 5,448,525,383.00)
id_30	1,921,989,821.00 (−1,837,570,762.00, 5,681,550,405.00)
id_31	1,843,064,069.00 (−1,897,504,663.00, 5,583,632,800.00)
id_32	1,718,581,718.00 (−2,033,858,473.00, 5,471,021,910.00)
id_33	1,721,998,130.00 (−2,037,562,454.00, 5,481,558,713.00)
id_34	1,645,687,976.00 (−2,113,872,607.00, 5,405,248,560.00)
id_35	1,679,817,490.00 (−2,067,076,993.00, 5,426,711,973.00)
Constant	−1,666,704,603.00* (−3,501,271,029.00, 167,861,823.00)
Time Fixed Effects	No

Notes:

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.
Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

Table 2-8: Robustness Regression Results – Model 2

	Number of Vaccine dose administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_moderna	122.88 (−551.67, 797.43)	
number_cov_chain_moderna	1,197.46 (−1,174.70, 3,569.62)	
number_formulation	99.99 (−250.09, 450.08)	
number_lipids	231.97 (−860.79, 1,324.72)	
number_fill	940.33 (−496.40, 2,377.07)	
number_top_1000	−10.08 (−221.41, 201.25)	
lag(reg_moderna, 1)		562.60** (60.18, 1,065.03)
lag(number_cov_chain_moderna, 1)		−4,664.65*** (−6,653.87, −2,675.42)
lag(number_formulation, 1)		879.00 (−878.42, 2,636.42)
lag(number_lipids, 1)		−5,019.53* (−10,359.20, 320.14)
lag(number_fill, 1)		2,382.85* (−359.91, 5,125.62)
Constant	−170.91 (−766.68, 424.85)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

When looking at the regional distribution of Moderna's vaccine, we can see that the result differs from model 2. The region results of the modelling Moderna vaccine distribution in 31 European countries. The result of the regression shows that when using Beck & Katz model the number of lipids production facilities in the country is significant at the 10 percent level with a positive impact on the number of Moderna vaccine administrated. Arellano & Bond's result state that 2 variables are significant. The lag number of Covid-19 vaccine formulation facilities in the country has a negative impact on the number of Moderna doses administrated in the country with a significant level of 10 percent. The second significant variable is the lag number of fill & finish facilities in the country with a level of 5 percent. The region results of the modelling Moderna vaccine distribution in Canada and the United States show that Beck & Katz model found no significant variable. Arellano & Bond's results state that 2 variables are significant. The lag number of Covid-19 vaccine formulation facilities in the country and the lag number of Moderna supply chain facilities in the country has a negative impact on the number of Moderna doses administrated in the country with a significant level of 1 percent. For South American countries, both Beck & Katz and Arellano & Bond models found that none of the variables were significant. The result for Moderna vaccine distribution in Asian countries shows that Beck & Katz model found no significant variable. Arellano & Bond's results state that 3 variables are significant. The lag number of fill & finish facilities, the lag of Moderna vaccine approval and the lag of the number of Moderna supply chain facilities in the country is significant at the 1 percent level. Both the lag of Moderna vaccine approval and the lag number of fill & finish facility has a positive impact on the number of Moderna dose administrated while the lag of the number of Moderna supply chain facilities in the country has a negative impact. Lastly, the result shows that for both Beck & Katz and Arellano & Bond models, none of the variables were significant when modelling for Moderna vaccine distribution in African countries *(Appendix F to J)*.

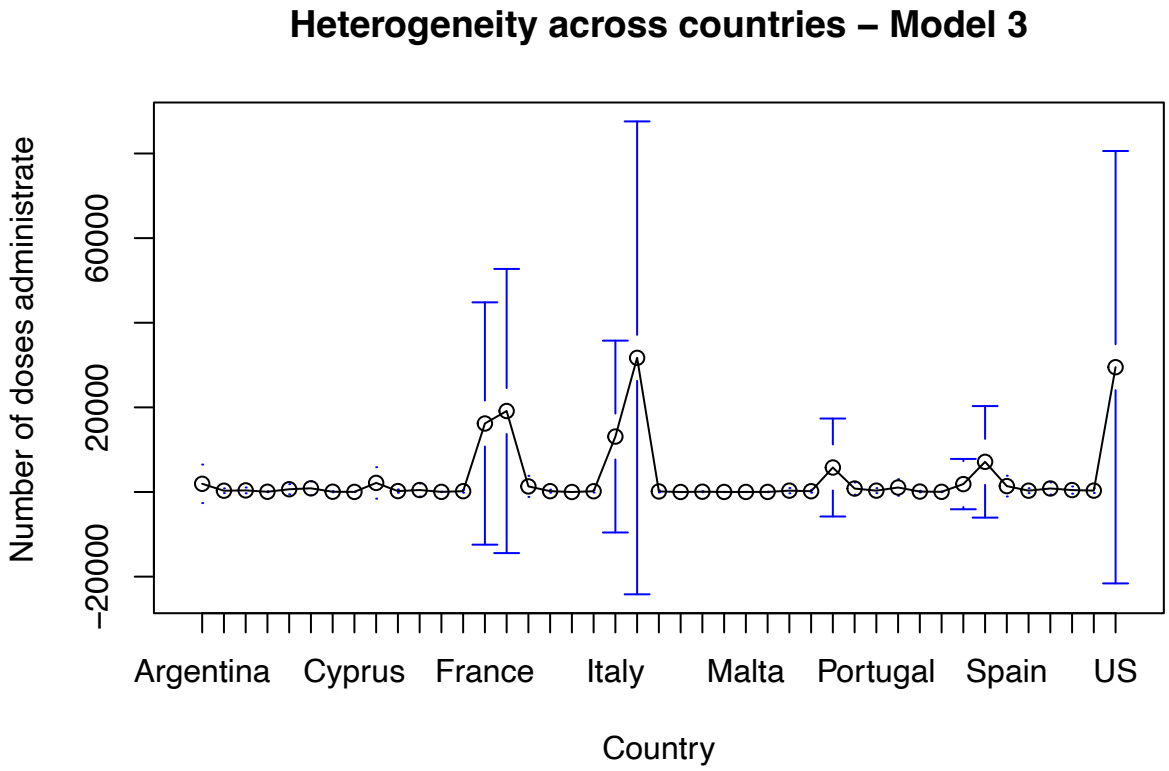
Model 3 - Pfizer

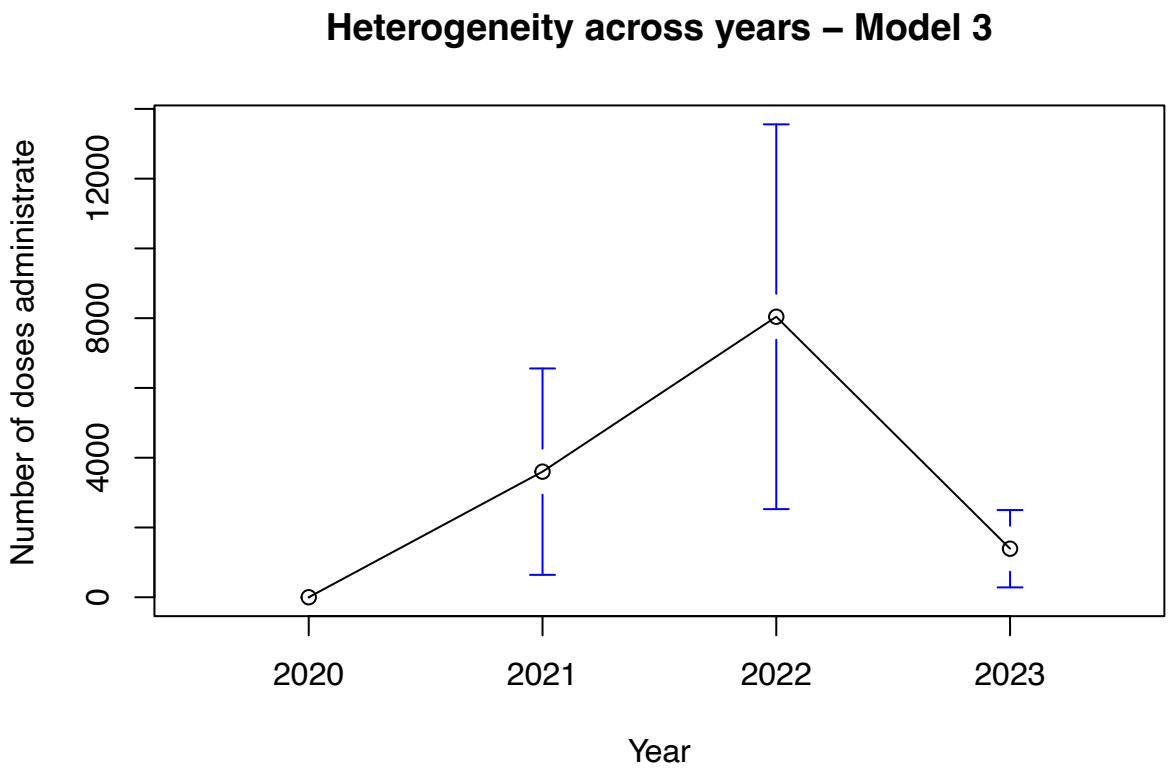
Table 2-9 is the statistical summary of our last dataset, which is for Pfizer's supply chain and the number of Pfizer vaccine doses administrated from December 2020 to March 2023, representing 172 observations or 34 countries. The average total number of doses administrated (dose_moderna) is 3,269.3 with a standard deviation of 10,705.2. The average number of Covid-19 vaccine formulation facilities per country is 0.9, the average number of Covid-19 vaccine lipids production is 0.3 and the average number of Covid-19 fill & finish facilities is 0.7 with a standard deviation of 2.3, 1.1 and 2 respectively. Furthermore, the average number of Moderna Covid-19 SC facilities per country is 0.5 with a standard deviation of 1.8.

Table 2-9: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
id	172	22.0	12.4	1	43
year	172	2,021.5	1.1	2,020	2,023
reg_pfizer	172	0.9	0.2	0	1
dose_pfizer	171	3,269.3	10,705.2	0.0	81,780.3
number_cov_chain_pfizer	172	0.5	1.8	0	10
number_formulation	172	0.9	2.3	0	13
number_lipids	172	0.3	1.1	0	7
number_fill	172	0.7	2.0	0	12
number_top_1000	172	10.8	25.9	0	163
rug_ver_pfizer	172	0.8	0.4	0	1

All the commands and algorithms are coded in R 4.2.2




Heterogeneity across years – Model 3
Number of doses administrate
0
4000
8000
12000
2020
2021
2022
2023
Year

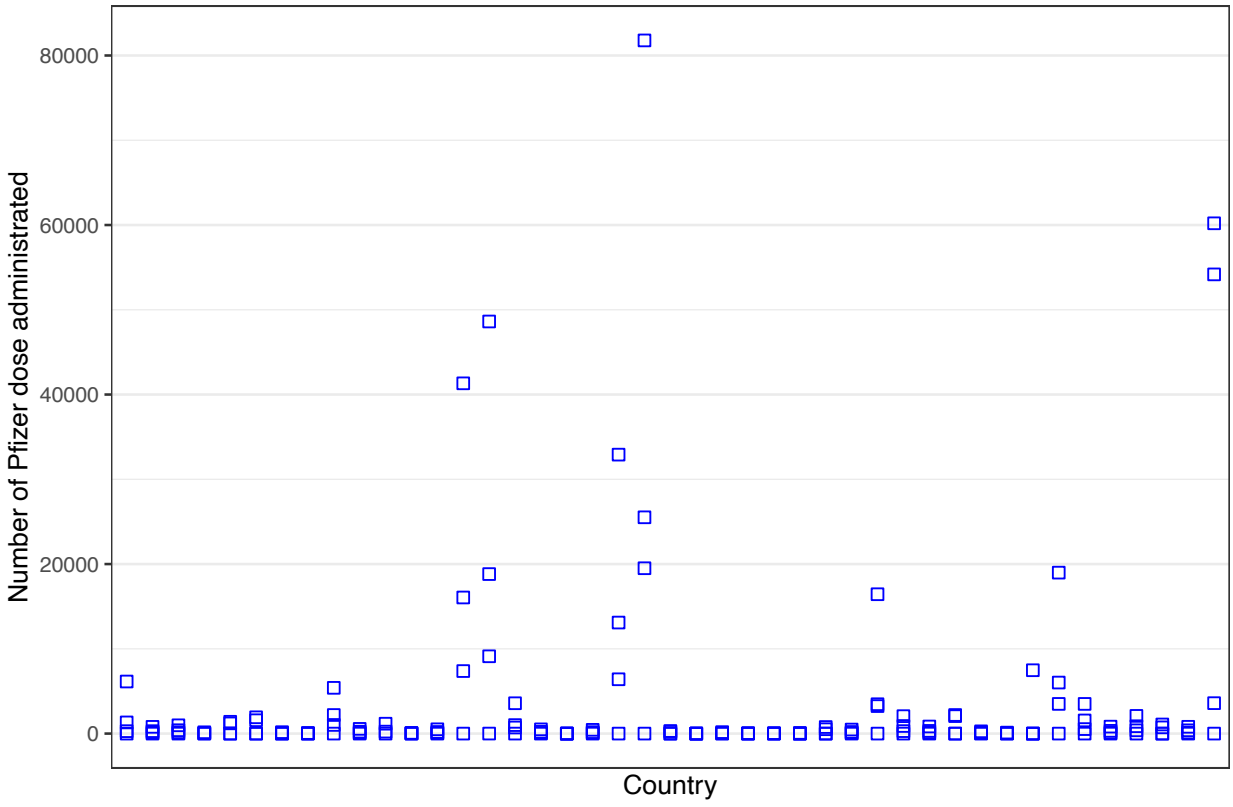

Number of Pfizer dose administrated
80000
60000
40000
20000
0
Country

Table 2-10 shows the linear regression results of Model 3. We can see that 4 variables are significant at different levels. The id_21 is significant at the 1 percent level and has a positive impact on the dependent variable. As the 3 other significant variables, they are significant at the 5 percent level. Those variables are reg_pfizer, number_cov_chain_pfizer and the constant variable. Additionally, both the reg_pfizer and number_cov_chain_pfizer have a positive impact on the dependent variable while the constant variable has a negative impact on it. *Table 2-11* shows the results of the test of robustness for Model 3. We can see that in Beck & Katz model only two of our six variable was significant. The number of Covid-19 vaccine formulation facilities in the country is significant at the 5 percent level and the number of fill & finish facilities in the country is also significant at the 5 percent level, both variables have a positive influence on the number of Pfizer vaccine administrated per country. Arellano & Bond's model found two significant variables. According to *Table 2-11*, the lag of the number of Pfizer supply chain facilities present in the country is significant at the 10 percent level and has a positive impact on the number of doses administrated. The lag number of Covid-19 vaccine lipids production facilities in the country is significant at the 5 percent level and has a negative impact.

Table 2-10: Linear Regression Results – Model 3

	Number of Vaccine dose administrated
reg_pfizer	6,819,761,954.00** (764,763,463.00, 12,874,760,444.00)
number_cov_chain_pfizer	5,289,817,471.00** (372,212,977.00, 10,207,421,965.00)
number_formulation	−3,380,071,451.00 (−8,394,068,531.00, 1,633,925,628.00)
number_lipids	−2,429,347,169.00 (−8,348,145,653.00, 3,489,451,315.00)
number_fill	1,958,887,697.00 (−1,520,631,893.00, 5,438,407,287.00)
number_top_1000	164,008,550.00 (−110,191,693.00, 438,208,793.00)
id_1	4,533,288,868.00 (−5,744,554,785.00, 14,811,132,521.00)
id_2	−909,372,631.00 (−10,145,416,863.00, 8,326,671,602.00)
id_3	−5,940,726,410.00 (−19,791,414,093.00, 7,909,961,272.00)
id_4	−725,009,499.00 (−9,383,203,270.00, 7,933,184,272.00)
id_5	3,191,650,413.00 (−6,365,197,182.00, 12,748,498,007.00)
id_6	254,896,519.00 (−8,395,935,925.00, 8,905,728,963.00)
id_7	−884,006,018.00 (−9,558,226,924.00, 7,790,214,889.00)
id_8	−594,095,733.00 (−9,244,928,177.00, 8,056,736,712.00)
id_9	4,748,813,151.00 (−5,529,030,502.00, 15,026,656,804.00)
id_10	−1,386,667,905.00 (−10,210,328,120.00, 7,436,992,310.00)
id_11	1,956,941,122.00 (−7,731,434,147.00, 11,645,316,391.00)
id_12	−587,221,108.00 (−9,238,053,552.00, 8,063,611,337.00)
id_13	−1,082,304,321.00 (−9,814,360,257.00, 7,649,751,616.00)
id_14	−1,667,792,089.00 (−19,030,030,563.00, 15,694,446,386.00)
id_15	−4,988,103,755.00 (−26,066,688,367.00, 16,090,480,858.00)
id_16	1,756,419,706.00 (−6,838,802,686.00, 10,351,642,099.00)
id_17	−915,561,736.00 (−9,614,427,686.00, 7,783,304,215.00)
id_18	−617,449,065.00 (−9,268,281,510.00, 8,033,383,379.00)
id_19	−4,664,752,522.00 (−15,889,896,235.00, 6,560,391,191.00)
id_20	−1,296,812,994.00 (−13,901,329,441.00, 11,307,703,453.00)
id_21	32,019,606,614.00*** (21,866,324,086.00, 42,172,889,142.00)
id_22	−488,223,984.00 (−9,139,056,428.00, 8,162,608,461.00)
id_23	−618,155,195.00 (−9,268,987,639.00, 8,032,677,249.00)
id_24	−550,261,321.00 (−9,201,093,765.00, 8,100,571,123.00)
id_25	−603,359,079.00 (−9,254,191,523.00, 8,047,473,366.00)
id_26	−603,150,843.00 (−9,253,983,287.00, 8,047,681,602.00)
id_27	1,105,408,479.00 (−7,449,660,581.00, 9,660,477,539.00)
id_28	9,754,124,401.00 (−9,460,936,238.00, 28,969,185,040.00)
id_29	−599,215,158.00 (−9,257,408,929.00, 8,058,978,613.00)
id_30	6,873,586,114.00 (−1,681,482,946.00, 15,428,655,174.00)
id_31	−791,957,263.00 (−9,615,617,478.00, 8,031,702,952.00)
id_32	−457,450,372.00 (−9,115,644,143.00, 8,200,743,400.00)
id_33	445,461,131.00 (−8,205,371,313.00, 9,096,293,576.00)
id_34	−513,826,773.00 (−9,164,659,217.00, 8,137,005,672.00)
id_35	−901,469,677.00 (−9,575,690,583.00, 7,772,751,230.00)
Constant	−6,201,262,679.00** (−12,099,033,382.00, −303,491,976.00)
Time Fixed Effects	No

Notes:

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.
Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

Table 2-11: Robustness Regression Results – Model 3

	Number of Vaccine dose administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_pfizer	3,016.35 (−7,715.41, 13,748.12)	
number_cov_chain_pfizer	−499.63 (−1,206.07, 206.82)	
number_formulation	882.97** (201.60, 1,564.33)	
number_lipids	−2,335.54 (−6,791.21, 2,120.13)	
number_fill	4,708.00** (880.85, 8,535.14)	
number_top_1000	−61.42 (−348.63, 225.78)	
lag(reg_pfizer, 1)		3,989.74 (−914.98, 8,894.47)
lag(number_cov_chain_pfizer, 1)		5,429.62* (−316.79, 11,176.02)
lag(number_formulation, 1)		1,814.97 (−1,599.70, 5,229.63)
lag(number_lipids, 1)		−12,962.11** (−25,306.35, −617.87)
lag(number_fill, 1)		−3,625.35 (−9,335.14, 2,084.44)
Constant	−2,140.00 (−12,469.55, 8,189.55)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

Similar to model 1 and model 2, we model Pfizer vaccine distribution for different regions. Starting with the result of the European countries. Beck & Katz's model found three significant variables. The number of Pfizer supply chain facilities in the country is significant at the 1 percent level and has a positive impact on the number of Pfizer vaccines administrated. The number of formulation and fill & finish facilities in the country is significant at the 5 percent level. Arellano & Bond's results state that 3 variables are significant. The lag number of Covid-19 vaccine fill & finish facilities in the country and the lag of the number of Pfizer supply chain facilities in the country are significant at the 5 percent level. The lag of Pfizer vaccine approval in the country is significant at the 1 percent level. After modelling vaccine distribution in North American countries, Canada and the United States, the result shows that Beck & Katz model found no significant variable. On the other hand, Arellano & Bond's model found that the lag of Pfizer vaccine approval in the country and the lag of having formulation facilities in the country are both significant at the 1 percent level and have a negative impact on the number of Pfizer doses administrated. The result shows that for both Beck & Katz and Arellano & Bond models, none of the variables were significant concerning the distribution of Pfizer's vaccine in South America. The results of the modelling Pfizer vaccine distribution in 4 Asian countries show that for both Beck & Katz and Arellano & Bond models, none of the variables were significant. In terms of distribution in African countries, the result shows that Beck & Katz model found no significant variable. On the other hand, Arellano & Bond's model found that the lag of Pfizer vaccine approval is significant at the 1 percent level and has a positive impact on the number of Pfizer vaccines administrated in the country *(Appendix K to O)*.

2.5 Discussion

The 2020 Covid-19 pandemic triggered a global health crisis that required the rapid development, production, and distribution of vaccines on a large scale. Efficient and equitable allocation of vaccines is crucial to counter the spread of the virus and mitigating its impact on public health and economies. As countries strive to protect their populations, ensuring equitable access to the Covid-19 vaccine supply chain has become a critical concern. Achieving an equitable distribution system is vital to address both the immediate health needs of individuals and the broader goal of containing the pandemic on a global scale.

The project aims to examine the elements related to VSC that affect the distribution of Covid-19 vaccines produced by leading manufacturers such as Pfizer and Moderna. We constructed econometric models, using the panel data analysis, that depict the connections between crucial variables in the vaccine supply chain. Furthermore, we will measure the influence of different factors on the availability and distribution of vaccines within a country, including current vaccine production facilities, vaccine approval policies, and more. With the literature, we formulated two hypotheses; H1) The presence of the vaccine producers' supply chain in a country affects the number of vaccine doses administrated in that country and H2) the high number of pharmaceutical companies present in a country has a positive impact on the number of doses administrated.

This project aims to identify whether supply chain factors affect the distribution of vaccines per country and what are the relevant factors. Thus, providing empirical insights on equitable vaccine distribution in the context of a global pandemic from a supply chain perceptive. The data collected provided information on the number of doses administrated over 3 years for 43 countries for Pfizer and Moderna. With dose data, we created three panel dataframes, one for each producer and one for both combined. Having panel data, we use the panel data analysis approach. First, we had to perform

a validity test on our data. We examined heterogeneity in our data set using the number of doses administered as the dependent variable and the country as the independent variable. After validating the data sets, we tested our data set for fixed and random effects using the Hausman test, resulting in the need for a fixed-effects model. With the result of the Hausman test, we decided to use both Beck and Katz model and Arellano model to identify the key variable influencing the dose of the Covid-19 vaccine administrated in a country creating 3 initial models. We also model the dose of the Covid-19 vaccine administrated in a country based on their geographical region, thus creating 15 models in total.

When looking at the results, we found that different variables were significant at different levels for our three models. For Model 1, the regression that combines data of Pfizer & Moderna, the number of fill & finish facilities in the country is significant at the 10 percent level and the variable of Pfizer Rugman's classification (rug_ver_pfizer) is significant at the 5 percent level when using Beck & Katz model. According to Arellano & Bond's model, it is the lag of the number of Moderna SC facilities present in the country is significant at the 1 percent level. For Model 2, Beck & Katz model found that none of our 6 variables was significant when determining which factors influence the number of Moderna's vaccines administrated in a country. However, Arellano & Bond's model found 4 significant variables at different levels, The lag(number_lipids) and the lag(number_fill) are both significant at the 10 percent level, the lag(reg_moderna,1) is significant at the 5 percent level and the lag(number_cov_chain_moderna,1) is significant at the 1 percent level. Lastly for Model 3, Beck & Katz model found that both the number of Covid-19 vaccine formulation facilities and the number of fill & finish facilities in the country are significant at the 5 percent level.

Our results contradict or confirm our first hypothesis (H1) depending on the model used. In the case of Model 1 the presence of both Pfizer and Moderna's supply chain in the country is not significant when using Beck & Katz model and Arellano & Bond model. However, for Model 2 and Model 3, the presence of a vaccine supply chain is significant when using Arellano & Bond model. This shows that the presence of VSC cannot be used as an indicator of vaccine distribution. For the second hypothesis (H2), our result contradicts it. For all the models, the number of top 1000 pharmaceutical companies in the country is not significant and it has a negative impact on the number of doses administrated.

The use of panel econometric models in analyzing the distribution of the Covid-19 vaccine has proven to be a valuable tool. Furthermore, panel econometric models enable researchers to analyze the impact of various determinants on vaccine distribution, such as population characteristics, healthcare infrastructure, government policies, and socio-economic factors. By incorporating multiple dimensions, these models offer a more comprehensive and nuanced analysis of the complex dynamics at play. Panel econometric models have demonstrated their utility in providing valuable insights into the distribution of the Covid-19 vaccine. Their ability to capture heterogeneity, control for unobserved factors, and estimate causal effects have enhanced our understanding of the factors influencing vaccine distribution and have supported evidence-based policymaking.

The panel econometric model can be a useful tool for analyzing vaccine supply chain dynamics, but it also has certain limitations, primarily the availability of data and the variables chosen. First, the panel econometric model requires a relatively large sample size to produce reliable estimates, which may be challenging for specific regions or countries with limited data availability. The data we collected was for 43 countries, mostly European countries, which skewed the accuracy of our model. Additionally, the variable choice is important as panel econometric models can provide insights into the associations between variables, but establishing causality can be challenging. Vaccine supply chains and vaccine distribution are influenced by various factors, some of which may not be observable or measurable.

Omitted variable bias can arise when relevant variables are excluded from the model, leading to biased parameter estimates and potentially incorrect inferences. Moreover, panel models assume that the relationships among variables are constant over time and across entities, which may not always hold in the context of vaccine distribution, where factors such as vaccine availability and public sentiment can vary significantly. Additionally, the interpretation of panel econometric results requires caution, as they are susceptible to omitted variable bias and endogeneity.

Moving forward, it is crucial to continue refining and advancing panel econometric modelling techniques to address the specific challenges posed by vaccine distribution. This includes incorporating more granular data, exploring dynamic relationships, and considering additional factors such as vaccine hesitancy and distribution logistics. By leveraging econometric models and addressing their limitations, policymakers and stakeholders can make informed decisions to ensure equitable and efficient distribution of vaccines, ultimately contributing to global efforts in controlling the pandemic.

References

Ahmed, F., Ahmed, N., Pissarides, C., & Stiglitz, J. (2022). Why Inequality Could Spread Covid-19. The Lancet, 395(10230), 809-810.

Asch, D. A., et al. (2021). Mobile Outreach for Covid-19 Vaccination. Journal of the American Medical Association, 325(13), 1251-1252.

Bollyky, T. J., et al. (2021). The Equitable Distribution of Covid-19 Vaccines. Journal of the American Medical Association, 325(10), 929-930.

Bollyky, T. J., Gostin, L. O., Hamburg, M. A., Piot, P., & Redlener, I. E. (2021). National Security and Global Health: Safeguarding Global Health in the 21st Century. Columbia University Press.

Brewer, N. T., Chapman, G. B., Rothman, A. J., Leask, J., & Kempe, A. (2017). Increasing Vaccination: Putting Psychological Science Into Action. Psychological Science in the Public Interest, 18(3), 149-207.

Delamater, P. L., et al. (2021). Vaccine Equity and Access: A Review of Interventions for Underserved Populations. Journal of Community Health, 46(4), 726-740.

Forrester, J. W. (1961). Industrial Dynamics. MIT Press.

Ghebreyesus, T. A., Moeti, M., Vaz, R. G., Bhatnagar, N., & Lal, S. (2021). Covid-19 Vaccination: Pathway to Equitable Global Access. World Health Organization.

Hotez, P. J., et al. (2021). Equity Must Accompany Covid-19 Vaccines in Global Health. PLoS Neglected Tropical Diseases, 15(1), e0008962.

Karafillakis, E., et al. (2019). Vaccine Hesitancy among Healthcare Workers in Europe: A Systematic Review. Vaccine, 37(36), 4761-4768.

Karakaya, E., Aksu, S., & Ustundag, A. (2012). Supply Chain Risk Assessment: A Case Study in the Pharmaceutical Industry. International Journal of Production Research, 50(21), 6162-6172.

Leask, J., et al. (2018). Target the Fence-Sitters. Nature, 555(7696), 446-447.

MacDonald, N. E., SAGE Working Group on Vaccine Hesitancy. (2015). Vaccine Hesitancy: Definition, Scope and Determinants. Vaccine, 33(34), 4161-4164.

Mclay, L. A., & Santos, M. (2011). Quantitative Models for Managing Supply Chain Risks: A Review. European Journal of Operational Research, 214(2), 283-293.

Mehrotra, A., et al. (2021). Addressing the Equity Barrier in Covid-19 Vaccination. Journal of General Internal Medicine, 36(4), 1036-1039.

Montgomery, D. C., & Runger, G. C. (2014). Applied Statistics and Probability for Engineers. John Wiley & Sons.

Moon, S., et al. (2020). Will Ebola Change the Game? Ten Essential Reforms before the Next Pandemic. The Report of the Harvard-LSHTM Independent Panel on the Global Response to Ebola. The Lancet, 386(10009), 2204-2221.

O'Leary, S. T., et al. (2021). Disparities in Vaccine Hesitancy: A Review of Multiple Measures, Causes, and Possible Interventions. Human Vaccines & Immunotherapeutics, 17(5), 1418-1428.

Oliver, S. E., Gargano, J. W., Marin, M., Wallace, M., Curran, K. G., Chamberland, M., … & Dooling, K. (2021). The Advisory Committee on Immunization Practices' Interim Recommendation for Use of Pfizer-BioNTech Covid-19 Vaccine in Adolescents Aged 12-15 Years—United States, May 2021. Morbidity and Mortality Weekly Report, 70(20), 749-752.

Phelan, A. L., & Oluyomi, A. (2021). Covid-19 Vaccine Equity in the United States: A Rapid Review. Journal of Racial and Ethnic Health Disparities, 1-10.

Phelan, A. L., et al. (2020). Vaccine Ethics and Equity: More Than Just Distribution. American Journal of Bioethics, 20(12), 7-9.

Vazquez-Prokopec, G. M., Mercado, M. I. G., & Dumonteil, E. (2021). The Challenge of Vector-Borne Diseases in Underserved Regions: A Scoping Review. Frontiers in Public Health, 9, 613540.

Wooldridge, J. M. (2015). Introductory Econometrics: A Modern Approach. Cengage Learning.

Chapter 3: Network Analysis Approach to the Covid-19 Vaccine Production Supply Chain - A Data Science Perspective

Koné Aïchata Souleymane, M.Sc. student, HEC
Montreal Supervision:
Bélanger Valérie, HEC
Montreal Warin Thierry, HEC
Montreal

HEC Montréal

Abstract

Network analysis has proven to be a powerful tool for optimizing the distribution and supply of the vaccine COVID -19. This study explores the application of network analysis techniques to the complex web of stakeholders involved in the vaccine supply chain. By mapping the connections and interactions among the actors involved in the chain, network analysis provides insights into the flow of vaccines and identifies potential bottlenecks. Visualizing the network helps to understand the dynamics of vaccine distribution and enables policy makers and health organizations to develop targeted strategies to ensure efficient distribution and equitable access. Network analysis is conducted using social network analysis (SNA) techniques. SNA will map the network structure of the vaccine supply chain, identify key nodes and links, and analyze the flow of vaccines through the network. The expected outcomes of this study are as follows. 1) A map of the network structure of the Covid-19 vaccine supply chain, including key players, nodes, and links. 2) Identification of bottlenecks and inefficiencies in the vaccine supply chain. 3) Suggest strategies to improve the efficiency and effectiveness of the vaccine supply chain.

3.1 Introduction

The Covid-19 pandemic has presented unprecedented challenges to global healthcare systems, economies, and societies (Wouters et al., 2021). The development, production, and distribution of Covid-19 vaccines have become critical endeavors in curbing the spread of the virus and mitigating its impact (Osterholm et al., 2021). The complexity of the vaccine supply chain, encompassing multiple stakeholders, intricate logistics, and global distribution networks, necessitates innovative approaches to optimize its efficiency, resilience, and equity.

Network analysis has emerged as a powerful methodology for understanding complex systems and has gained increasing attention in the context of the Covid-19 vaccine supply chain. By examining the interconnections and interactions among key actors and nodes, network analysis offers valuable insights into the structure, dynamics, and performance of the vaccine supply chain (Wang et al., 2021). It facilitates the identification of bottlenecks, vulnerabilities, and opportunities for improvement, thereby enabling policymakers and stakeholders to make informed decisions to enhance the effectiveness of vaccine distribution efforts. Furthermore, network analysis allows for the evaluation of the resilience of the vaccine supply chain. By simulating disruptions, identifying critical nodes or links, and assessing the cascading effects of failures, researchers can assess the system's ability to withstand shocks and propose strategies to enhance its resilience (Wu et al., 2021). Moreover, network

analysis can facilitate the assessment of the efficiency and equity of vaccine distribution, enabling the identification of areas where resources can be better allocated to ensure fair and timely access to vaccines for all populations.

The main objective of this study is to conduct a network analysis of the Covid-19 vaccine supply chain. The network analysis will be conducted using social network analysis (SNA) techniques. The SNA will involve mapping the network structure of both vaccine supply chains, identifying the key nodes and links, and analyzing the flow of vaccines through the network. The expected results of this study are as follows: 1) A map of the network structure of the Covid-19 vaccine supply chain, including the key stakeholders, nodes, and links. 2) Identification of bottlenecks and inefficiencies in the vaccine supply chain. 3) Proposals for strategies to improve the efficiency and effectiveness of the vaccine supply chain. The data will be collected from various sources, including official reports, news articles, and academic papers on the Pfizer and Moderna Covid-19 vaccine supply chain as of June 2021.

Understanding the network structure of the vaccine supply chain is crucial for identifying critical nodes and relationships that can significantly impact the production process. By mapping out the network, researchers can discern the key manufacturers involved, and analyze the flow of vaccines between them. Such analysis can reveal patterns of centralization or fragmentation in the network, shedding light on potential vulnerabilities or inefficiencies that need to be addressed. Thus, the results of this study can inform policymakers and stakeholders in the vaccine supply chain on strategies for improving the distribution and delivery of vaccines, ultimately contributing to the global effort to control the spread of Covid-19.

3.2 Literature Review

The efficient and reliable distribution of vaccines is critical to global public health. Vaccination programs have saved countless lives by preventing and mitigating the spread of infectious diseases. However, the successful implementation of vaccination campaigns depends on a well-functioning vaccine supply chain. Bottlenecks in supply chain management refer to specific points or stages within a supply chain where the flow of goods, information, or processes is hindered, or slowed down, leading to inefficiencies, delays, and potential disruptions. Identifying and addressing bottlenecks is crucial for optimizing supply chain performance and ensuring the smooth flow of products from suppliers to consumers. Previous papers provide an overview of the concept of bottlenecks in supply chain management and provide key research findings.

Bottlenecks in supply chain management can take various forms as previously mentioned. Goldratt and Cox (1984) introduced the Theory of Constraints (TOC), which emphasizes identifying and mitigating bottlenecks as a fundamental principle for improving supply chain efficiency. They argue that a chain is only as strong as its weakest link, making bottleneck detection and resolution essential. One of the most common types of bottlenecks in supply chains is capacity constraints. Research by Hopp and Spearman (2000) discusses the impact of limited production capacity on supply chain performance. They emphasize the need for effective capacity management and scheduling to alleviate bottlenecks and maintain a balanced flow of products. Furthermore, excessive inventory levels can also lead to bottlenecks due to storage constraints, increased carrying costs and information distortion. This distortion can lead to another concept in supply chain management, the “bullwhip effect” (Lee & Billington, 1992). Lastly, Transportation and logistics bottlenecks can occur due to delays in transit,

customs clearance, or insufficient transport capacity. Researchers, such as Notteboom and Rodrigue (2005), investigate port-related bottlenecks and suggest strategies for addressing them. They emphasized the importance of efficient transportation networks in global supply chains.

Detecting and addressing bottlenecks in the vaccine supply chain is essential to ensure timely and equitable access to vaccines. Thus, various techniques and methodologies employed for bottleneck detection in vaccine supply chains have been proposed by researchers. Among bottleneck detection techniques, Cold chain monitoring plays a pivotal role in vaccine distribution due to the sensitivity of vaccines to temperature variations. Kumari et al. (2016) highlight in their paper the importance of real-time temperature monitoring using sensors and data loggers to detect temperature excursions. Deviations from recommended temperature ranges can signal bottlenecks in cold chain infrastructure and transportation. Another proposed methodology is the use of Data analytics and predictive modelling. These methods have seen an increase in their use to detect bottlenecks in vaccine supply chains. Indeed, Takian et al. (2020) explore the application of machine learning algorithms to analyze historical vaccine distribution data and predict potential bottlenecks. Predictive models can help decision-makers proactively allocate resources and optimize supply chain operations. Lastly, Ribeiro et al. (2017) demonstrate how Geographic Information Systems (GIS) analyses can help pinpoint areas with limited access to vaccination facilities, enabling targeted interventions. With GIS, they were able to identify bottlenecks related to geographical constraints and accessibility.

Network analysis

Network analysis has become a fundamental tool for understanding the complex patterns of interconnected entities in a wide range of domains, including social sciences, biology, computer science, and engineering (Newman, 2010; Barabási, 2016). With the advent of large-scale data and the increasing availability of computational resources, network analysis has gained prominence as a powerful approach to uncovering hidden structures, relationships, and dynamics within complex systems. By applying graph theory and advanced analytical techniques, network analysis provides valuable insights into the structure, function, and behavior of networks.

The foundation of network analysis lies in graph theory, which provides a mathematical framework for modelling and analyzing interconnected structures (Diestel, 2017). Networks, represented as graphs consisting of nodes and edges, capture the relationships and interactions between entities. The nodes can represent individuals, organizations, genes, web pages, or any other discrete unit of interest, while the edges represent the connections, dependencies, or interactions between these entities. By studying the patterns and properties of nodes and edges, network analysis uncovers important structural features and dynamics that are often not evident through traditional analytical approaches.

Network analysis encompasses various fundamental concepts and measures. For instance, centrality measures, such as degree centrality, betweenness centrality, and eigenvector centrality, quantify the relative importance or influence of nodes within a network (Newman, 2010). These measures allow to identify key nodes that act as connectors or central hubs within a network. Another important concept is community detection, which aims to identify cohesive groups of nodes that exhibit strong internal connections and weaker connections with other groups (Fortunato, 2010). Community detection algorithms enable the identification of functional substructures within complex networks, providing insights into modular organization and information flow.

The methods and techniques used in network analysis have evolved rapidly in recent years. Network visualization plays a crucial role in understanding network structures and patterns (Borgatti, Everett, & Johnson, 2018). Visualization techniques allow to explore and interpret complex networks, facilitating the identification of clusters, subgroups, and relationships between nodes. Network modelling involves constructing mathematical models that capture the characteristics and properties of real-world networks (Barabási & Albert, 1999; Watts & Strogatz, 1998). These models help understand and explain the emergence of specific network properties, such as the small-world phenomenon or scale-free degree distributions.

The applications of network analysis are diverse and span multiple disciplines. In social sciences, network analysis has been applied to study social relationships, influence dynamics, information diffusion, and the spread of opinions and behaviors within social networks (Wasserman & Faust, 1994; Borgatti, Mehra, Brass, & Labianca, 2009). In biology, network analysis has been instrumental in uncovering molecular interactions, gene regulatory networks, protein-protein interactions, and ecological networks (Barabási, 2016; Albert, Jeong, & Barabási, 2000). Network analysis also finds applications in transportation systems, urban planning, technological networks, and communication systems, aiding in optimizing network efficiency, identifying critical nodes, and improving infrastructure design (Ortúzar & Willumsen, 2011; Cohen & Havlin, 2010).

Network structures or topologies play a crucial role in the design and analysis of networks. Several studies have examined the strengths and weaknesses of different network structures to understand their impact on network performance and reliability.

One commonly studied network structure is the bus topology. According to Tanenbaum (2011), the bus topology is simple to implement and cost-effective. However, a significant drawback of the bus topology is that the failure of the main link can bring down the entire network, making fault identification challenging. In contrast, the star topology has gained popularity due to its centralized control and easy management (Forouzan, 2013). In this structure, a central node connects all other nodes in the network, providing individual fault isolation. However, the central node becomes a single point of failure, potentially affecting the entire network if disrupted. Another well-known network structure is the ring topology, where nodes are connected in a closed loop (Kurose & Ross, 2020). This topology offers a specific sequence for pathing and reduces redundancy. However, a single node or link failure can disrupt the entire network, and troubleshooting in a closed loop is challenging. Mesh topology provides the highest level of redundancy as each node is connected to every other node (Stallings, 2013). This structure ensures multiple paths for data transmission, allowing for seamless communication even if one or more connections fail. However, implementing a mesh network can be costly due to the number of linkages. Hybrid topologies have gained attention due to their flexibility and scalability (Comer, 2019). By combining different network structures, hybrid topologies can be tailored to meet specific needs. However, the complexity of managing and maintaining multiple topologies within a hybrid structure should be considered. Lastly, the tree topology combines the features of the bus and star topologies. It allows for the extension of networks by connecting multiple star topologies. It provides a hierarchical structure that can be easily managed (Tanenbaum, 2011). However, failure of the central hub or main backbone can disrupt the entire network. The scalability of the tree topology is limited by the number of available ports on the central hub (Tanenbaum, 2011).

Centrality measure

In network analysis, centrality measures are used to identify the most important nodes in a network. Centrality measures capture different aspects of node importance, such as the number of connections, the degree of influence, and the ability to bridge different parts of the network.

Degree Centrality: Degree centrality is the most basic centrality measure, and it is based on the number of connections a node has in a network. The more connections a node has, the higher its degree of centrality. Degree centrality has been used in various domains, such as social network analysis (Scott, 2000) and protein-protein interaction networks (Jeong et al., 2001).

Degree centrality possesses several strengths that contribute to its wide usage. Firstly, it is simple and interpretable, allowing researchers to intuitively understand the importance of nodes based on their degree values (Borgatti, Everett, & Freeman, 2002). Secondly, it is computationally efficient, making it applicable to large-scale network analysis (Brandes, 2001). Thirdly, degree centrality is robust to random failures and noise, as nodes with high degrees tend to have redundancy in their connections (Cohen, Havlin, & Ben-Avraham, 2003). Lastly, degree centrality captures both local and global node importance, providing insights into both individual node influence and the overall network structure (Estrada & Rodríguez-Velázquez, 2005).

Despite its strengths, degree centrality has some limitations. Firstly, it is insensitive to the position of the node within the network, ignoring the role of neighbouring nodes (Barrat et al., 2004). Secondly, degree centrality fails to capture indirect or long-range effects in the network, which are important in certain contexts (Opsahl et al., 2008). Thirdly, degree centrality is vulnerable to intentional attacks and targeted removals, as important nodes can be easily identified and compromised (Albert et al., 2000). Lastly, degree centrality overlooks the importance of node attributes or contextual factors, which may significantly influence node importance in certain applications (Borgatti, 2006).

Betweenness Centrality: Betweenness centrality measures the degree to which a node acts as a bridge between different parts of a network. Nodes with high betweenness centrality are important for maintaining the connectivity of the network. Betweenness centrality has been used in various domains, such as transportation networks (Brandes et al., 2008) and social network analysis (Freeman, 1979).

Betweenness centrality possesses several strengths. Firstly, it identifies nodes that play a crucial role in mediating interactions and information flow between other nodes, making it particularly useful in understanding the network's communication efficiency (Freeman, 1977). Secondly, it captures the strategic position of nodes in controlling the flow of information, making it valuable for identifying potential bottlenecks or vulnerable points in the network (Brandes, 2001). Thirdly, betweenness centrality can be applied to both unweighted and weighted networks, allowing for a flexible analysis of different types of networks (Borgatti, 2005).

Betweenness centrality has some limitations. Firstly, it tends to favour nodes in networks with a higher number of shortest paths, potentially overlooking nodes with alternative pathways or longer-range effects (Newman, 2018). Secondly, betweenness centrality is computationally expensive to calculate, particularly in large networks, making it challenging to apply to massive-scale datasets (Brandes, 2001). Thirdly, betweenness centrality does not consider the importance of node attributes or context, which can be relevant in certain applications (Borgatti, 2006).

Closeness Centrality: Closeness centrality measures the degree to which a node is close to all other nodes in a network. Nodes with high closeness centrality can spread information quickly and efficiently throughout the network. Closeness centrality has been used in various domains, such as social network analysis (Sabidussi, 1966) and ecological networks (Bodin et al., 2006).

This centrality measure provides insights into the accessibility and efficiency of information or influence flow within a network, allowing the identification of nodes that have faster access to the network's resources (Bavelas, 1950). Secondly, closeness centrality can be calculated for both weighted and unweighted networks, providing flexibility in analyzing different types of networks (Opsahl, 2013). Thirdly, closeness centrality captures the local perspective by focusing on the node's immediate neighbours, making it useful for understanding local information diffusion (Opsahl, 2013).

However, closeness centrality assumes that the shortest path is the only pathway for information or influence flow, which may not always reflect the complexity of real-world networks (Opsahl, 2013). Secondly, closeness centrality is sensitive to disconnected nodes or components in a network, potentially distorting centrality rankings (Sabidussi, 1966). Thirdly, closeness centrality may not adequately capture the influence of longer-range connections or indirect pathways in certain network contexts (Borgatti, 2005).

Eigenvector Centrality: Eigenvector centrality measures the degree to which a node is connected to other high-degree nodes in a network. Nodes with high eigenvector centrality are important because they are connected to other important nodes in the network. Eigenvector centrality has been used in various domains, such as social network analysis (Bonacich, 1972) and metabolic networks (Ma et al., 2003).

Eigenvector centrality considers not only the number of connections a node has but also the centrality of its neighbouring nodes. It assigns higher centrality to nodes that are connected to other important nodes in the network (Bonacich, 1972). This captures the idea that being connected to influential nodes increases the importance of a node. It also considers the entire network structure when calculating centrality. It considers the connections and importance of all nodes, allowing for a holistic view of node centrality within the network (Bonacich, 2007). This makes it particularly useful for identifying nodes that have indirect influence or are well-positioned within the network. finally, Eigenvector centrality is less susceptible to manipulation compared to degree centrality. In degree centrality, one can artificially increase their centrality by connecting to many low-degree nodes. However, in eigenvector centrality, the importance of a node depends not only on its connections but also on the centrality of its neighbours. This makes it more robust against intentional manipulations of the network structure (Bonacich, 2007).

Inversely, eigenvector centrality calculation requires expensive computation power. The process involves solving an eigenvector equation, which may require significant computational resources (Newman, 2018). This limits the practicality of eigenvector centrality in analyzing very large networks. Eigenvector centrality can also be influenced by the size of the network. In very small networks, eigenvector centrality may result in equal or similar centrality values for all nodes, providing limited discrimination between nodes (Newman, 2018). On the other hand, in very large networks, the eigenvector calculation may become computationally challenging, limiting the scalability of the measure. Moreover, it is primarily designed for undirected networks. While adaptations have been proposed for directed networks (e.g., PageRank algorithm), eigenvector centrality may not fully capture the complexities of directed interactions and influence flows in such networks (Newman, 2018). Lastly, Eigenvector centrality heavily relies on the network's structure, particularly the connectivity patterns and node interactions. In networks with certain structural

characteristics, such as disconnected components or highly clustered nodes, eigenvector centrality may not accurately capture the importance of nodes (Newman, 2018). It may assign high centrality to nodes that are part of small, isolated clusters or penalize nodes in sparsely connected regions of the network.

Network analysis application for vaccine supply chains

Several studies have used network analysis to assess the complexity of the vaccine supply chain. For example, Li et al. (2021) used social network analysis to map the vaccine supply chain in China and identified the key nodes and links in the chain. Zingales et al. (2021) used network analysis to identify the critical nodes and links in the vaccine supply chain in Italy. In the context of the Covid-19 pandemic, multiple studies have applied network analysis to the Covid-19 vaccine supply chain, using data from various sources to map the relationships between key actors and nodes.

Network analysis can also be used to evaluate the effectiveness of vaccine distribution strategies. For example, Liu et al. (2021) used network analysis to evaluate the impact of different distribution strategies on the vaccine supply chain in China. A study by Wang et al. (2021) used social network analysis to examine the distribution of Covid-19 vaccines in China. The study found that the vaccine distribution network was highly centralized, with a few key players controlling the majority of vaccine distribution channels. The study also identified several bottlenecks in the supply chain, including insufficient transportation capacity and inadequate vaccine storage facilities. De Giovanni and Scalera (2021) conducted a comprehensive review of the Covid-19 pandemic using network analysis and identified the bottlenecks in the vaccine supply chain. Zhu et al. (2021) used network analysis to identify the critical nodes and links in the vaccine supply chain in China and proposed strategies for improving the efficiency of the chain.

Another study by Hu et al. (2021) used network analysis to examine the distribution of Covid-19 vaccines in the United States. The study analyzed data from the Centers for Disease Control and Prevention (CDC) and identified several key nodes in the vaccine distribution network, including vaccine manufacturers, distributors, and healthcare providers. The study found that the network was highly centralized, with a few key players controlling most vaccine distribution channels. The study also identified several areas for improvement, including increasing the number of distribution channels and improving vaccine allocation strategies.

A study by Canales et al. (2021) used network analysis to examine the distribution of Covid-19 vaccines in Latin America and the Caribbean. The study analyzed data from the Pan American Health Organization (PAHO) and identified several key nodes in the vaccine distribution network, including vaccine manufacturers, national regulatory authorities, and healthcare providers. The study found that the network was highly fragmented, with limited coordination and communication between key actors. The study also identified several challenges in the supply chain, including limited vaccine supply, inadequate transportation capacity, and insufficient vaccine storage facilities.

In addition to mapping the relationships between key actors and nodes in the Covid-19 vaccine supply chain, network analysis can also be used to evaluate the performance of the supply chain. A study by Wu et al. (2021) used network analysis to assess the resilience of the Covid-19 vaccine supply chain in China. The study analyzed data from the National Health Commission of China and identified several key nodes in the vaccine distribution network, including vaccine manufacturers, distributors, and healthcare providers. The study found that the network was highly resilient, with multiple backup nodes and redundant pathways for vaccine distribution. The study also identified several factors that

contributed to the resilience of the supply chain, including strong government support, effective communication and coordination between key actors, and robust vaccine storage and transportation infrastructure.

Questions

From the literature review, we were able to formulate three questions:

- **Q1**: By applying network analysis to the Covid-19 vaccine supply, are we able to represent it as a complex model?
- **Q2**: Can network analysis identify bottlenecks, inefficiencies, and redundant nodes within the Covid-19 vaccine supply chain?
- **Q3**: Will applying Centrality measurement techniques to the Covid-19 vaccine supply chain reveal key nodes and their significance, provide insights into critical entities and improve the overall resilience and efficiency of the supply chain?

3.3 Methodology

This study uses a quantitative approach using R to analyze the Covid-19 vaccine supply chain as of June 2021 using network analysis. The networks created are undirected and all the data used for this project was collected from the article of Bown and Bollyky on the Covid-19 vaccine titled "How Covid-19 vaccine supply chains emerged amid a pandemic" (2021). The article identified companies involved in the supply chain of six Covid-19 vaccine producers: Pfizer, Moderna, Johnson&Johnson, AstraZeneca, Novavax, and CureVac. From this article, we created three supply chain datasets, one for Pfizer, one for Moderna and the last one for all the companies present in all six Covid-19 vaccine producer supply chains. The data will include information on the vaccine supply chain, such as the key company involved in both supply chains and the stage of production they are involved in (formulation, lipid production & fill and finish).

The data will be analyzed using network analysis techniques to map the Covid-19 vaccine supply chain for both Pfizer and Moderna. The analysis will include the following steps:

1. Identification of key nodes/links in the vaccine supply chain: The nodes represent both the key actors involved in the vaccine supply chain and the three stages of production presented in the vaccine supply chain. The links represent the relationships between the company and the stage of production nodes, such as that one company is in charge of the fill and finishing in the chain. Bridging nodes/links will be deemed as the key nodes/links.
2. Identification of key nodes/links in the vaccine supply chain: The nodes represent the key actors involved in the vaccine supply chain. The links represent the relationships between the nodes, such as the flow of vaccines between different actors in the production sequence. Bridging nodes/links will be deemed as the key nodes/links.
3. Assessment of network structure: The network structure will be assessed based on measures such as degree centrality, betweenness centrality, closeness centrality, eigenvector centrality and subgraph centrality. These measures will help to identify the most important nodes in the

network and the bottlenecks in the vaccine supply chain. These measures will then be used to re-create the second network with the centrality measures.

Basic Network

As mentioned previously, our first group of networks will map the links between companies and their function in the chain. This will be done for Pfizer SC, Moderna SC and the six main vaccine producers, creating in total three of the first networks. For this part of the project, the nodes were both the companies in the chain and the stage of production (formulation, lipids production, adjuvant production and fill & finish. While the edges represent the belonging of a company to a production stage. This was done by creating a dataset with the two variables: "From" and "To". The "From" variable represents the companies while the "TO" variable is the stage of production as presented in *Table 3-1*.

Table 3-1: Format of the dataset for the first network of the relation between the stage of production and the companies

From	To
Pfizer	Lipids production
Moderna	Formulation
Catalent	Fill and Finish

The same company can be involved in multiple stages of production. This was the case for Pfizer, and other companies, which were involved in all three stages of vaccine production. Pfizer had three facilities that were in charge of formulation, three other facilities for fill and finish and one location in charge of lipids production. This multi-stage element was considered by creating multiple observations for the same company. In the case of Pfizer, seven observations were created, three of them leading to the formulation node, three to the fill and finish and one leading to the production of the lipids as shown in *Table 3-2*.

Table 3-2: Example of the dataset of multi-relation between the stage of production and the companies

From	To
Pfizer	Formulation
Pfizer	Formulation

From	To
Pfizer	Formulation
Pfizer	Fill and finish
Pfizer	Fill and finish
Pfizer	Fill and finish
Pfizer	Lipids formulation

The second group of networks shows the movement of the vaccine and vaccine components in the supply chain of Pfizer, Moderna and the six main vaccine producers. Contrary to the first network, the node of this network is the companies only and the edges is the movement of product between companies.

The movement of the product follows this order: lipids producer send their product to the formulation manufacturer, as lipid is one of their inputs to create the vaccine, then the formulation manufacturer send the finished vaccine to the fill and finish facilities, where the vaccine is bottled and sealed. To simplify this project, we are assuming that all the lipid producers can send their products to all the formulation manufacturers and that all formulation manufacturers are sending their products to all the fill and finish facilities. This assumption was crucial when creating our three new datasets for the second network. Taking Pfizer as an example, *Table 3-3* shows what the dataset looked like for the movement of vaccines between Pfizer's facilities. This assumption generated a weighted network for the vaccine supply chain.

Table 3-3: Format of the dataset for the second network of the flow of Covid-19 vaccine within its supply chain

From	To
Pfizer (Lipids)	Pfizer (formulation 1)
Pfizer (Lipids)	Pfizer (formulation 2)
Pfizer (Lipids)	Pfizer (formulation 3)
Pfizer (formulation 1)	Pfizer (fill and finish 1)
Pfizer (formulation 1)	Pfizer (fill and finish 2)
Pfizer (formulation 1)	Pfizer (fill and finish 3)
Pfizer (formulation 2)	Pfizer (fill and finish 1)
Pfizer (formulation 2)	Pfizer (fill and finish 2)
Pfizer (formulation 2)	Pfizer (fill and finish 3)
Pfizer (formulation 3)	Pfizer (fill and finish 1)
Pfizer (formulation 3)	Pfizer (fill and finish 2)
Pfizer (formulation 3)	Pfizer (fill and finish 3)

Centrality Measurement

For the second part of the project, centrality measures are important parts of network analysis as gives information on the importance of certain node over other. We decided to only perform measure the centrality of the network of the flow of vaccines with the supply chain rather than the network of the stage of production, as we believe that this measurement is more relevant for the flow of vaccines between companies. To illustrate our steps for this part of our network analysis, we will use Pfizer's second dataset as an example after the basic network step has been completed for Pfizer's second network.

The first step is to create a new dataframe with all the node and their centrality measures in a new code chunk. *Table 3-4* represent the five centrality measures used in this project with their equation: degree centrality, betweenness centrality, closeness centrality, eigenvector centrality and subgraph centrality.

Table 3-4: Degree centrality, betweenness centrality, closeness centrality, eigenvector centrality and subgraph centrality with their formulas

Centrality measure	Equation
Degree centrality	$\deg(v) = \lvert w \in V : (v,w) \in E \rvert$
Betweenness centrality	$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma(s,t\lvert v)}{\sigma(s,t)}$
Closeness centrality	$C_C(v) = \frac{1}{\sum_{u \in V} d(u,v)}$
Eigenvector centrality	$C_E(v) = \frac{1}{\lambda} \sum_{u \in (v)} C_E(u)$ $Ax = \lambda x$
Subgraph centrality	$C_s(v) = \sum_{k=2}^{n} \frac{1}{(k-1)!} \sum_{S \subseteq N_k(v)} I(G_S)$

To create a network with the node's centrality measure, we first need to create a dataset with the nodes and their centrality measures as variables using the formulas in the table above. Once this step is completed, we can plot our network with the size of the node depending on its degree centrality score. This means that a node with a high degree score will be bigger than a node with a smaller score. Additionally, the width of an edge depends on its betweenness score ratio, therefore a higher ratio means a thicker line.

3.4 Results

In the following section, the results of our networks will be discussed and analyzed. The results are presented in three parts one for each supply chain. For each network, we will start by analyzing the structure of the network by identifying the structural holes in the networks. Then we will discuss the centrality measures of each node of the network.

Pfizer

Figure 3-1 is the first network created for Pfizer, where graph A is the network with the company's name and graph B is the structure of Pfizer's network. We can see that the plotted network resembles a tree type network with a central node. In this network, only Pfizer fills the structural hole present. It is the only company that takes part in all the stages of production for vaccine manufacturing. However, to a different degree, we can see that Pfizer has three facilities that oversee the formulation and the fill and finish, while they only have one facility that produces lipids. With only one company acting as a broker, this makes that network less robust to change when deleting enough nodes and/or links.

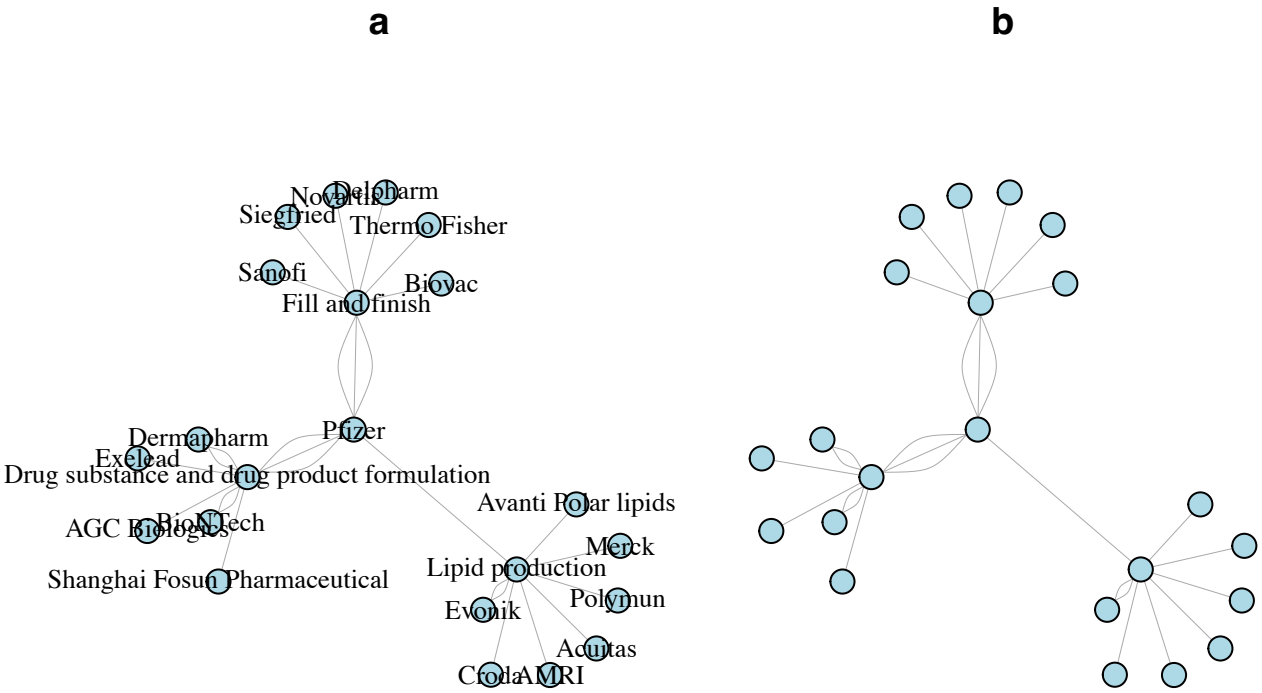


Figure 3-1: Network Graph with Stage of Production - Pfizer

A second network was created for Pfizer, and it resembles a star shape network like Moderna's second network (*Figure 3-2*). Graph A, in *Figure 3-2*, is the network with the company's name and graph B is the structure of Pfizer's network. We can also see that the companies that are in charge of the fill and finish and the lipids production are located on the outer bound of the network, while the formulation is embedded within the network. BioNTech, Pfizer and Dermapharm are located at the center of the network.

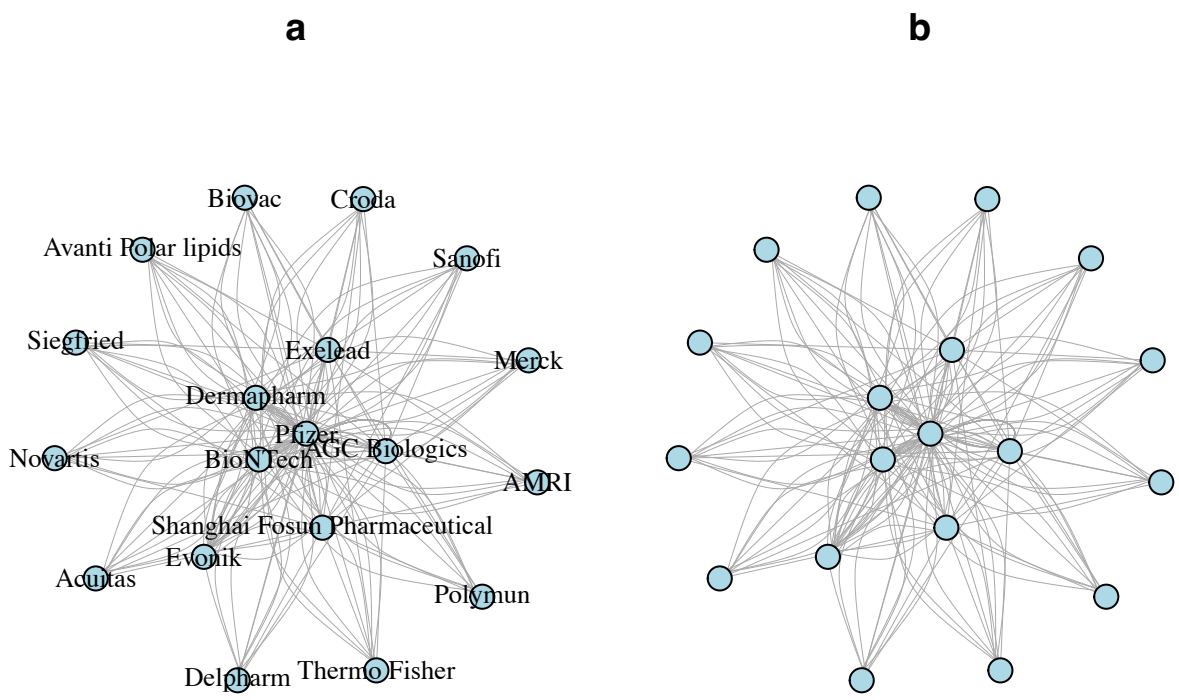


Figure 3-2: Network Graph - Pfizer

When looking at the centrality measures of our second network represented in *Figure 3-3*, we can see that Pfizer has the highest degree of centrality, betweenness centrality, eigenvector centrality and subgraph centrality, with values of 94, 40.1, 0.05555556, 1 and 47183999272 respectively. BioNTech has an eigenvector measure of 0.3245553 which means that it is connected to many nodes that themselves have high eigenvector scores. AGC Biologics, Exelead, BioNTech, Dermapharm and Shanghai Fosun Pharmaceutical, all formulation companies, have the second highest closeness centrality measure of 0.04545455, this means that those a node acts as a bridge between other nodes in the network. With the centrality measures of each node, we create a final network of Pfizer's supply chain, which resulted in *Figure 3-4*. *Figure 3-4* emphasize important nodes by enlarging them based on their degree of centrality, where graph A is the network with the company's name and graph B is the structure of the network. Thus, representing Pfizer with a larger node has it has the highest degree of centrality measure in its supply chain as seen in *Figure 3-3*.

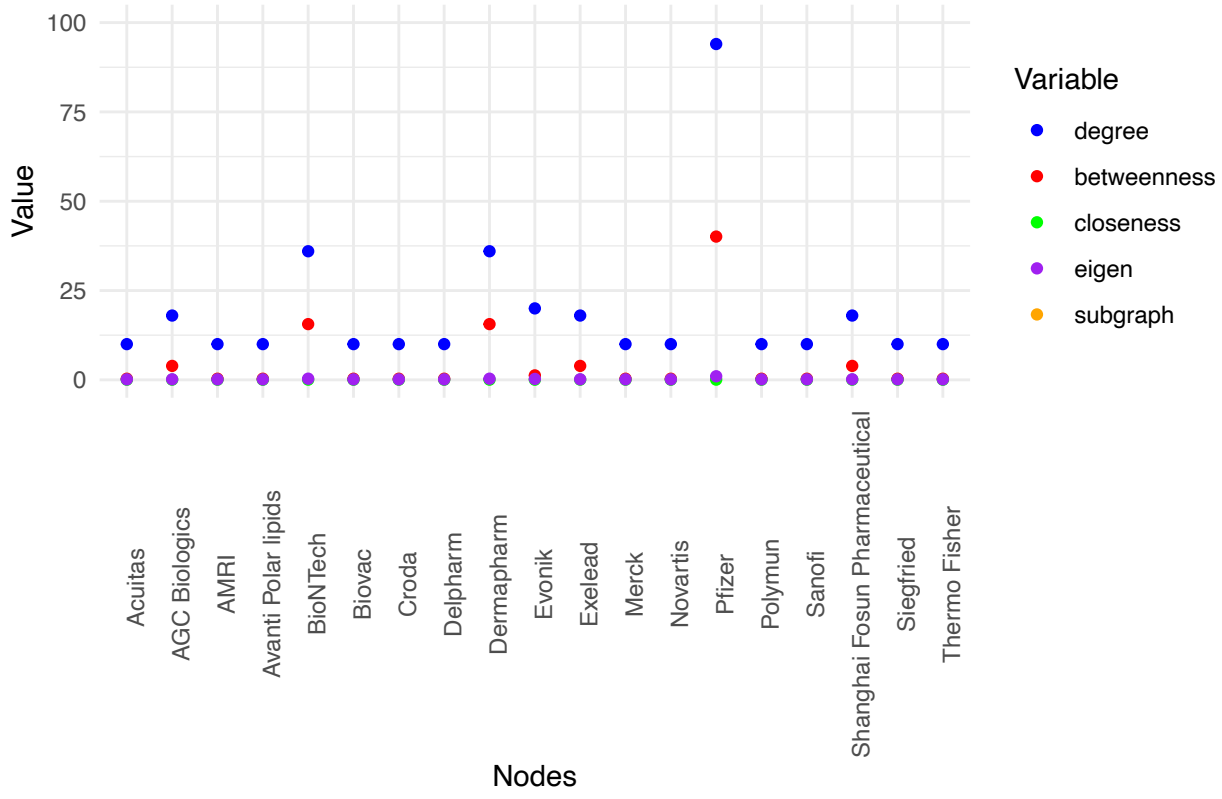


Figure 3-3: Centrality Measures - Pfizer Supply chain

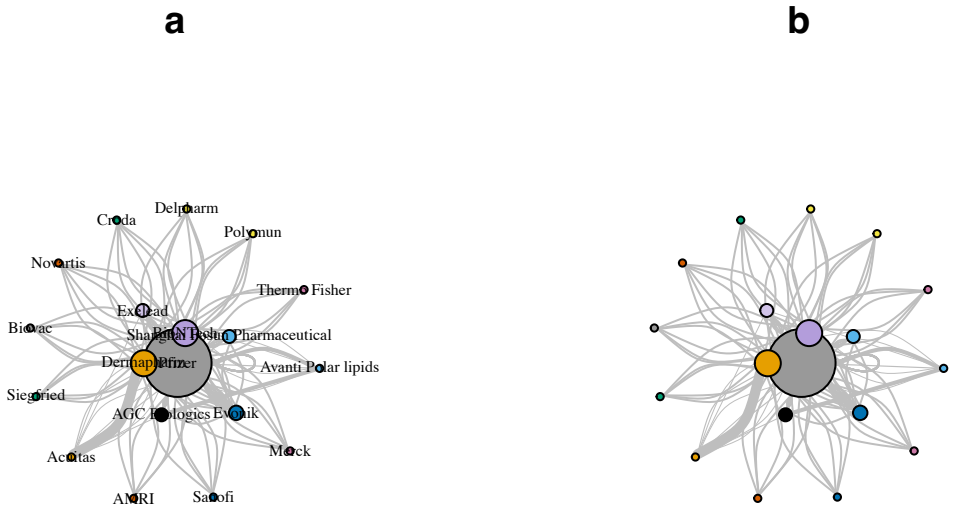


Figure 3-4: Network Graph with Centrality Measures - Pfizer

Moderna

When looking at the first network created for Moderna, we can see that the plotted network resembles a tree type network (*Figure 3-5*). Graph A, in *Figure 3-5*, is the network with the company's name and graph B is the structure of Moderna's network. We can also see that Rovi is the only company in Moderna's supply chain that fill the structural hole between the fill and finish and the formulation. Additionally, only one company, Corden Pharma, is allocated to the lipids production of their vaccine this part of the supply chain is also disconnected from the rest of the network. This disconnection indicates that the limited number of companies in the fill and finish is a bottleneck present in Moderna's Covid-19 supply chain. The lack of bridging in this network affects the production and the capacity of certain key nodes or links were to disappear. For example, if one of Corden Pharma's facilities were to close down, this would slow down Moderna production.

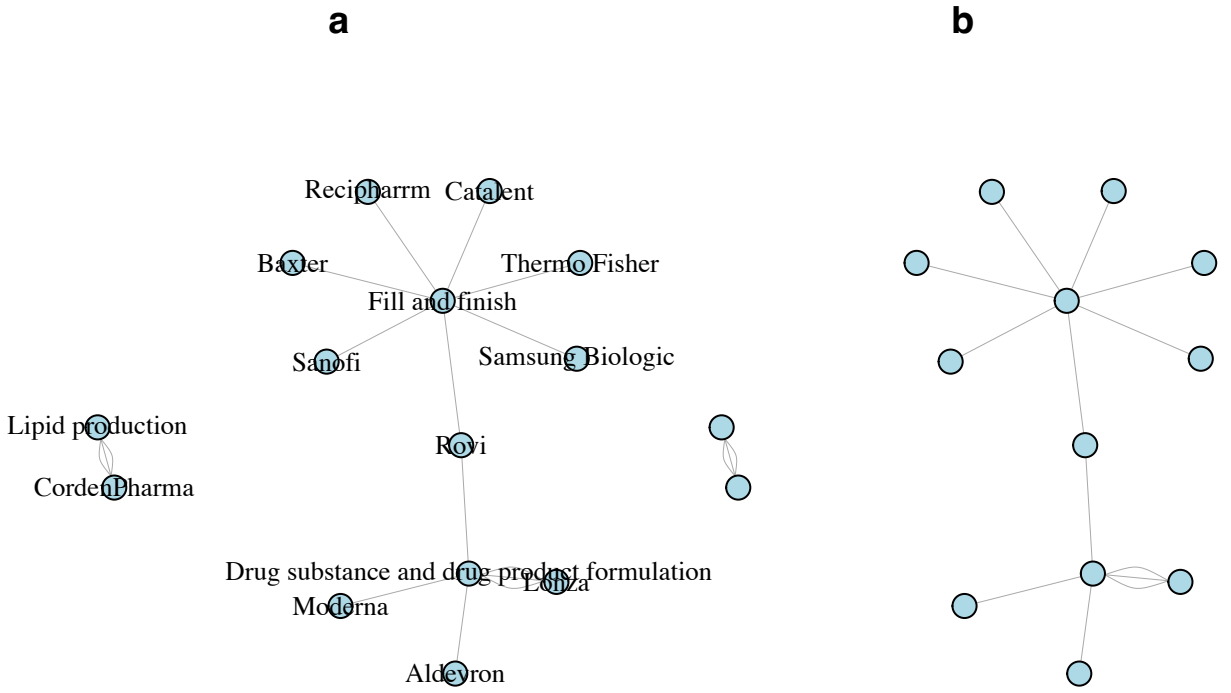


Figure 3-5: Network Graph with Stage of Production - Moderna

Figure 3-6 is the second network where graph A is the network with the companies' names and graph B is the structure of Moderna's network. *Figure 3-6* presents Moderna's supply chain in a form that resembles most of a star shape network. We can also see that the companies that are in charge of the fill and finish are located on the outer bound of the network, while the formulation and the lipids production are embedded within the network. Additionally, Lonza, one of the formulation companies, is located at the center of the network.

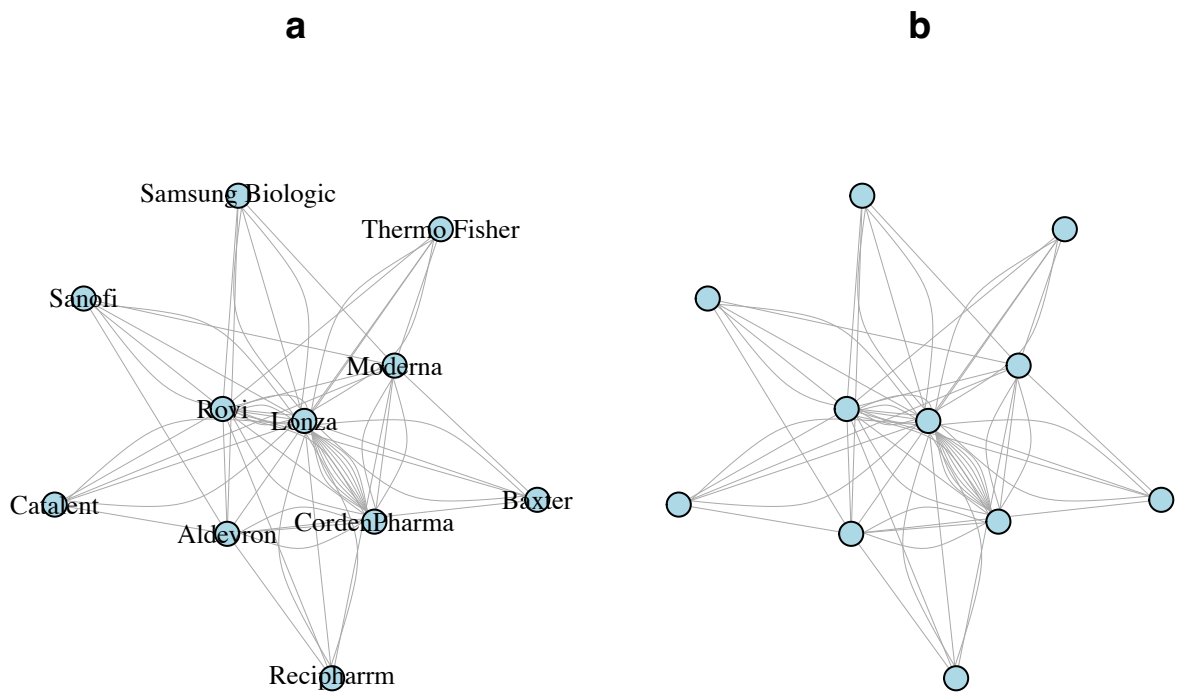


Figure 3-6: Network Graph - Moderna

If we look at the centrality measures of our second network, we can see that Lonza has the highest degree of centrality, betweenness centrality, eigenvector centrality and subgraph centrality, with values of 30, 15.75, 0.08333, 1 and 803107.66 respectively, as represented in *Figure 3-7*. Furthermore, CordenPharma has an eigenvector measure of 0.866 which means that it is connected to many nodes that themselves have high eigenvector scores. Moderna, Aldevron has the same closeness centrality measure as Lonza, this means that a node acts as a bridge between other nodes in the network. With the centrality measures of each node, we create a final network of Moderna's supply chain, which resulted in *Figure 3-8*, graph A and graph B the network with the company's name and the structure of the network. *Figure 3-8* emphasize important nodes by enlarging them based on their degree of centrality. Thus, representing Lonza with a larger node has it has the highest degree of centrality measures in Moderna's supply chain as seen in *Figure 3-7*.

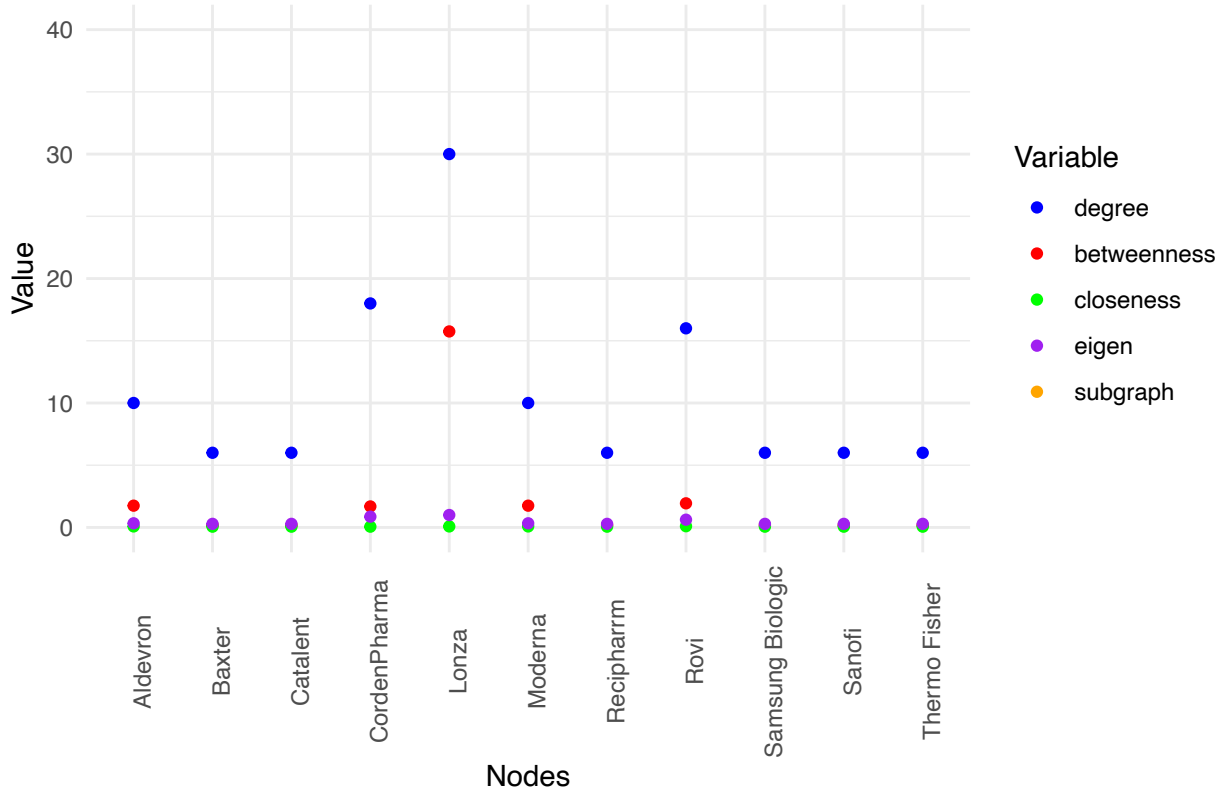


Figure 3-7: Centrality Measures - Moderna Supply chain

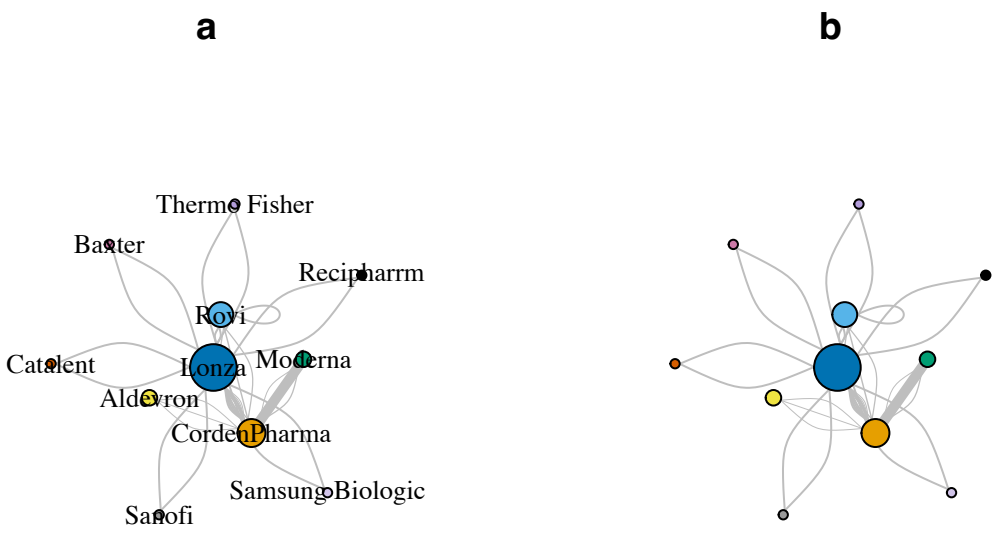


Figure 3-8: Network Graph with Centrality Measures - Moderna

All Covid-19 vaccine producers

When looking at the first network created for all Covid-19 vaccine producers, we can see that the plotted network resembles a tree type network (*Figure 3-9*). Graph A, in *Figure 3-9*, is the network with the company's name and graph B is the structure of the network. Only seven nodes fill the structural hole present in this network. Two of the seven companies, Merck, and Pfizer, take part in the three stages of production for vaccine manufacturing (lipids production, formulation & fill and finish), one, Biologics, is involved in two stages of production for vaccine manufacturing (adjuvant production and formulation), and four, Catalent, Rovi, CSL and FUJIFILM Diosynth Biotechnologie, are involved in two stages of production for vaccine manufacturing (formulation & fill and finish). However, we can see that Pfizer has three facilities that are in charge of the formulation and the fill and finish, while they only have one facility that produces lipids while Merck has one facility for each stage of production. With only a few companies acting as brokers, this makes that network less robust to change when deleting key nodes and/or links.

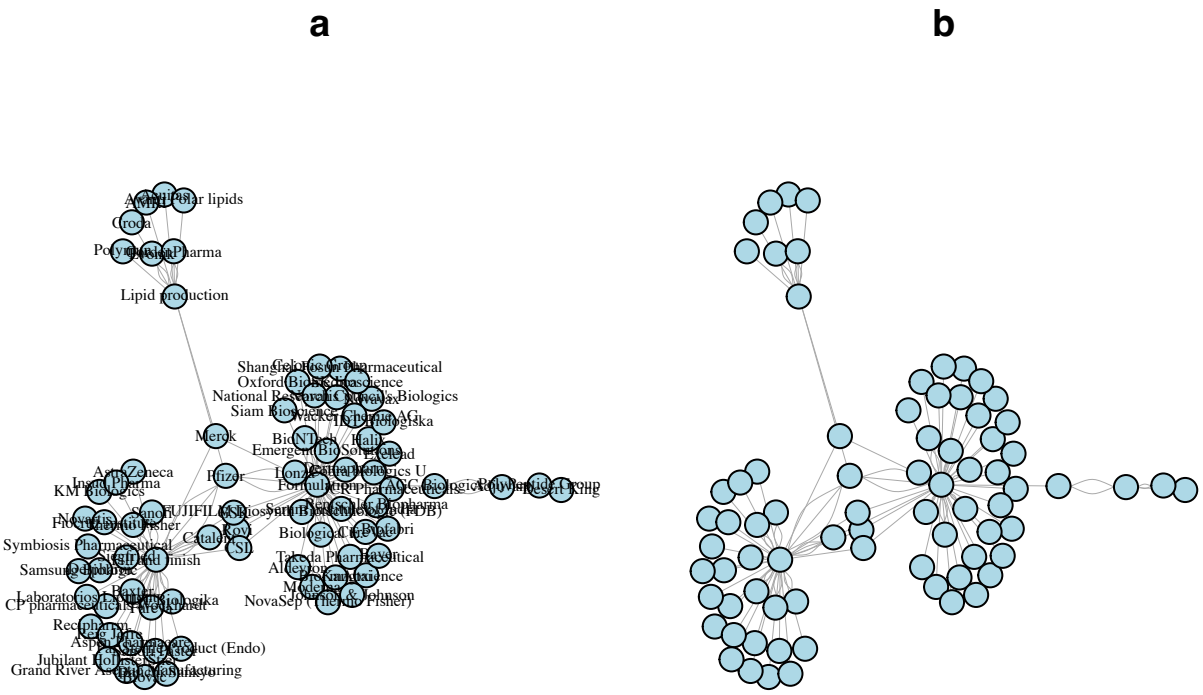


Figure 3-9: Network Graph with Stage of production - All Covid-19 vaccine producers

Figure 3-10 is the second network of all the Covid-19 vaccine producers, where graph A is the network with the company's name and graph B is the structure of the network. *Figure 3-10* presents all Covid-19 vaccine producers as a single supply chain and resembles a complete network. While most nodes are grouped, three nodes - Desert King, Biofabri and PolyPeptide Group - are separated from the group.

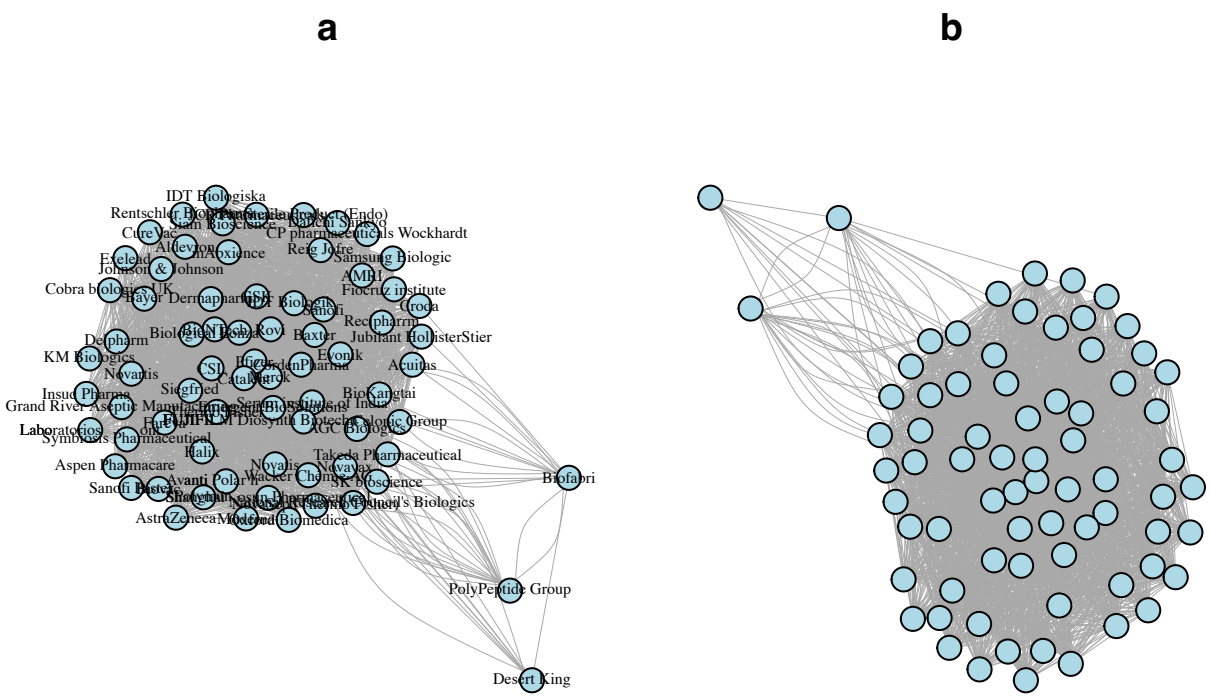


Figure 3-10: Network Graph - All Covid-19 vaccine producers

When looking at the centrality measures of our second network, shown in *Figure 3-11*, we can see that Pfizer has the highest degree of centrality and betweenness centrality, with values of 363, and 182.04 respectively. We can also see that the three company that located outside of the cluster have the smallest degree of centrality and closeness centrality. Polypeptide Group has a degree centrality of 20 with a closeness centrality of 0.006, Desert King has a degree centrality of 10 with a closeness centrality of 0.006 and Biofabri has a degree centrality of 17 with a closeness centrality of 0.0074. With the centrality measures of each node, we create a final network of the supply chain, which resulted in *Figure 3-12* where graph A is the network with the companies' names and graph B is the structure of the network. *Figure 3-12* emphasize important nodes by enlarging them based on their degree of centrality. Thus, representing Pfizer with a larger node has it has the highest degree of centrality measure in its supply chain as seen in *Figure 3-11*.

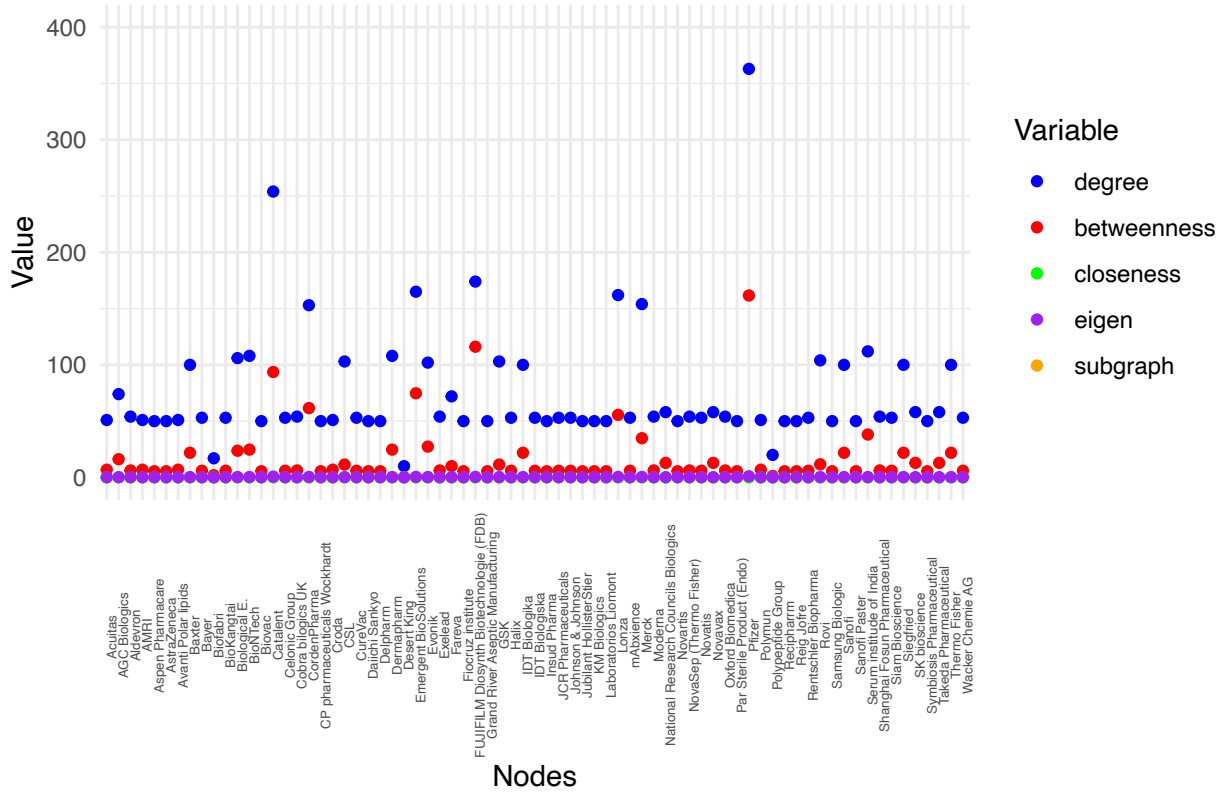


Figure 3-11: Centrality Measures - All Covid-19 vaccine producers

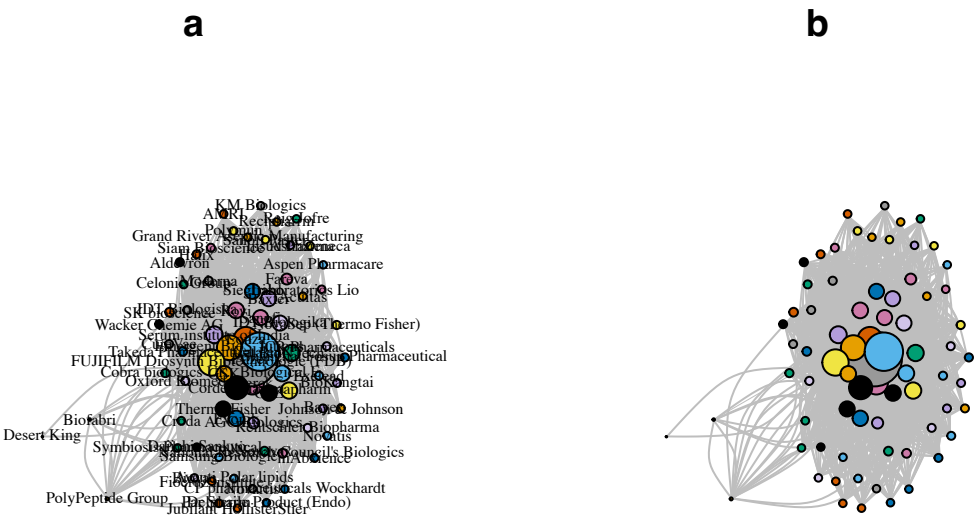


Figure 3-12: Network Graph with Centrality Measures - All Covid-19 vaccine producers

3.5 Discussion

The Covid-19 pandemic has presented unprecedented challenges to healthcare systems, economies, and societies worldwide. Covid-19 vaccine development, manufacturing, and distribution are critical to combating the virus and its impact. The complicated vaccine supply chain involves multiple players, complex logistics and global distribution networks. Innovative approaches are therefore needed to improve supply chain efficiency, resilience, and fairness.

Network analysis has proven to be an invaluable tool in optimizing the distribution and delivery of the Covid-19 vaccine. By examining the intricate connections among various stakeholders, including vaccine manufacturers, distributors, healthcare providers, and government agencies, network analysis provides a comprehensive understanding of the vaccine supply chain. This analysis enables policymakers and healthcare organizations to identify bottlenecks, predict potential shortages, and develop targeted strategies to ensure efficient vaccine delivery. By visualizing vaccine flow and identifying critical nodes within the network, network analysis helps streamline the distribution process, minimize waste, and ensure that vaccines reach the communities that need them most. It plays a critical role in maximizing the impact of vaccination efforts and ultimately contributes to the global fight against the Covid-19 pandemic. With the literature, we formulated three questions; H1) By applying network analysis to the Covid-19 vaccine supply, we can represent it as a complex model, H2) Network analysis can identify bottlenecks, inefficiencies, and redundant nodes within the Covid-19 vaccine supply chain and H3) Centrality measurement techniques applied on the Covid-19 vaccine supply chain will reveal key nodes and their significance, providing insights into critical entities and improving the overall resilience and efficiency of the supply chain

As network analysis is being used more and more to understand the complex patterns of interconnected entities in the Covid-19 vaccine supply chain in recent papers, this paper aims to complete the existing literature on the application of network analysis on the Covid-19 vaccine supply chain, by 1) mapping the supply chain as a weighted network, by 2) mapping a weighted network of the function and companies to identify the vulnerable function in the chain and by 3) providing a network that features the centrality measures in the mapping. Thus, contributing to the existing

literature by giving a different perspective in analyzing the key nodes and links in the Covid-19 supply chain.

In this project, network analysis was used to identify bottlenecks in the VSCs of Pfizer, Moderna and the six major Covid-19 vaccine manufacturers combined. After collecting the relevant data, we mapped our networks in R. We did this by mapping 1) the relationship between the companies and their function in the supply chain, 2) the movement of vaccines between companies in the chain, and 3) centrality with the second network.

After creating our network, we first analyzed the structure of the network, which helped us identify the structural holes within the network. We then examined the centrality measures to identify key nodes and potential bottlenecks. With the first group networks, we found that structural holes are filled by a few companies making the whole supply vulnerable if those nodes or links were to be disrupted. In the cases of Moderna's Network Graph with Stage of Production, the structural hole between lipids production and the rest of the network remains unfilled, isolating a function of the supply chain. Furthermore, we also found that the second group of networks has a structure that is like a star shape network. This tells us that the network is centralized, which provides control and easy management. However, this structure makes the network susceptible to disruption if the central nodes were to fail. The centrality measure helps to identify the central and key nodes of our networks using degree centrality, betweenness centrality, closeness centrality, eigenvector centrality and subgraph centrality. The results of our research validate all our questions. As shown in figures 1 to 8, the Covid-19 vaccine supply chain is represented as a complex network when applying network analysis techniques, thus supporting Q1. Using both network analysis and centrality measurement we were able to identify the cause of bottlenecks, redundancy, and inefficiencies in the Covid-19 VSC, thus confirming Q2 and Q3.

The main conclusion of this research is that bottlenecks in the Covid-19 vaccine supply chain have two causes: the absence of gateways in the network, and the dependence on certain companies to increase production capacity. Pfizer is the only company to establish gateways between the different production stages of its networks, making it one of the main nodes in the network. In the case of the Moderna supply chain, Lonza owns three facilities that increase production capacity, making Lonza a key node in the network. In addition, Rovi and CordenPharma are also key nodes in the Moderna network, as Rovi is the only link node in the network and CordenPharma is the only lipid producer in the chain. In conclusion, to reduce risks and limit bottlenecks in CVPs, pharmaceutical companies should reduce their dependence on key companies/suppliers, by increasing the number of production facilities, and increasing the number of relay companies in their CVPs, by ensuring that several companies operate at several production stages.

In conclusion, network analysis is a powerful tool for studying the Covid-19 vaccine supply chain. By mapping the relationships between key actors and nodes in the supply chain, network analysis can provide valuable insights into the efficiency, resilience, and performance of the system. It can help identify bottlenecks and opportunities for improvement, and inform decision-making by policymakers, manufacturers, and distributors. Overall, network analysis has the potential to play a critical role in improving the Covid-19 vaccine supply chain and ensuring equitable access to vaccines. As the world continues to grapple with the pandemic, researchers, policymakers, and stakeholders need to continue exploring innovative ways to analyze and improve the vaccine supply chain.

This study is not without limitations. This project may be limited by the availability and reliability of data, especially in low-income countries with limited health infrastructure. The Covid-19 vaccine

supply chain is a rapidly changing and dynamic system, which may make it difficult to capture and analyze all relevant data points. Additionally, changes to the supply chain (e.g., new vaccine manufacturers, changes to distribution channels) may not be reflected in the data promptly. Furthermore, network analysis can only capture relationships between actors and nodes that are included in the dataset and other factors that may impact vaccine distribution, such as political and economic factors, may not be captured in the analysis. The findings of a network analysis of the Covid-19 vaccine supply chain may not be directly applicable to other vaccine supply chains or public health challenges, as this supply chain is in the context of the early response global pandemic. Researchers need to acknowledge and address these limitations in their studies to ensure that the findings are accurate, relevant, and useful for informing policy and practice in the vaccine supply chain.

Future directions for this research would include linkage probability based on geographic localization to the network. As stated previously, some of our networks were created under the assumption that all companies/facilities are sending their product to all the others following the vaccine production flow. However, this is unlikely since facilities within the same geographic location have a high chance of sending their product to one another, due to their proximity factor. On the other hand, sending products from one geographic location to another will serve as a support or safety net, this was the case for one of the fill and finish companies present in Pfizer's supply chain. Therefore, lowers the probability of such links in the network.

References

Baden, L. R., El Sahly, H. M., Essink, B., Kotloff, K., Frey, S., Novak, R., … & Follmann, D. (2021). Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. New England Journal of Medicine, 384(5), 403-416.

Canales, C., Ortiz, M., Rodríguez, P., & Lagos, R. (2021). A network approach to understanding Covid-19 vaccine distribution in Latin America and the Caribbean. PLoS ONE, 16(4), e0249883.

Goldratt, E. M., & Cox, J. (1984). "The Goal: A Process of Ongoing Improvement." North River Press.

Haji Ali Afzali, H., Alemi, H., Khodayari, Zarnaq, R., & Alavi, S. S. (2021). Network analysis of Covid-19 vaccine supply chain in Iran: a social network analysis. Journal of Pharmaceutical Policy and Practice, 14(1), 1-8.

Hopp, W. J., & Spearman, M. L. (2000). "Factory Physics: Foundations of Manufacturing Management." Irwin/McGraw-Hill.

Hu, T., Liao, S., Wu, T., Xue, H., & Zhao, Y. (2021). Network analysis of Covid-19 vaccine distribution in the United States. Journal of Medical Systems, 45(6), 1-7.

Krause, P. R., Fleming, T. R., Longini, I. M., & Peto, R. (2021). Covid-19 vaccine trials should seek worthwhile efficacy. The Lancet, 397(10284), 1607-1609.

Lee, H. L., & Billington, C. (1992). "Managing supply chain inventory: Pitfalls and opportunities." Sloan Management Review, 33(3), 65-73

Notteboom, T., & Rodrigue, J. P. (2005). "Port regionalization: Towards a new phase in port development." Maritime Policy & Management, 32(3), 297-313

Osterholm, M. T., Kelley, N. S., Sommer, A., & Belongia, E. A. (2021). Efficacy and effectiveness of Covid-19 vaccines. New England Journal of Medicine, 385(23), 2338-2349.

Ozawa, S., Stack, M. L., Publicover, N. G., & Walker, D. G. (2017). The economics of vaccine supply chains in developing countries. Vaccine, 35(17), 2279-2285.

Polack, F. P., Thomas, S. J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., … & Pérez, J. L. (2020). Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. New England Journal of Medicine, 383(27), 2603-2615.

Ribeiro, H. V., Rodrigues, F. A., Oliveira, M. J., & Zorzenon dos Santos, R. M. (2017). "Mapping accessibility in the Brazilian vaccination coverage network: A spatial and network analysis." PloS ONE, 12(6), e0179593

Smith, J. C., & Sivaramakrishnan, K. (2020). Covid-19 and vaccine logistics: Challenges and opportunities. Journal of Global Health, 10(2).

Takian, A., Kazempour-Ardebili, S., & Zakeri, H. (2020). "Artificial intelligence for COVID-19 vaccine supply chain management." Journal of Medical Systems, 44(8), 135.

Verma, S., Gokhale, M., & Bhatia, R. (2021). Global vaccine supply chains: Current status, challenges and opportunities. Vaccine, 39(7), 1016-1027.

Wang, J., Jing, R., Lai, X., Zhang, H., Lyu, Y., Knoll, M. D., & Fang, H. (2021). Distribution of Covid-19 vaccine in China: a social network analysis of distribution channels. Vaccine, 39(7), 1052-1056.

World Health Organization. (2020). Guidelines on the management of vaccine supply chains during the Covid-19 pandemic. Geneva: World Health Organization.

Wouters, O. J., Shadlen, K. C., Salcher-Konrad, M., Pollard, A. J., Larson, H. J., Teerawattananon, Y., … & Glassman, A. (2021). Challenges in ensuring global access to Covid-19 vaccines: production, affordability, allocation, and deployment. The Lancet, 397(10278), 1023-1034.

Wu, S., Li, T., Liu, Y., Wu, L., Zhang, Z., & Liu, W. (2021). Resilience assessment of Covid-19 vaccine supply chain in China: A network analysis. Vaccine, 39(22), 2943-2950.

Xiong, H., Chen, H., & Zhao, Y. (2021). A comprehensive review of the global vaccine development initiatives in response to the Covid-19 pandemic. Human Vaccines & Immunotherapeutics, 17(6), 1846-1854.

Yadav, P., & Lyngdoh, T. I. (2020). Covid-19 vaccine supply chain: Challenges and strategies. Indian Journal of Public Health, 64(Supplement), S152-S156.

Zhang, Y., Li, Y., & Lu, J. (2021). From emergency to sustainability: Building resilient vaccine supply chains against Covid-19 and future pandemics. Globalization and Health, 17(1), 1-10.

Conclusion

This project looks at the initial stages of a vaccine supply chain created amid a global pandemic, and the lessons that can be learned from such an event. With global pandemics, similar to SARS-CoV-2 predicted to become more and more likely, this project aims to steer research in the right direction to learn as much as possible about a pandemic vaccine supply chain, which has not been the case in the literature on vaccine supply chains pre-Covid-19.

This thesis was presented in the form of three articles that analyze the emerging Covid-19 VSC. In the first article, we were able to categorize the upstream and downstream dimensions of six main Covid-19 vaccine producers and found that none of them were global in both dimensions. In fact, the majority of producers were classified as having both a bi-regional upstream production and downstream distribution which tells us those producers are susceptible to market concentration risk, dependency on regional factors, and limited knowledge accessibility. In the second article, we used panel data analysis and found that Beck & Katz model and Arellano & Bond model identify different factors that influence the distribution of Pfizer's vaccine, Moderna's vaccine and the total of Pfizer and Moderna vaccine in a country. For Moderna's vaccine, the lag(number_lipids) and the lag(number_fill) are both significant at the 10 percent level, the lag(reg_moderna,1) is significant at the 5 percent level and the lag(number_cov_chain_moderna,1) is significant at the 1 percent level. For Pfizer's vaccine, the number of Covid-19 vaccine formulation facilities and the number of fill & finish facilities in the country are significant at the 5 percent level. For the total vaccine, the number of fill & finish facilities in the country and the variable of Pfizer Rugman's classification are significant at the 10 and 5 percent level (Beck & Katz model) and the lag of the number of Moderna SC facilities present in the country is significant at the 1 percent level (Arellano & Bond model).

The last article found that both Pfizer and Moderna, despite going with a different strategy to increase production, both supply chains as star-sharped networks with few nodes centralized nodes. In both Pfizer cases, they and BioNTech, their partner, are at the center of the network. However, in Moderna's chain, Lonza, an outsourced formulation company, that located at the center of its network. Another finding of this article is that only two companies have facilities that operated in two or more stages of vaccine production. This company is Pfizer for their network and Rovi for Moderna's. Moderna's supply chain relies solely on one company, Corden Pharma for the lipid production required for their vaccine, making it the bottleneck of this chain.

The study of the COVID-19 vaccine supply chain proved to be of paramount importance in managing and mitigating the global pandemic. The unprecedented demand for vaccines, coupled with logistical challenges, has highlighted the essential role played by an efficient and resilient supply chain in ensuring equitable access and successful vaccination campaigns worldwide. By examining the Covid-19 vaccine supply chain, researchers, policymakers, and stakeholders gain valuable insights into the complexities of vaccine manufacture, distribution, and administration in times of pandemic. This knowledge helps identify potential bottlenecks, vulnerabilities, and areas for improvement, ultimately leading to more effective strategies and interventions in times of crisis. By better understanding the Covid-19 VSC, pharmaceutical companies, government and non-governmental agencies will be able to optimize production and distribution processes, ensuring a smooth flow of vaccines from manufacturers to final recipients in the event of future pandemics. Thus, enabling and simplifying contingency plans to be drawn up to deal with disruptions, such as measures to be implemented to reinforce the security and integrity of the supply chain. Lastly, an efficient vaccine supply chain is

essential for equitable vaccine distribution, particularly in times of pandemic which was one of the fears during the Covid-19 pandemic. By understanding the intricacies of the supply chain, policymakers can design strategies that better address the challenges of vaccine availability and accessibility. This ensures that vaccines reach every corner of the globe, contributing to the collective effort to achieve global herd immunity.

This project was based on the Covid-19 vaccine supply chain as of June 2021. Therefore, we suggest for future research to access the Covid-19 vaccine supply chain over time which will give insight into the evolution of a pandemic vaccine supply which has not been researched to this day.

Bibliography

Ahmed, F., Ahmed, N., Pissarides, C., & Stiglitz, J. (2022). Why Inequality Could Spread Covid-19. The Lancet, 395(10230), 809-810.

Asch, D. A., et al. (2021). Mobile Outreach for Covid-19 Vaccination. Journal of the American Medical Association, 325(13), 1251-1252.

Baden, L. R., El Sahly, H. M., Essink, B., Kotloff, K., Frey, S., Novak, R., … & Follmann, D. (2021). Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. New England Journal of Medicine, 384(5), 403-416.

Bollyky, T. J., & Bown, C. P. (2021). The Tragedy of Vaccine Nationalism: Only Cooperation Can End the Pandemic. Foreign Affairs.

Bollyky, T. J., et al. (2021). The Equitable Distribution of Covid-19 Vaccines. Journal of the American Medical Association, 325(10), 929-930.

Bollyky, T. J., Gostin, L. O., Hamburg, M. A., Piot, P., & Redlener, I. E. (2021). National Security and Global Health: Safeguarding Global Health in the 21st Century. Columbia University Press.

Brewer, N. T., Chapman, G. B., Rothman, A. J., Leask, J., & Kempe, A. (2017). Increasing Vaccination: Putting Psychological Science Into Action. Psychological Science in the Public Interest, 18(3), 149-207.

Canales, C., Ortiz, M., Rodríguez, P., & Lagos, R. (2021). A network approach to understanding Covid-19 vaccine distribution in Latin America and the Caribbean. PLoS ONE, 16(4), e0249883.

Contractor, F. J., Kumar, V., & Kundu, S. K. (2007). Nature of the relationship between international expansion and performance: The case of emerging market firms. Journal of World Business, 42(4), 401-417.

Delamater, P. L., et al. (2021). Vaccine Equity and Access: A Review of Interventions for Underserved Populations. Journal of Community Health, 46(4), 726-740.

Dembek, Z. F., Smith, C. B., Wagner, M., & Dausey, D. J. (2021). Vaccine nationalism and the role of international organizations in the COVID-19 pandemic response. Global Public Health, 16(5), 685-699.

Dunning, J. H. (1988). The eclectic paradigm of international production: A restatement and some possible extensions. Journal of International Business Studies, 19(1), 1-31.

Evenett, S. J., Hoekman, B., Rocha, N., & Ruta, M. (2021). The COVID-19 vaccine production club: Will value chains temper nationalism? *Policy Research Working Papers*. https://doi.org/10.1596/1813-9450-9565

Fidler, D. P. (2021). COVID-19: The Politics of Vaccination. Foreign Affairs

Forrester, J. W. (1961). Industrial Dynamics. MIT Press.

Ghebreyesus, T. A., Moeti, M., Vaz, R. G., Bhatnagar, N., & Lal, S. (2021). Covid-19 Vaccination: Pathway to Equitable Global Access. World Health Organization.

Goldratt, E. M., & Cox, J. (1984). "The Goal: A Process of Ongoing Improvement." North River Press.

Gostin, L. O. (2020). International Pandemic Response: Key Legal Issues. JAMA.

Gupta, S. (2020). COVID-19 Vaccine Supply Chains: Implications for Global Health Governance. BMJ Global Health.

Haji Ali Afzali, H., Alemi, H., Khodayari, Zarnaq, R., & Alavi, S. S. (2021). Network analysis of Covid-19 vaccine supply chain in Iran: a social network analysis. Journal of Pharmaceutical Policy and Practice, 14(1), 1-8.

Hopp, W. J., & Spearman, M. L. (2000). "Factory Physics: Foundations of Manufacturing Management." Irwin/McGraw-Hill.

Hotez, P. J., et al. (2021). Equity Must Accompany Covid-19 Vaccines in Global Health. PLoS Neglected Tropical Diseases, 15(1), e0008962.

Hu, T., Liao, S., Wu, T., Xue, H., & Zhao, Y. (2021). Network analysis of Covid-19 vaccine distribution in the United States. Journal of Medical Systems, 45(6), 1-7.

Karafillakis, E., et al. (2019). Vaccine Hesitancy among Healthcare Workers in Europe: A Systematic Review. Vaccine, 37(36), 4761-4768.

Karakaya, E., Aksu, S., & Ustundag, A. (2012). Supply Chain Risk Assessment: A Case Study in the Pharmaceutical Industry. International Journal of Production Research, 50(21), 6162-6172.

Krause, P. R., Fleming, T. R., Longini, I. M., & Peto, R. (2021). Covid-19 vaccine trials should seek worthwhile efficacy. The Lancet, 397(10284), 1607-1609.

Krause, P. R., Fleming, T. R., Longini, I. M., & Peto, R. (2021). Covid-19 vaccine trials should seek worthwhile efficacy. The Lancet, 397(10284), 1607-1609.

Leask, J., et al. (2018). Target the Fence-Sitters. Nature, 555(7696), 446-447.

Lee, H. L., & Billington, C. (1992). "Managing supply chain inventory: Pitfalls and opportunities." Sloan Management Review, 33(3), 65-73

MacDonald, N. E., SAGE Working Group on Vaccine Hesitancy. (2015). Vaccine Hesitancy: Definition, Scope and Determinants. Vaccine, 33(34), 4161-4164.

Mclay, L. A., & Santos, M. (2011). Quantitative Models for Managing Supply Chain Risks: A Review. European Journal of Operational Research, 214(2), 283-293.

Mehrotra, A., et al. (2021). Addressing the Equity Barrier in Covid-19 Vaccination. Journal of General Internal Medicine, 36(4), 1036-1039.

Montgomery, D. C., & Runger, G. C. (2014). Applied Statistics and Probability for Engineers. John Wiley & Sons.

Moon, S., et al. (2020). Will Ebola Change the Game? Ten Essential Reforms before the Next Pandemic. The Report of the Harvard-LSHTM Independent Panel on the Global Response to Ebola. The Lancet, 386(10009), 2204-2221.

Mullard, A. (2020, November 30). *How covid vaccines are being divvied up around the world*. Nature News. https://www.nature.com/articles/d41586-020-03370-6

Ngo, A. (2021). Global Vaccine Nationalism: A Threat to Collective Security. Harvard Kennedy School.

Notteboom, T., & Rodrigue, J. P. (2005). "Port regionalization: Towards a new phase in port development." Maritime Policy & Management, 32(3), 297-313

O'Leary, S. T., et al. (2021). Disparities in Vaccine Hesitancy: A Review of Multiple Measures, Causes, and Possible Interventions. Human Vaccines & Immunotherapeutics, 17(5), 1418-1428.

Oliver, S. E., Gargano, J. W., Marin, M., Wallace, M., Curran, K. G., Chamberland, M., … & Dooling, K. (2021). The Advisory Committee on Immunization Practices' Interim Recommendation for Use of Pfizer-BioNTech Covid-19 Vaccine in Adolescents Aged 12-15 Years—United States, May 2021. Morbidity and Mortality Weekly Report, 70(20), 749-752.

Osterholm, M. T., Kelley, N. S., Sommer, A., & Belongia, E. A. (2021). Efficacy and effectiveness of Covid-19 vaccines. New England Journal of Medicine, 385(23), 2338-2349.

Ozawa, S., Stack, M. L., Publicover, N. G., & Walker, D. G. (2017). The economics of vaccine supply chains in developing countries. Vaccine, 35(17), 2279-2285.

Phelan, A. L., & Oluyomi, A. (2021). Covid-19 Vaccine Equity in the United States: A Rapid Review. Journal of Racial and Ethnic Health Disparities, 1-10.

Phelan, A. L., et al. (2020). Vaccine Ethics and Equity: More Than Just Distribution. American Journal of Bioethics, 20(12), 7-9.

Phelan, A. L., et al. (2020). What SARS-CoV-2 Reveals About the Global Health Security Agenda. The Lancet.

Pisani-Ferry, J., et al. (2021). A Proposal for the COVID-19 Vaccine Supply Chain. Project Syndicate.

Polack, F. P., Thomas, S. J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., … & Pérez, J. L. (2020). Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. New England Journal of Medicine, 383(27), 2603-2615.

Riaz, M.M.A., Ahmad, U., Mohan, A. *et al.* Global impact of vaccine nationalism during COVID-19 pandemic. *Trop Med Health* **49**, 101 (2021). https://doi.org/10.1186/s41182-021-00394-0

Ribeiro, H. V., Rodrigues, F. A., Oliveira, M. J., & Zorzenon dos Santos, R. M. (2017). "Mapping accessibility in the Brazilian vaccination coverage network: A spatial and network analysis." PloS ONE, 12(6), e0179593

Rugman, A. M., & Verbeke, A. (2004). A perspective on regional and global strategies of multinational enterprises. Journal of International Business Studies, 35(1), 3-18.

Smith, J. C., & Sivaramakrishnan, K. (2020). Covid-19 and vaccine logistics: Challenges and opportunities. Journal of Global Health, 10(2).

Smith, K. F. (2021). Global vaccine alliances and the role of vaccine diplomacy in addressing global challenges. The Lancet, 397(10294), 1034-1042.

Smith, M. J. (2020). "The Politics of a Covid-19 Vaccine." The New England Journal of Medicine, 383(22), e129.

Sridhar, D., Gurdasani, D., & Borelli, G. (2020). Will vaccine nationalism hinder efforts to control the COVID-19 pandemic? The Lancet, 395(10236), 1339-1341.

Takian, A., Kazempour-Ardebili, S., & Zakeri, H. (2020). "Artificial intelligence for COVID-19 vaccine supply chain management." Journal of Medical Systems, 44(8), 135.

Thompson, M. (2021). Vaccine Nationalism: The Game of Strategic Ethics. Journal of Bioethical Inquiry.

Vazquez-Prokopec, G. M., Mercado, M. I. G., & Dumonteil, E. (2021). The Challenge of Vector-Borne Diseases in Underserved Regions: A Scoping Review. Frontiers in Public Health, 9, 613540.

Verma, S., Gokhale, M., & Bhatia, R. (2021). Global vaccine supply chains: Current status, challenges and opportunities. Vaccine, 39(7), 1016-1027.

Wang, J., Jing, R., Lai, X., Zhang, H., Lyu, Y., Knoll, M. D., & Fang, H. (2021). Distribution of Covid-19 vaccine in China: a social network analysis of distribution channels. Vaccine, 39(7), 1052-1056.

Wooldridge, J. M. (2015). Introductory Econometrics: A Modern Approach. Cengage Learning.

World Bank. (2021). Global Economic Prospects. World Bank Group

World Health Organization. (2020). Guidelines on the management of vaccine supply chains during the Covid-19 pandemic. Geneva: World Health Organization.

Wouters, O. J., Shadlen, K. C., Salcher-Konrad, M., Pollard, A. J., Larson, H. J., Teerawattananon, Y., … & Glassman, A. (2021). Challenges in ensuring global access to Covid-19 vaccines: production, affordability, allocation, and deployment. The Lancet, 397(10278), 1023-1034.

Wu, S., Li, T., Liu, Y., Wu, L., Zhang, Z., & Liu, W. (2021). Resilience assessment of Covid-19 vaccine supply chain in China: A network analysis. Vaccine, 39(22), 2943-2950.

Xiong, H., Chen, H., & Zhao, Y. (2021). A comprehensive review of the global vaccine development initiatives in response to the Covid-19 pandemic. Human Vaccines & Immunotherapeutics, 17(6), 1846-1854.

Yadav, P., & Lyngdoh, T. I. (2020). Covid-19 vaccine supply chain: Challenges and strategies. Indian Journal of Public Health, 64(Supplement), S152-S156.

Zhang, Y., Li, Y., & Lu, J. (2021). From emergency to sustainability: Building resilient vaccine supply chains against Covid-19 and future pandemics. Globalization and Health, 17(1), 1-10.

Appendix A - Econometric model 1.1 (Total Pfizer & Moderna - Europe)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_pfizer	-1,199,855,738.00 (-3,004,589,376.00, 604,877,899.00)	
reg_moderna	1,962,033,172.00 (-1,102,106,931.00, 5,026,173,274.00)	
number_cov_chain_pfizer	1,492,283,726.00 (-2,397,859,686.00, 5,382,427,137.00)	
number_cov_chain_moderna	-1,336,981,569.00 (-5,864,500,848.00, 3,190,537,709.00)	
number_formulation	-1,095,200,524.00 (-2,546,700,985.00, 356,299,937.00)	
number_lipids	253,718,096.00 (-3,815,996,179.00, 4,323,432,372.00)	
number_fill	3,846,213,330.00** (387,404,706.00, 7,305,021,954.00)	
number_top_1000	17,009,844.00 (-437,022,667.00, 471,042,356.00)	
lag(reg_pfizer, 1)		-1,367,418,454.00*** (-1,945,855,018.00, -788,981,889.00)
lag(reg_moderna, 1)		432,410,103.00 (-146,026,461.00, 1,010,846,667.00)
lag(number_cov_chain_pfizer, 1)		3,485,352,029.00*** (1,538,452,481.00, 5,432,251,576.00)
lag(number_cov_chain_moderna, 1)		3,086,511,307.00** (517,104,082.00, 5,655,918,533.00)
lag(number_formulation, 1)		-3,137,593,328.00** (-5,627,566,578.00, -647,620,078.00)
lag(number_lipids, 1)		-1,300,353,540.00 (-4,559,377,446.00, 1,958,670,365.00)
lag(number_fill, 1)		1,319,307,009.00 (-1,653,884,175.00, 4,292,498,193.00)
lag(number_top_1000, 1)		79,548,243.00 (-36,139,069.00, 195,235,556.00)
Constant	-0.0000 (-1,007,999,357.00, 1,007,999,357.00)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix B - Econometric model 1.2 (Total Pfizer & Moderna - North America)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_pfizer	34,014.00 (-82,108.33, 150,136.30)	
number_cov_chain_pfizer	5,521.92 (-169.64, 11,213.48)	
number_formulation	1,192.63 (-1,180.93, 3,566.18)	
lag(reg_pfizer, 1)		-39,854.79
lag(number_formulation, 1)		-621.79
Constant	-34,013.94 (-152,102.70, 84,074.78)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix C - Econometric model 1.3 (Total Pfizer & Moderna - South America)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_pfizer	581.23 (-4,150.71, 5,313.17)	
reg_moderna	3,016.93 (-2,372.18, 8,406.04)	
number_formulation	2,509.14 (-2,843.45, 7,861.72)	
number_top_1000	-129.41 (-3,242.65, 2,983.83)	
lag(reg_pfizer, 1)		1,500.03 (-1,909.38, 4,909.45)
lag(reg_moderna, 1)		654.67 (-2,736.75, 4,046.09)
Constant	43.14 (-5,262.94, 5,349.22)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix D - Econometric model 1.4 (Total Pfizer & Moderna - Asia)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_pfizer	-777.24 (-6,354.95, 4,800.46)	
reg_moderna	3,553.17 (-11,900.57, 19,006.90)	
number_cov_chain_moderna	-40,807.27 (-129,625.40, 48,010.92)	
number_formulation	-41,697.10 (-187,732.40, 104,338.20)	
number_fill	50,022.99 (-55,007.21, 155,053.20)	
number_top_1000	1,324.35 (-2,746.63, 5,395.34)	
lag(reg_pfizer, 1)		980.1*** (980.10, 980.10)
lag(reg_moderna, 1)		-628.75
lag(number_cov_chain_moderna, 1)		-23,457.08*** (-23,457.08, -23,457.08)
lag(number_fill, 1)		29,142.02*** (29,142.02, 29,142.02)
Constant	-2,648.71 (-12,633.72, 7,336.30)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix E - Econometric model 1.5 (Total Pfizer & Moderna - Africa)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_pfizer	0.00 (-14,647.87, 14,647.87)	
reg_moderna	3,736.77 (-8,948.65, 16,422.19)	
lag(reg_pfizer, 1)		7,473.54*** (7,473.54, 7,473.54)
lag(reg_moderna, 1)		-7,473.54*** (-7,473.54, -7,473.54)
Constant	-0.00 (-10,357.60, 10,357.60)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.

All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix F - Econometric model 2.1 (Moderna - Europe)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_moderna	405.80 (-299.99, 1,111.59)	
number_cov_chain_moderna	-338.37 (-807.46, 130.73)	
number_formulation	-93.14 (-338.96, 152.67)	
number_lipids	255.70* (-42.46, 553.86)	
number_fill	833.02* (-123.74, 1,789.78)	
number_top_1000	-2.60 (-120.81, 115.61)	
lag(reg_moderna, 1)		196.07 (-66.17, 458.31)
lag(number_cov_chain_moderna, 1)		144.33 (-949.60, 1,238.26)
lag(number_formulation, 1)		-733.47* (-1,543.26, 76.31)
lag(number_lipids, 1)		56.73 (-1,434.48, 1,547.94)
lag(number_fill, 1)		1,782.9** (404.52, 3,161.43)
Constant	-229.99 (-646.71, 186.73)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.
Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix G - Econometric model 2.2 (Moderna - North America)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
number_cov_chain_moderna	7,694.38 (-11,906.57, 27,295.33)	
number_formulation	326.25 (-343.24, 995.73)	
number_lipids	-6,586.40 (-32,775.26, 19,602.46)	
lag(number_cov_chain_moderna, 1)		-3,700.0*** (-3,700.08, -3,700.08)
lag(number_formulation, 1)		-113.23*** (-113.23, -113.23)
Constant	6,586.40 (-19,063.41, 32,236.21)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.
Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix H - Econometric model 2.3 (Moderna - South America)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_moderna	165.80 (-104.98, 436.59)	
number_formulation	1,303.83 (-237.37, 2,845.03)	
number_top_1000	0.00 (-157.22, 157.22)	
lag(reg_moderna, 1)		193.44 (-126.99, 513.86)
Constant	-0.00 (-157.22, 157.22)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.

All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix I - Econometric model 2.4 (Moderna - Asia)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_moderna	1,023.03 (-2,041.94, 4,088.01)	
number_cov_chain_moderna	-8,699.66 (-28,680.38, 11,281.06)	
number_formulation	-8,530.30 (-39,871.49, 22,810.89)	
number_fill	10,142.36 (-12,102.94, 32,387.66)	
number_top_1000	277.97 (-627.48, 1,183.42)	
lag(reg_moderna, 1)		313.41*** (313.41, 313.41)
lag(number_cov_chain_moderna, 1)		-7,853.13*** (-7,853.13, -7,853.13)
lag(number_fill, 1)		8,977.55*** (8,977.55, 8,977.55)
Constant	-848.37 (-3,772.91, 2,076.17)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.

All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix J - Econometric model 2.5 (Moderna - Africa)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_moderna	0.00 (0.00, 0.00)	
number_fill	0.00 (0.00, 0.00)	
lag(reg_moderna, 1)		0.00 (0.00, 0.00)
lag(number_fill, 1)		0.00 (0.00, 0.00)
Constant	0.00 (0.00, 0.00)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.

All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix K - Econometric model 3.1 (Pfizer - Europe)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_pfizer	252.60 (-1,315.97, 1,821.17)	
number_cov_chain_pfizer	1,985.39*** (585.26, 3,385.53)	
number_formulation	-1,145.02** (-2,217.79, -72.24)	
number_lipids	-360.58 (-1,420.29, 699.13)	
number_fill	2,989.81**(85.63, 5,893.99)	
number_top_1000	-5.69 (-411.42, 400.05)	
lag(reg_pfizer, 1)		-708.40***(-708.40, -708.40)
lag(number_cov_chain_pfizer, 1)		1,477.41**(184.93, 2,769.89)
lag(number_formulation, 1)		-1,413.34 (-3,284.54, 457.85)
lag(number_lipids, 1)		-1,646.18 (-5,350.17, 2,057.82)
lag(number_fill, 1)		2,434.45**(556.20, 4,312.69)
Constant	-0.00 (-995.35, 995.35)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix L - Econometric model 3.2 (Pfizer - North America)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_pfizer	21,135.44 (-47,973.29, 90,244.17)	
number_cov_chain_pfizer	3,117.90 (-207.13, 6,442.92)	
number_formulation	866.38 (-845.13, 2,577.89)	
lag(reg_pfizer, 1)		-20,753.77***(-20,753.77, -20,753.77)
lag(number_formulation, 1)		-508.55***(-508.55, -508.55)
Constant	-21,135.38 (-91,668.42, 49,397.66)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix M - Econometric model 3.3 (Pfizer - South America)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_pfizer	2,196.47 (-2,445.85, 6,838.79)	
number_formulation	737.37 (-3,485.53, 4,960.28)	
number_top_1000	-1,497.18 (-4,966.47, 1,972.11)	
lag(reg_pfizer, 1)		2,212.48 (-1,860.98, 6,285.94)
Constant	499.06 (-4,328.62, 5,326.74)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.

All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix N - Econometric model 3.4 (Pfizer - Asia)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_pfizer	812.91 (-806.15, 2,431.96)	
number_formulation	-3,244.10 (-46,261.20, 39,772.99)	
number_fill	25,156.41 (-14,371.26, 64,684.08)	
number_top_1000	-285.63 (-1,053.06, 481.80)	
lag(reg_pfizer, 1)		591.96 (-249.15, 1,433.07)
lag(number_fill, 1)		14,584.58 (-546.75, 29,715.90)
Constant	571.26 (-1,977.67, 3,120.20)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,
with control for potential serial correlation, contemporenous correlation and heteroskedasticity.
All the commands and algorithms are coded in R 3.5.3 using the plm package.

Appendix O - Econometric model 3.5 (Pfizer - Africa)

Robustness Regression Results

	Number of Vaccine doses administrated	
	Beck and Katz 2	Arellano and Bond 2
	(1)	(2)
reg_pfizer	2,491.18 (-7,274.06, 12,256.42)	
lag(reg_pfizer, 1)		3,736.77***(3,736.77, 3,736.77)
Constant	-0.00 (-8,456.95, 8,456.95)	
Time Fixed Effects	No	No

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Time-fixed effects estimations based on Beck and Katz,

with control for potential serial correlation, contemporenous correlation and heteroskedasticity.

All the commands and algorithms are coded in R 3.5.3 using the plm package.