Doctoral Programme in Economics

UNIVERSITÀ DEGLI STUDI FIRENZE

UNIVERSITÀ DI SIENA

UNIVERSITÀ DI PISA

Department of Economics and Statistics

University of Siena

Director: Prof. Ugo Pagano

Essays in Applied Economics: Disease Outbreaks and Gravity Model Approach to Bovines movement network in Italy

Mahdi Gholami

Supervisor: Prof. Tiziano Razzolini

Thesis submitted for the degree of Doctor of Philosophy in Economics

March 2017

In the Name of God

THIS THESIS IS DEDICATED

WITH RESPECT AND AFFECTION TO MY BELOVED PARENTS

Habib and Ziba

Acknowledgment

I would like to thank Professor Razzolini for his invaluable guide in the writing of this thesis. Not only has he very carefully read and commented this study; and even more importantly, since the beginning he was always available to listen, discuss, and inspire my research path. Actually, I could not have done this work without his great helps. The possibility to learn from Prof. Razzolini is the most important thing I take with me from my academic experience here in Siena. My immense appreciation to him goes far beyond what I can possibly say in these few lines.

I am indeed so grateful to Italian National Animal Identification and Registration Database instituted by Ministry of Health at CSN of Istituto G.Caporale, based in Teramo for delivering us the valuable dataset on nation-wide registered Bovine movements. In addition, I would like to thank The Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise "G. Caporale" for providing SIMAN[1] system, a national information system for the notification of animal diseases in Italy. Without them, this work could not be fulfilled. This work has also benefited from very useful comments and suggestions of Prof. Paolo Pin (Bocconi University).

There was not anybody more important to me in the pursuit of my doctoral degree than the members of my family. I would like to thank my parents, whose love and their prayers are with me in anything I pursue. Words cannot express how thankful I am to you, my mother and father, for all of the sacrifices that you have made on my behalf. Also, I would like to thank all of my friends who supported me in my career and encouraged me to struggle towards my goal. Finally, yet importantly, I would like to thank all people who contributed to making my

[1] Sistema Informativo Nazionale delle Malattie Animali

Ph.D. period a splendid and unforgettable life event in the wonderful city of Siena. Besides education, I found the great opportunity to visit and learn many amazing Italian cultural and historical heritages, too. I am deeply indebted to all of the people who have been involved in both my academic and non-academic adventures in the wonderful land of Italy.

Needless to say, that I am only responsible for all the errors and omissions of this work.

Table of Contents

Summary

Generally, the movement of cattle within any country including Italy is very essential to the economics of livestock industry. However, this transmission can also carry and spread the risk of contracting the disease (infectious diseases) by other cattle in various geographical areas. For example, this pattern of animal movements brought about an outbreak of foot-and-mouth infectious disease throughout the UK in 2001.[2] Also as Taylor et. al (2001) mentioned that these diseases can give rise to the productivity decline and even can lead to some threats to human health. Therefore, to reduce the risk and economic loss of this kind of contagious diseases, authorities should be able to manage and control them (Andesron, 2002). This control measures can consist of monitoring the animal trades, inspecting entry to and exit from premises, adopting some eradication programs (say, sending sick cattle to slaughterhouses), quarantining cattle, and etc.

To properly assess this kind of control measures, we need a comprehensive and detailed information and regulation about the cattle movement pattern. To address these issues, European Economic Community (EEC) devised some regulations and imposed them on its member states to adhere to these measures (EU traceability framework). Actually, these measures originate from the public health and food health concerns, which are related to animal health and the economic impacts of the outbreak of infectious diseases. The EEC issued Council Directive 92/102/EEC in 1992 (with latest modifications in 2013), which obliged member states to record the origin and destination point of each cattle. Also, each cattle should be tagged by an ear tag to be traceable[3]. Specifically, European Parliament and European council in 2000 tried to simplify and implement

[2] UK Department for Environment, Food and Rural Affairs (DEFRA), Origin of the UK Foot and Mouth Disease Epidemic 2001 [2002]. available at: http://agris.fao.org/agris-search/search.do?recordID=US201300078648

[3] Proposal for a regulation of the European parliament and of the council on Animal Health available at: http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52013PC0260

this process through a digital framework that allows countries to identify and register their bovines[4]. In Italy, this procedure was done by Italian National Animal Identification and Registration Database which proposes us a rich and valuable dataset, including the all required parameters, to perform and analyze some important determinants of Bovines movements and the effects of this kind of disease outbreaks on the pattern of bovines trade among different holdings.

In recognition of the importance of national and international bovines' trade, this thesis attempts to assess and investigate some relevant determinants of bovines movements among Italian holdings and Italian provinces.

This study consists of three chapters. In the first chapter, we introduced a structural gravity model of trade and then we linked it to the Italian bovine trade system. Then, we assessed two important determinants of any animal movements, i.e. feed prices and financial literacy rate of farmers. In addition, we tried to analyze the interaction of these determinants on the movements of bovines among Italian provinces. We found that feed (corn, in our case) price shocks and financial literacy rate of farmers could significantly affect on the pattern of bovines movement. Furthermore, our findings suggest that this two factors have a close relationship together and can offset each other's effects, in the sense that the enhancing financial literacy rate of farmers can somehow immune them to the unexpected price shocks and actually undermine the effects of unfavorable price shocks on their business.

In the second chapter, we used again the structural gravity model and employed it to investigate the risk of the outbreak of bovine diseases among Italian provinces. We found that the

[4] Regulation (EC) No 1760/2000 of the European Parliament and of the Council of 17 July 2000 establishing a system for the identification and registration of bovine animals and regarding the labelling of beef and beef products and repealing Council Regulation (EC) No 820/97. Available at: http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32000R1760

disease incidence rate has a significant positive effect of the movement of bovines from origin nodes to destination points. Also, we tried to merge our findings with the feed (corn) price effects and saw that these two factors act in a different direction. That is the more is the incidence rate, the effect of feed prices become less relevant to the movement of bovines among different provinces.

In chapter 3, we tried to have a more detailed view on the effects of disease outbreaks on the movements of bovines among farms and slaughterhouses. In general, we found that the disease status has a negative effect on bovines movement from farms to farms, and a positive influence on the movements of bovines from farms to slaughterhouses. In addition, we found that the ownership has a significant role in determining the pattern of trade among holdings, in the sense that the results are reliable only if two trading partners have different owner/keeper. Although, if the effects were driven by movements between farms with the same owner, it would be a rational decision by the owners in order to separate and protect healthy bovines from sick bovines. However, we found that in the case of positive disease tests, distance has a positive effect on the movement of bovines between various farms. This somehow shows us that some farms may act in such an opportunistic behaviors that can even lead to the spread of diseases among different regions.

Finally, we analyzed some network characteristics of the bovines movements to see the effect of network structure and its interaction with disease status on the pattern of movements. More specifically, we used the most commonly used feature of any network i.e. indegree, outdegree, and degree and found that in the case of diseases the farms tend to send more bovines to the slaughterhouses. Actually, the results on the interaction of indegree and disease tests show us that the farms that have received more (probably sick) bovines in the previous period are more likely to send their bovines to slaughterhouses that farms. This can be seen as a sign that shows us the movement network changes to avoid the spread of diseases among bovines in different farms. Also, we should note that the indegree analysis can help us tracing back the diseases to the nodes that are more exposed to receiving bovines and consequently more subject to contracting diseases.

On the other hand, we found the outdegree is also important only in case of the trade from farms to slaughterhouses. In the sense that the farms that have had a positive disease test records in a period, prefer to extend their slaughterhouse partners and send more bovines to them. So, this means that the market share of previous partners to decrease in the current period. Thie result again is consistent with the fact that holdings trade together in a manner to reduce the risk of an outbreak. Also, knowing the effects of outdegree can help us finding potentially dangerous nodes in the sense that these nodes can be a source of spreading diseases among different premises.

Chapter I

Bovines' movement in Italy: a gravity model approach

1. Introduction

Isaac Newton (1687) initially found out the Universal Gravitation law in physics, which states that attractive force between two physical objects has a direct proportion to the mass of those objects and a reverse relation with the distance between them. Henry Carey, the American sociologist, in his "Principles of Social Science (1858)" used this law to explain some of the social phenomena. Since then, social scientists have been using a version of Newton's Law of Gravitation to explain the movement of people, information, and goods between different locations. Because, in general, we can see some elements of size and distance in social science issues which are similar to the mass and distance of objects in physics.

Gravity model has been used a lot in recent decades in the field of international economics and especially econometric analysis of trade patterns and actually has achieved a substantial amount of success in explaining the trade phenomenon. This model has become an important workhorse in the literature of international economics and many scholars have been employed it to illustrate the interaction among entities including countries, peoples, etc. It has been applied in a series of many subjects containing international trade, migration, foreign direct investment (FDI) and so on and so forth. The basic idea of this model is that geographical and spatial patterns in the economic activities can be explained by two main factor. The first factor is the size or economic weight or can be called the economic mass of the units which for example in the case of countries, mostly represented by Gross Domestic product (GDP), and the second factor which can be somehow interpreted as an obstacle to trade is distance among economic units. These factors are going to be interpreted that the closer the units are, the more trade between them and the bigger the weight (or size or mass) of the entities are, the more trade we have between them.

Some scholars are also used gravity models of international trade to analyze economic and social issues such as the movement of peoples in terms of migration flows. Actually, the simplest versions of gravity models that have been used in the context of bilateral emigration/immigration movements are to consider the relative size of the origin and destination countries and the distance between them; however, there are some other factors that can influence migration patterns. For this reason, gravity models are expanded with other variables associated with different migration pull and push factors; for example, better economic opportunities and actually better future in the destination country (it can be for instance prospects for higher wages or lower unemployment rates, more secure conditions, and higher political freedom, and so on.

Although, some limitation regarding available data hindered the use of gravity model in the literature of migration movements. However, some improvement in accessing the data made it possible to assess this migration flows on the related economies. For example, the effect of emigration/immigration flows on the host/home economies is a highly discussed issue in political and social fields. In fact, this kind of movements importantly influences at the level of the internationalization of the related economies. A quite big academic literature since Gould (1994) and Rauch and Trinidade (2002) has been published and resulted in a positive and significant impact of immigration in promoting trade and investment between the origin and destination countries.

We have seen in recent times that the gravity model has been used in other sub-field of international economics such as international agricultural trade literature. The reason is that this model could have been showed its relevance in describing the variables that influence international agricultural trade. Accordingly, many empirical studies in areas of agricultural policy, economic size, trade potentials, economic integration, and other obstacles and stimulus of agricultural trade have been done by exploiting the gravity model.

International agricultural trade is generally aimed at exchanging agricultural goods between or among entities that might have some kind of agreement and understanding of each other and the purpose of this exchange is to improve their cost-benefit situations and in general their economic conditions. Of course, as United States Agency for International Development (USAID, 2010) pointed out the agricultural trade is important to the poor in developing countries because most of the world's poor live in rural areas where agriculture is the key source of income and consumption. Actually, Agricultural trade produces a source of growth and agricultural growth motivates growth in other sectors of the economy. Moreover, the supply costs of agricultural products/commodities are determined to some extent by natural conditions such as weather conditions, soil quality and etc., while some more important factors can affect in this issue, for instance per capita income, resources endowments and stocks, transportation costs, population of units, level of development (like financial literacy rate), and etc. Thus, identification and evaluation of the determinants of agricultural trade potentials in Italy such as animals, animal products, and etc., is considered an important research focus so as to contribute to the body of information available towards policy implications and strengthening of the Italian regional economies.

As we said in previous paragraphs, the analysis of the people movement in a migration context uses gravity models where the stocks of immigrants are considered as a representative number or proxy for the size of the immigrant flows and are explained as factors that decrease some barriers to trade between the origin and the destination countries. Also, theoretical development, as well as empirical advancement in the estimation methods of gravity model in recent years, made it possible to assess more determinant of this migration-trade patterns. Specifically, after the seminal work by Santos-Silva and Tenreyro (2006), Poisson maximum likelihood estimation has become almost the workhorse in the literature, although there is not a commonly accepted method among scholars. In fact, they offered a Pseudo-Poisson Maximum Likelihood (PPML) estimator which capture the problems arise by heteroscedasticity and zero-trade problem. One of our contributions is to use this methodology in explaining the bovine-trade patterns.

Another contribution is related to the focus on sub-national units of analysis that are provinces, which is not novel in itself, but rather in the way its implications for theory-consistent and applied policy making are discussed and made somewhat clear.

So far, we have mentioned that there have been many academic attempts to assess the determinants of transfer of people in forms of emigration, immigration and so on, but there have not been such studies about the movement of livestock in literature. One obstacle was that the lack of reliable data on this kind of bovine movement. Fortunately, and thanks to Italian ministry of agriculture we have access to a rich dataset containing the overall cattle movements through Italian provinces. We will present them in section 3.

There are so many exogeneous and endogenous factors that can influence the movement of livestock in a country. Here in this chapter, we are going to assess the effects of the two of this factors and their interaction on the transfer of bovines among selected Italian provinces. The external factor is feed (corn) price and the internal factor is the risk level of the farmers, which can be captured through financial literacy measures. We will explain them more in next parts.

According to the report done by "Global Agricultural Information Network" on the "Italian Livestock and Products Outlook 2012"[5], the Italian cattle farms have been facing hard times for the last few years due to increasing input costs, difficulties accessing credit, high energy and transportation costs, heavy bureaucratic regulatory affairs, and falling demand. In addition, biogas demand for corn, as well as the need for more land required to realize the EU nitrates directive, increases the competition for land specifically in the northern parts of Italy. Accordingly, to this kind of changes, the farms and feedlots should manage their costs to can survive during this situation.

As the abovementioned report states, one consequence of these changes can be in the stock of cattle herds in different locations and on the transfer of cattle heads among different regions. The reports show that during recent years, Italy's cattle heads decreased slightly due to high input costs and low farm-gate prices. We can see an overview of cattle production, supply, and demand in table 1.

Moreover, the report states that the feed prices have risen during marketing year 2010/11 while less demand brought about prices to drop during the first 6 months, and then it had a negative effect on profitability. As a result, farmers were compelled to decrease their supply to make up the decreasing demand. However, feed prices slightly declined over the summer and cattle prices started to increase again. In September, farmers began to increase their cattle herds due to the improved market situation. As a result, calf production is expected to increase slightly, while prices are expected to remain almost constant.

---

[5] Actually, this report includes assessment of commodity and trade issues made by USDA staff on the program of Global Agriculture Information Network.

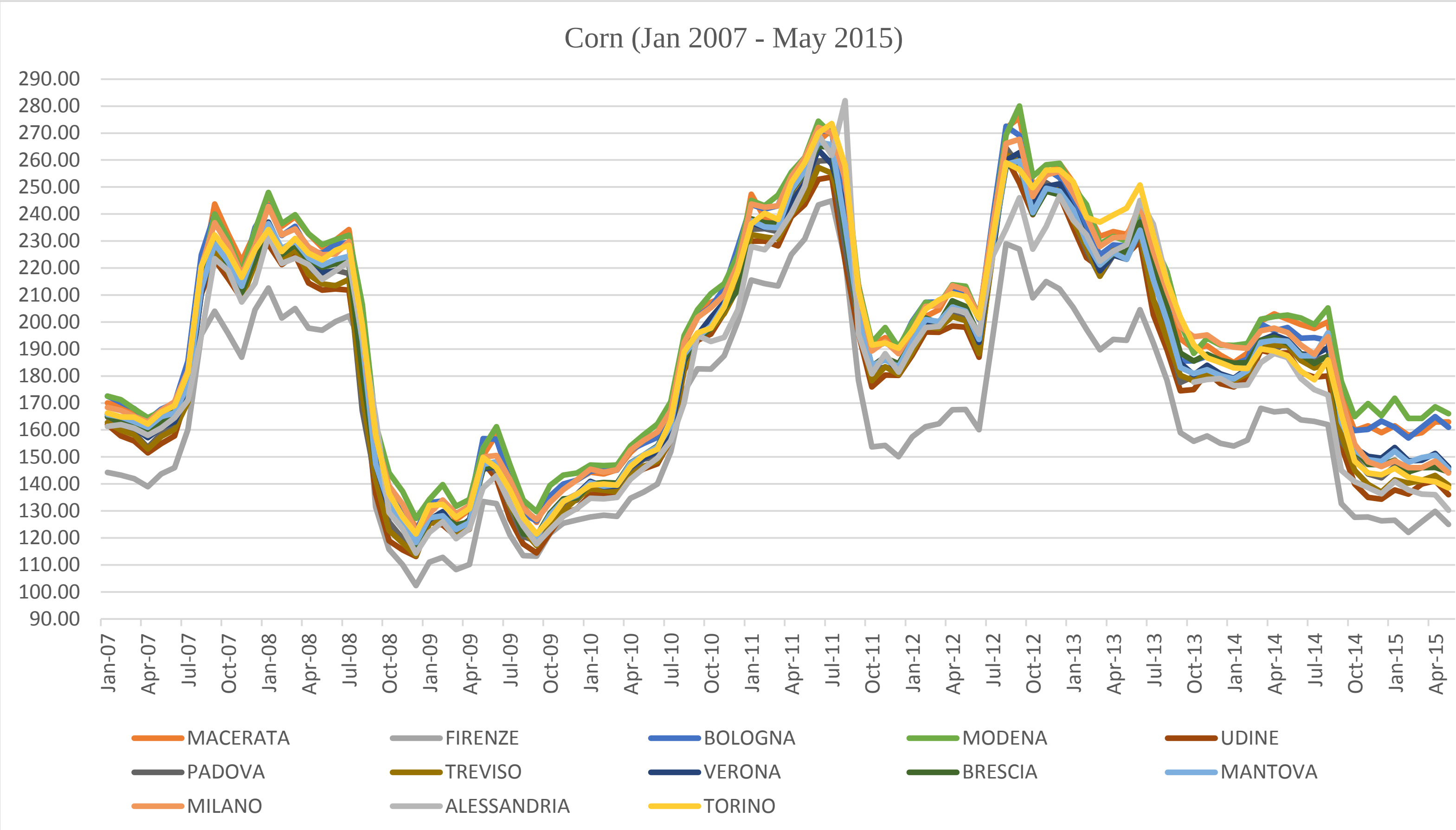
One of the input costs, which as we said is somewhat an external determinant factor, that affect on these changes in the stock and movement of cattle herds is feed price. For example, We can see the significant of the effect of feed price in the United States agricultural industry (Hebert, C.J., and B.C. Anderson., 2011). In fact, they show that the exogenous effect of corn (one of the feeds in agriculture) prices has a negative effect on the price of beef. Therefore, the farmers prefer to sell off their cattle earlier when prices of feed increase.

In the field of crops, there are two categories of hay that is used in Italian cattle industry. One is alfalfa hay and another is meadow stable hay. Alfalfa hay, wheat, and corn are several cattle feed in Italy. Among them, corn is considered the most common cattle feed in Italy. In figure 1, we can see the fluctuations of the corn price as an input cost which can be considered as a determinant in explaining the variations of the movement of cattle and livestock industry in recent years.

Table 1. Cattle Production, Supply, and Demand ('000 Head). (Source: Italian Livestock and Products Outlook report)

Cattle	2010 Marketing year begin: Jan 2010 Current post data	2011 Marketing year begin: Jan 2011 Current post data	2012 Marketing year begin: Jan 2012 Current post data
Total Cattle beginning stocks	6,447	6,198	6000
Dairy Cows beginning stocks	2,099	1,991	1,995
Beef cows beginning stocks	374	372	370
Production (calf crop)	2,350	2,269	2,582
Intra EU-27 imports	1,372	1,345	1,320
Extra EU-27 imports	0	0	0
Imports	1,372	1,345	1,320
Total Supply	**10,169**	**9,812**	**9,902**
Intra EU-27 exports	39	39	38
Extra EU-27 exports	1	4	4
Exports	40	42	42
Cow slaughter	524	520	525
Calf slaughter	932	885	895
Total slaughter	3,862	3,700	3,750
Loss	70	70	70
Ending inventories	6,198	6,000	6,040
Total distribution	**10,169**	**9,812**	**9,902**

Figure 1: Corn prices changes of 13 Italian provinces (Jan 2007 to May 2015)



The other factor that is somehow internal to the farm's decisions and affects the decision-making process of the farms is economic understanding of the farmers, which usually captures by financial literacy measure. Financial literacy is becoming more important as countries face with an increased access to financial markets and services. Removing some technological, geographical and generally, national barriers can result in accessing of more people to financial products. For example, Mobile phone systems, urbanization, the increasing microfinances, etc., made it possible for poor and rural areas to can easily access to a bank account from far distances. However, some scholars such as Jappelli (2010) and Lusardi & Mitchell (2011c;2014) found some evidence in

Europe and United States which shows some low level of financial literacy and lack of the knowledge of people from basic economic and financial concepts.

In fact, lack of financial literacy and enough economic understanding can result in a decline in individual welfare and accordingly, many countries have adopted some strategies to enhance the level of the financial literacy of their citizens[6]. Supporting financial involvement and preparing access without creating literacy programs may result in problems for new bank account holders who do not understand how to invest and budget their savings. Generally, we can say that the financial literacy can be considered as a public good that can be increased through foreign collaborations with financial service institutions, trainers and educators and in general, civil society.

According to International Handbook of Financial Literacy (2016), Financial literacy is described as "*the ability to deal effectively with money and financial matters—not just for professionals in the sector of investment and banking, but for every person responsible for managing his/her financial affairs in everyday life*". This can include basic numeracy, risk diversification, inflation, etc. In fact, someone who is financially literate can understand value changes of money over time, interest rate functions, mechanics of credit and debt, and how these various financial/economic tools can affect their decision process. Moreover, those who are not quite financially literate are not able to make knowledgeable decisions about their personal financial matters.

With respect to the farmers, they are facing many risks related to their business. So, understanding risks and importance to hedge the risks can be an influential factor in determining

[6] For further details, see OECD (2013a).

their business behaviors. In fact, a (commercial) farmer can be considered as a typical hedger. Because the market prices of the feeds (i.e. corn, wheat and etc.) may fluctuate through the time (this in itself can be due to some other factors like weather conditions, lack of the rain, changes in supply and demand of the feeds and so on) and in this respect, the farmers might decide that buying feeds is a good idea at one time and a possible loss in another time. So, as we see, the price changes can be a source of the risk for farmers and farmers may use various financial instruments to reduce or hedge their risks. One such instruments can be the forward contracts. *"Forward contracts are mutual agreements to deliver a certain amount of a commodity at a certain date for a specified price and each contract is unique to the buyer and seller"* (Shinde, 2016[7]). In our case study, the farmer can sell a number of forward contracts corresponding to the amount of corn that he/she expects to need and basically lock in the current price of the corn. Anyway, This kind of behaviors needs some level of financial literacy that national and local entities can provide them to the farmers.

Also, One other important source of risk for any farm and any agricultural economic sector is pests and diseases. Accordingly, governments can adopt some strategies to improve the economic and financial literacy of their farmers in order to decrease the possible losses in case of the spread of (infectious) diseases in their cattle and generally livestock industry. For example, some useful strategies could be investing in vaccination programs for their animals or some collaboration with other farms to fight against pest and diseases. In fact, financial literacy related programs can help and enable farms to monitor crucial factors in their business and decide which was is better in

[7] For more details, refer to "Basic of commodities market: The target of success" by Deepak Shinde (2016)

planning/saving for the unexpected events. Therefore, as we can guess that the financial literacy situation of the farmers can have a significant role in determining their trade decisions.

In this chapter, we are going to have a general view on the level of financial literacy of the households in some Italian provinces and examine the role of this measure in the decision-making process of farms. Actually, we measure the level and distribution of financial literacy among Italian provinces (most important ones based on the frequency of bovine trade among themselves) and investigate its effects on bovine movement behavior. Our findings show that there is a considerable difference in financial literacy rate of different Italian regions. This is compatible with the work of Fornero and Monticone (2001), where they studied the detailed distribution of financial literacy among the population and asserted a specific pattern in Italian financial literacy rate, which is a regional difference in financial literacy between residents of the north, center, and south.

Thus, we would like to analyze whether financial literacy matters in the decision-making process of people to transfer bovines among different holdings and how the interaction of this measure with feed prices can explain the pattern of trade among different holdings in different provinces. Our empirical analysis is based on the survey done by Bank of Italy in 2010 on Household income and wealth (SHIW[8]).

In sum, based on the above consideration, the specific objective of this chapter is to find out effective factors that are significant in the movement of bovines among Italian provinces and specifically what is the effect of feed price shocks on the pattern of this movements.

The rest of the chapter is organized as follows. In section 2, we will review the theoretical literature of the gravity models, and then in section 3, we will go over the empirical literature of

[8] For more details on the ingredients of financial literacy, refer to: https://www.bancaditalia.it/pubblicazioni/indagine-famiglie/bil-fam2010/index.html

gravity models. In this part, we explore on the several methods in the econometrics literature to estimate a gravity model and describe the pros and cons of each method. Then in section 4, we explain our adopted theoretical model, which is a structural gravity model of trade based on Anderson & van Wincoop (2003). In the fifth part, we present our dataset, material and the relevant method that we have adopted to estimate our model. Then, we have our results and so, we can discuss them. At the end, in section 6, we conclude the chapter.

2. Theoretical literature on gravity models

The economic model of gravity equation was first introduced by Tinbergen (1962) and then was expanded by Linnemann (1966) and as Anderson & van Wincoop (2003) pointed out that It has been accepted, both theoretically and empirically, a popular model among economic scholars. Anderson and van Wincoop (2003) also mentioned that the empirical gravity models have been very successful in economics in general. In fact, after several criticisms about some drawbacks of theoretical justifications, like lack of micro-foundation basis and etc., it has had restoration in recent couple of years with many new theoretical advancements and practical applications such as remarkable works of Anderson, (1979); Helpman and Krugman, (1985); Bergstrand, (1989), Deardoff, (1998), and Anderson and van Wincoop (2003).

After more than two decades of increasing efforts to provide models that justify theoretical applications for the empirical success of the gravity equation, it almost has been shown that the empirical results of the gravity equation are specified nicely on theoretical foundations. One important contribution in this regard was associated with the structural form of the gravity equation, which was developed by Anderson and van Wincoop (2003) and the implication of misspecification or omitted variable bias. These connect to the way that trade costs and firm heterogeneous behavior is included in the gravity equation. Actually, they included multilateral trade resistance trade cost into a micro-founded gravity equation. Accordingly, we also use this structural gravity model proposed by Anderson and van Wincoop (2003) as our theoretical model to incorporate relative feed price as another relative cost. We will describe this model in more detail in section 4.

A second important contribution of the structural gravity equation is associated with the firm heterogeneous behavior that was incorporated into the equation. This methodological was led by Melitz (2003) and Bernard et al. (2003) and relates to the presence and behavior of heterogeneous companies which are functioning in international markets.

Firm heterogeneity occurs when a minority of living firms in a country exports and not all of them (Bernard et al, 2003). Moreover, an amount of existing firms and not all firms export to all other countries in the rest of the world. In fact, they are only active for some specific countries and may select not to export some products to specific markets (because of the lack of means and so on.). The cause of this kind of heterogeneous behavior is because of the fixed costs are market specific and almost are higher for international trade than for domestic markets. Therefore, just most productive firms can overcome these costs, and firms' inability to sell its product to other countries may be due to the high cost included. As a result, the bilateral trade matrix will not be full and we will have zero cases in many cells. Of course, since we are going to assess the movement of cattle at a subnational level, so, we are not so concerned with this heterogeneity problem. Because as we see in our dataset, almost all provinces trade with each other in a different period of time. In fact, we have less than 30% zero in our dataset. Nevertheless, the zero trade values problem is an important issue and it should be addressed in the literature.

The widespread presence of zero bilateral trade flows has influential implication to model the gravity equation. One reason that some scholars such as Helpman, Melitz, and Rubinstein (2008) propose are that this zero values in itself may contain significant information about the countries, for example, why they are not tending to trade with each other, while we should be able to take into account this information to achieve efficient estimations.

Standard gravity models commonly ignore the issue of the predominance of zero bilateral trade flows and predict theory consistent with only positive bilateral trade flows. However, some theoretical gravity equation was derived by some scholars such as Helpman, Melitz, and Rubinstein (2008); Melitz and Ottaviano (2008); Chen and Novy (2011), etc., which take into account this issue and propose theoretical interpretations for it. For example, the ‘new new’ trade model of international trade with firm heterogeneity which was constructed by Metlitz (2003) is commonly considered for reaching to a theoretical basis for gravity equation.

Also, Helpman et al. (2008) discuss that “by disregarding countries that do not trade with each other, these studies give up important information contained in the data” (Helpman et al. 2008 p442), and so the symmetric relationship presumed by the standard gravity model biases the estimates and therefore it is inconsistent with the data. To resolve this problem (i.e. biased estimation), Helpman et al (2008) propose a theoretical gravity equation that includes both firm heterogeneity and positive asymmetry, and it can provide predictions and consistent results in the case of both positive and zero trade values between unit pairs. In addition, the more recent attempts in the literature of the gravity models are the improvement of theoretical gravity equation that provides a theoretical basis when we are facing with both observed and unobserved zero trade flows (Kareem et al. 2016). This is associated with an advent of theoretically consistent estimation techniques and those that would take into account of zero trade values.

In sum, what we are looking for in this study is that to use the most relevant theoretical gravity model in an agricultural movement context (like, people movement or livestock movement) which is structural gravity model which considers the bilateral trade resistance term as well as other trade cost such as distance, and lead to consistent and efficient estimation methods. We will explain this model in section 4.

3. Empirical literature on gravity models

As we know, the gravity model is becoming very popular in recent decades to explain trade patterns. This is because of its empirical success in describing such trade relations. Nevertheless, in spite of its popularity, there is not much common census about the appropriate econometric estimation method that brings about consistent estimates, when for example we have so many zeros in our dependent variable or logarithmic transformation of the equation. In fact, in this section, we briefly explore some of these methods and pros and cons of each. Then, we will see that based on our available dataset and the regular assumptions of gravity models, the method proposed by Santos Silva and Teneyro (2006) is more appropriate for our work.

As we said until now, the gravity model is very accepted among scholars to analyze trade relations. Firstly, this is because of the strict theoretical base that this model has and second which is somehow more important is the successfulness of this model in explaining foreign trade patterns. Nevertheless, there has still been some doubts about the appropriate specification of the equation and the appropriate econometric estimation method(s) that would provide consistent estimates when we are facing the problem, mentioned above, like zero trade values etc. In this part, we are going to delve into the various estimation methods used to address these concerns. Our detailed focus is on the advantages and disadvantages of each method. Then, at the end of this section, by reviewing the methods used by scholars in the empirical studies, we will find the Poisson method as a workhorse in estimating structural gravity equations.

3.1 Some discussion in the gravity model estimation methods

Early empirical studies like Anderson (1979), Bergstrand (1985), Deardoff (1998), etc., were based on cross-sectional data to estimate a gravity equation. Actually, cross-sectional analysis was accepted as an economic framework for the gravity models. The common technique for estimating this cross-sectional analysis is the ordinary least square (OLS) method, but as some scholars such as Carrere (2006) have mentioned, the traditional cross-sectional approach is influenced by serious misspecification problems and so, previous estimates cannot be trusted. This is due to the fact that the traditional cross-sectional gravity model normally contains constant variables -such as same language, distance, border effects, cultural dummies and etc., - but the misspecification problem arises when it does not take into account time-invariant unobservable effects of unit-pairs. Thus, this unobservable unit –which can be national or subnational units- specific time-invariant variables embed in the error term. Since these unobserved variables probably are correlated with observed ones, the OLS estimation method can be lead to inconsistent estimators, which invalidates one of the classical assumptions of OLS. Moreover, since OLS estimation method does not take into account heterogeneity in individual units, this can result in biased estimators. This is due to the fact that the estimated parameters may change depending on the units considered. Thus, without considering these unit specific unobservable effects, our estimating method would be problematic and can be lead to biased estimators.

However, in recent decades and years, there has been an increasing wave of studies which have used panel data in their gravity models and so they exploited panel econometric techniques for estimation of respective models[9]. One advantage of panel data specification is that it gives rise to more degree of freedom and so as a result, we will have more precise estimates. Also, as Baltagi (2008) has pointed out, panel framework enables us to model changes in variables through time and space and therefore we can control for omitted variable bias in form of unobserved heterogeneity. Moreover, in panel data analysis, we can include unit-specific (country-pairs) effects such as time dummies, to model time-invariant unobserved trade effects and then as a result of this process we are not facing with the consistency problem which we cited above.

In dealing with panel data, the two common methods which are used in analyzing the data are the fixed effects and random effect estimation techniques, where the choice between the two depends on the presumptions. In fixed effect method, we enforce time-invariant effects for each unit that are possibly correlated with the independent variable. Actually, we assume that the unobserved heterogeneity is correlated with the error term. When in random effect method, we assume that the unobserved heterogeneity is completely exogenous and so, there is not any correlation between the individual effects and the independent variables. Also, the most common method to analyze panel data have been OLS method and if the relation between the dependent variable and independent variables does not change over time, we can have more exact estimators with pooled OLS estimation method.

Early studies on gravity model used ordinary least square methods to estimate the gravity equation and to do so, an accepted practice was to log-linearize the multiplicative form and the estimate the model. The validity of this logarithmic transformation is based on the homoscedasticity assumption (as the log and the error term should be statistically independent of

[9] See also, Baltagi et. al. (2003), Rose and van Wincoop (2001), Egger and Pfaffermayr (2004), Melitz (2007), and many others.

other repressors). Nevertheless, in recent years, Santos Silva and Tenreyro in their influential works (mainly in Silva & Teneyro (2006)) have pointed out some drawbacks of this practice. Actually, they said due to the fact that most trade data are basically subject to heteroscedasticity and also the abundance of zero values, log-linearizing the gravity equation and then applying OLS estimation method could lead to some problems.

As we said in the previous paragraph, one of the problems of logarithmic transformation is because of the trade data are mostly subject to heteroscedasticity. Nevertheless, Santos Silva and Tenreyro (2006) noted that OLS estimation method that is usually used after log-linearizing of gravity equation is not suitable. The reason behind their argument was that the expected values of the log-linearized error term will depend on the covariates of the regression and therefore OLS will not be consistent anymore[10] because with transforming the gravity equation into a logarithmic form, the properties of error term will change. To state this matter differently, as we know the OLS technique will present consistent estimates provided to the error term (ε_{ij}) of the log linear transformation ($\ln\varepsilon_{ij}$) is a linear function of independent variables (i.e. if $E\{\ln(\varepsilon_{ijt}|x_{ij})\}=0$), that is the homoscedasticity assumption. But, the logarithmic specification produces estimates of $E(\ln\varepsilon_{ij})$ and not $\ln E(\varepsilon_{ij})$, And we now that this two are not equal to each other, And also, we know from Jensen's equality that "the expected value of the logarithm of a random variable is different from the logarithm of its expected value", and here we have $\ln E(\varepsilon_{ijt}|x_{ij})=0$; $E(\ln(\varepsilon_{ijt}|x_{ij}))\neq 0$.[11]

[10] Santos Silva and Teneyro (2006) also argued that even if all observations of the dependent variables were strictly positive, OLS estimators are still inconsistent.

[11] Jensen's inequality is called after Johan Jensen, the Danish mathematician who in 1906 discovered that: the secant line of all convex function (i.e., the means of the convex function) lies above graph of the function (i.e., the convex function of the weighted means) at every point. The reverse is true for a concave function. His inequality has come into many contexts and an example in this case is the arithmetic mean inequality. Thus, in simple terms, his inequality states that the convex transformation of a mean is less or equal to the mean after a convex transformation. Similarly, the concave transformation of a mean is greater or equal to the mean after a concave transformation. Thereafter, Economists have adopted his intuition to show that the logarithm transformation of an equation generates the expected

Therefore, because of Jensen's inequality, we cannot say that the error terms in both specifications are equal and actually, the error terms in the log-linear transformed model of the gravity equation are not statistically independent of the repressors but are rather heteroskedastic which result in inconsistent estimates of the elasticity coefficients.

Santos Silva and Tenreyro (2006) start with this Jensen's inequality and mention that we are essentially faced with the heteroscedasticity problem in the case of log-linear transformation and so by using OLS estimation method, we will have inefficient and biased estimators. They point out that although the scholars were aware of this Jensen's inequality and the concavity of the logarithmic functions which can lead to biased estimators in case of OLS method, however, this influential flaw has been ignored in trade studies. They verified their reasoning as they figured out some evidence of heteroscedasticity and inconsistency in the log-linear transformation of a gravity equation. This problem can lead to inefficient and inconsistent estimates of the least square estimation method.

Another problem of logarithmic transformation which is somehow more troublesome is the presence of zero value trade data and so the choice of a suitable estimation method. Although in the Newtonian gravity theory, which is the origin of a gravity model, we can have very little gravitational force but not zero ones, nevertheless, this is not the case in trade context literature. Because there may have some zero trade value in bilateral trade data, which if we apply the log-linearize transformation method, we will face with both theoretical and methodological problems. For example, Frankel (1997) cited that some source of these zero values which can be as a result

---

value (mean) of the logarithmic transformation of the dependent variable $E(\ln Y_{ij})$ and not the logarithm of the mean of the dependent variable $\ln E(Y_{ij})$; and $E(\ln Y_{ij}) \neq \ln(E(Y_{ij}))$

of no-trade between countries or even we will have zero by rounding errors when trade between two country does not exceed a minimum bound and so they amount of trade can be rounded down to zeros. Also, we should note that this problem would be intensified, if a number of zeros in the dataset were exceeded the normal bound which is usually 50% of the data.

As we have said until now, the common practice in the literature to estimate a gravity equation is to first log-linearize the equation and the use the linear regression methods. Nevertheless, in the case of excessive zero trade value in trade matrix and specifically when this amount of zeros exceeds more than 50% and these zeros include pertinent information, this transformation of the dependent variable is not appropriate anymore because as we know from the mathematics the logarithm of zero is indeterminate or not feasible.

However, among the various methods, two methods have been considered as a common methodology in the economic literature. First, the truncation and second, the censoring methods. These methods are applied in the economic literature to overcome the problem of zero value trade data, and then using the linear methods to estimates the model.in the first one, i.e. truncation method, the zero value trade data totally delete from the trade matrix while in the second method, the zero values replaced by an infinitesimal arbitrary small positive number.

But as some scholars such as Linders and Groot (2006), Burger et al. (2009), Gomez-Herrera (2013), and etc., have shown that since these methods are arbitrary and they don't have any strong foundation (theoretically and empirically), so the result would not be so reliable and actually we will have somehow inconsistent estimates. Moreover, for example, Flowerdew and Aitkin (1982) found in their study that the results change when their constant replaced by another number (in censoring method). Actually, they changed the substituted constant between 0.01 and 1 and found that the regression coefficient changes, as well. Therefore, this shows that the results in censoring methods are highly sensitive to different chosen numbers that may twist our results, as well. In addition, Eichengreen and Irwin (1998) figured out that if we delete the zero values, we might lose some important information especially when these zero trade values are not randomly distributed and contain relevant information. So, this can produce the biased estimators. Another problem with deleting these zero values is sample selection bias (Heckman (1979) and Helpman et. al., (2008)). Gomez-Herrera (2013) mentioned that the information loss due to deleting zeros trade values, decrease the efficiency and generate biased estimates or as Xiong and Beghin (2013) have shown, this kind of deleting zero observations in truncation methods can be problematic for describing an economic interpretation of the model and actually we cannot say anything for future trade relations. Similarly, these concerns about the truncation and censoring methods have come into the literature by others scholars, as well. For example, Linder and Groot (2006) emphasized that the zero value trade flows may contain significant information about the trade relations and we cannot delete them in advance because since these zero values are not distributed randomly, our results would be biased. They argue that if for example the independent variables such as GDP, distance, etc., lead to less or even no trade, by deleting these zero values, we actually will have a sample selection bias and, accordingly we are facing with a downward bias in OLS method.

Therefore, by considering the problems of linear regression methods related to the logarithmic transformation of gravity equation and the presence of zero trade values, there has been much attention in recent times on more suitable estimation methods to can overcome these two issues in the literature of gravity model. Some of the main methods that are employed in this regards are the Nonlinear Least square (NLS) method, Tobit and Probit models, and Poisson method. All these methods have been employed for dealing with the problems of the log-normal specification and zero value trade data.

Most early works done by scholars have employed the Tobit model to consider the zero trade flows problem. For example, Anderson and Marcouiller (2002) and Rose (2004) used the Tobit model to take into account the zero values in the estimation. They said that these zeros can be results of actual trade flows or can be due to the measurement errors. Anyway, they applied the Tobit estimator to fit their data when data are only observable over some period. It is used in cases of measurement errors (for example, rounding up) or when actual outcomes do not express the desired outcomes. Actually, The Tobit censoring method includes rounding (censoring) part of the observation to zero or rounding up the zero trade flows below some positive number. Nevertheless, some other scholars like Linder and Groot (2006) argued that in the gravity model, the zero trade values cannot be censored at zero as the desired trade cannot be negative in the gravity equation; and this case can only happen if the GDP of one or country pair is equal to zero which we don't have anything like this in real world. They also discuss this kind of censoring by the help of the UN COMTRADE dataset for trade flows. They say that the censoring at a positive value is not correct because most of zero values are the result of economic matters and not from rounding up. Thus, the model is not suitable for considering zero trade values. Moreover, Martin and Pham (2008) compare the two model of truncation in OLS and censored Tobit, and conclude that the both model give biased results, though the latter propose more biased estimators. As a result, they recommend that the threshold Tobit model which was proposed by Eaton and Tamura (1994) is more appropriate and gives the lowest bias estimators. Nevertheless, Santos Silva and Tenreyro in their paper in 2011 figured out that this threshold Tobit model is also subject to large biases and even this bias increases with sample size, as well. Therefore, it can lead to an inconsistent estimator.

Until now, we showed that most estimation method for estimating gravity model have been subjected to some difficulties. Thus, scholars tried to find other techniques for dealing with such problems of log-linearizing and zero trade value. Santos Silva and Teneyro (2006) did one important and influential work where they propose to use the Poisson specification for estimating a gravity model and we can say, albeit some criticism, this method has become such a "role model" in the literature of gravity model. Actually, Santos Silva and Tenreyro (2006; 2011) used the Poisson Pseudo Maximum Likelihood (PPML) method to take into account the problems caused by zero value trade data and logarithmic transformation of the gravity equation. For example, Santos Silva and Tenreyro (2011) pointed out the truncated OLS and censored OLS methods give rise to inconsistent and biased estimates in case of zero value trade data and logarithmic transformation of gravity equation, and furthermore, if the sample size increase, this problem will intensify, as well. One advantage of this method that avoids the problem caused by using OLS under logarithmic transformation of gravity equation is that it uses the dependent variable itself instead of the logarithm of it. In fact, the technique estimates the values of a dependent variable in levels instead of taking its logarithm.

In general, Santos Silva and Teneyro discuss almost three reasons to show that their approach is more suitable than other techniques. Firstly, this Poisson estimation technique takes into account of observed heterogeneity. Secondly, by using the fixed effect Poisson pseudo maximum likelihood method, we can deal with zero value trade observations due to its multiplicative form and the least but not last, since this approach produces the estimates of trade flows and not the logarithm of them, we will not have under-prediction for large trade flows. Also, Santos Silva and Tenreyro in their 2006 persuasive work show that using PPML, we don't need to employ the logarithmic transformation of the dependent variable. In addition, by a simulation procedure, they show that the PPML estimator would be the best performing estimator that is consistent and gives rise to lowest bias among the other estimators. As a result, they claim that their method can be a “new workhorse” in estimating the typical constant elasticity models, such as gravity model (like the structural gravity model that we are going to use in our study).

Of course, albeit the advantages of the Poisson estimation method, there has been some challenges in the literature regarding this technique that although they are mostly have been responded by Santos Silva and Teneyro, but it worthwhile to mention some of them here. For instance, Burger et al. (2009) recognize some restriction of Poisson maximum likelihood method. Their point focused on the sensitivity of the Poisson model to the problem of overdispersion in the dependent variable and also vulnerability when we have too much zero trade values. They also said that on other restriction of Poisson model is that the model does not take into account unobserved heterogeneity and it just consider observed ones.[12] They argue that in this case, we will have an over-dispersion in the dependent variable and as Burger, et al. (2009) and Turkson (2010) have said, this overdispersion can lead to consistent estimates but not efficient estimates of trade flows.

Nevertheless, Santos Silva and Teneyro in their 2011 paper, responded to this criticism and showed that the Poisson Pseudo Maximum likelihood method leads to consistent and in general proper results even in the case of existence of over-dispersion (which means conditional variance is not equal to conditional mean) in the dependent variable and so, the criticism of over-dispersion which was brought by Burger, et al. (2009) is not such valid. They also argued that the prevalence of zero value trade flows does not influence the performance of our estimates. Moreover, some

[12] An important condition of the PPML is the equi-dispersion assumption in the dependent variable. It means that the conditional variance is equal to the conditional mean.

other scholars such as Soren and Bruemmer (2012) also showed that the PPML generate proper results even in the case of overdispersion.

All in all, we can see nonlinear estimation methods perform much better for estimation of gravity models and among them the Poisson Pseudo Maximum likelihood (PPML) method which initially proposed by Santos Silva and Teneyro (2006) gives more consistent and efficient estimates in case of logarithmic transformation and excess zero trade data. Also, as Silva and Teneyro have shown in their papers, this technique generally performs well for constant elasticity models and, since our theoretical model would be a structural gravity model which is derived from a Constant elasticity (CES) utility function, so this method is appropriate for our work. In addition, we do not have much zero value in our dataset, so it is not a serious concern for us and we can use this technique to estimate our model.

4. A structural Gravity model for Bovine movement

In this section, we develop a gravity model of Italian Bovine trade and uses it to investigate the relation and impact of feed price shocks on this movement. This kind of gravity models somehow differs from the "naive[13]" gravity models. In fact, they use a clear theoretical model of consumer decision maker to develop an econometric model and then it leads to gravity equation with multilateral resistance term included in it. This kind of models has brought about remarkable methodological progress in recent years. Also, these models have employed by many epidemiologists to model and assess the spread of infectious diseases among populations. In this

[13] Tinbergen (1962) says that the gravity model which derives from the physics similarity, has not a micro-founded base and it's actually empirically driven. Head and Mayer (2014), in their review paper, called this equation as "naive".

part, we are going to fit a structural gravity model to a panel dataset of Italian cattle movement to assess the effect of feed price on the movement. The units of trade in our case are provinces.

Anderson & van Wincoop (1979) introduced the theory of structural gravity models in the first place, where they augmented traditional model by a multilateral resistance term and from that time, there have been many advancements to international trade. In fact, this Model is the first structural gravity trade model that clearly takes into account the general equilibrium effects. It means that they considered relative trade costs- in addition to absolute trade costs- into the model.

Anderson and van Wincoop (2003) use this theoretical foundation and focus on the economics distance of trading partners with respect to a weighted average of economic distance to all other trading partners (which are "provinces" in our case). Then, they define this theoretically appropriate average trade barrier as "multilateral resistance" (Anderson and van Wincoop, 2003, p. 170).

Anyway, one necessary condition in this model is that the trading partners ("provinces" in our case) should have identical and homothetic preferences. Because it shows that, the provinces have the same demand function and use up the same quotient of their income on the commodities from another province. Homothetic preferences mean that the amount of demand for a good depends on the relative prices of that good and not on the income level. This property is fairly captured in a constant elasticity of substitution (CES) utility function. The utility function is given by:

$$U = \left(\sum_{i} c_{ij}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \qquad (1)$$

Where c_{ij} is the consumption of province j in goods from province i, and σ is the elasticity of substitution. Consumers in province j maximize this utility function subject to the budget constraint. This budget constraint shows the total expenditure of province j which has been resulted from the buying of goods from province i. So:

$$\sum_i c_{ij} p_{ij} = Y_j \quad or \quad \sum_i c_{ij} t_{ij} p_i = Y_j \qquad (2)$$

In the above equation, p_{ij} is the price of province i cattle for province j consumers (which are farms in our case) and it can be substituted by $t_{ij}p_i$, because the prices differ between provinces due to trade costs and one of our goals is to recognize these costs. So, we assume p_i is price in province i and t_{ij}>1 stands for the trade costs factor between two provinces, i.e origin node i and destination node j. Also, $c_{ij}t_{ij}p_i$ is the nominal value of movement from province i to j which Anderson and van Wincoop name it as X_{ij}. Therefore, summation of $c_{ij}t_{ij}p_i$ over all goods imported from other provinces is equal to the total expenditure of province j. In addition, since the provinces can only consume what they earn, Y_j can be interpreted as province j's income.14

$$market\ clearance: Y_i = \sum_j X_{ij}$$

Also, by assuming iceberg trade costs, these trade costs are proportional to the value of traded goods and not described by the volume of trade. It means that in our case these trade costs are equivalent to the feed price, which is proportional to cattle price.

Now, it is time to derive the gravity model. In the first step, and by using utility function and budget constraint, province j's demand for cattle originating from province I should be determined. In the next step, we should aggregate this demand over all receiving provinces from sending

[14] We assume that we have a one-sector economy and therefore, a province's total income Y_j is equal to its total expenses Ej.

province i. Hence, to get equilibrium trade and aggregate trade flows, this demand function is replaced in the market clearing condition.

To do the first stage we should maximize the CES utility function subject to the budget constraint. The lagrangian is given by:

$$L = \left(\sum_i {c_{ij}}^{\sigma-1/\sigma}\right)^{\sigma/\sigma-1} + \lambda\left(Y_j - \sum_i c_{ij}t_{ij}p_i\right)$$

Then, we should differentiate this lagrangian with respect to c_{ij} because our goal is to maximize utility (i.e. consumption c_{ij}). Therefore, the first order conditions are:

$$\frac{\partial L}{\partial c_{ij}} = \frac{\sigma}{\sigma-1}\left(\sum_i {c_{ij}}^{\sigma-1/\sigma}\right)^{\frac{\sigma}{\sigma-1}-1} \frac{\sigma-1}{\sigma} {c_{ij}}^{\frac{\sigma-1}{\sigma}-1} - \lambda t_{ij}p_i = 0$$

Then,

$$\lambda t_{ij}p_i = \left(\sum_i {c_{ij}}^{\sigma-1/\sigma}\right)^{1/\sigma-1} {c_{ij}}^{-1/\sigma} \qquad (3)$$

Now, we can multiply both sides by c_{ij} and then aggregate over all i. Therefore, we have:

$$\lambda \sum_i c_{ij}t_{ij}p_i = \left(\sum_i {c_{ij}}^{\sigma-1/\sigma}\right)^{1/\sigma-1} \sum_i {c_{ij}}^{\sigma-1/\sigma}$$

Now we can substitute λ from equation (3) and the summation of left-hand side from equation (2) and then obtain $t_{ij}p_i$ (=p_{ij}).

$$\left(\sum_i {c_{ij}}^{\sigma-1/\sigma}\right)^{1/\sigma-1} {c_{ij}}^{-1/\sigma}\left(t_{ij}p_i\right)^{-1} Y_j = \left(\sum_i {c_{ij}}^{\sigma-1/\sigma}\right)^{1/\sigma-1} \sum_i {c_{ij}}^{\sigma-1/\sigma}$$

The first terms on both sides, cancel out each other, and then, we have:

$$t_{ij}p_i = \frac{{c_{ij}}^{-1/\sigma}Y_j}{\sum_i {c_{ij}}^{\sigma-1/\sigma}} \tag{4}$$

Anderson and van Wincoop (2003) define a consumer price index of j, given by:

$$P_j = \left(\sum_i \left(p_i t_{ij}\right)^{1-\sigma}\right)^{1/1-\sigma} \tag{5}$$

By multiplying both sides of equation (4) with $t_{ij}p_i$ and using the price index of province j in equation (5) and then, some mathematical manipulations, we obtain:

$$\left(t_{ij}p_i\right)^{1-\sigma} = \frac{t_{ij}p_i c_{ij} {Y_j}^{-\sigma}}{{Y_j}^{1-\sigma}} {P_j}^{1-\sigma}$$

Now, we can substitute $t_{ij}p_i c_{ij}$ with X_{ij} and solve it for X_{ij}. Therefore, we obtain the demand function of province j for province i's cattle:

$$X_{ij} = \left(\frac{p_i t_{ij}}{P_j}\right)^{1-\sigma} Y_j \tag{6}$$

The general equilibrium structure of models requires market-clearing condition, which implies:

$$Y_i = \sum_j \left(\frac{p_i t_{ij}}{P_j}\right)^{1-\sigma} Y_j$$

Now, we can derive the equilibrium market price p_i:

$$p_i{}^{1-\sigma} = \frac{Y_i}{\sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma}} \times \frac{1}{Y_j} \tag{7}$$

Now, we can substitute this market price into equation (6) to obtain the demand function:[15]

$$X_{ij} = \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} \frac{Y_j Y_i}{Y_w} \left(\sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} \frac{Y_j}{Y_w}\right)^{-1} \tag{8}$$

Also, to simplify the equation, Anderson and van Wincoop define the last term within bracket as Π_i as:

$$\Pi_i = \left(\sum_j \left({t_{ij}}/{P_j}\right)^{1-\sigma} \frac{Y_j}{Y_w}\right)^{\frac{1}{1-\sigma}} \tag{9}$$

Then, if we substitute Π_i in equation above into demand function (8) and doing some mathematical manipulations, we will have the final structural gravity model:

$$X_{ij} = \frac{Y_i Y_j}{Y_w} \left(\frac{t_{ij}}{P_j \Pi_i}\right)^{1-\sigma} \tag{10}$$

In addition, by substituting the market equilibrium price in equation (7) into the CES price index in equation (5) and rearranging it, we obtain:

$$P_j = \left(\sum_i \left(\frac{t_{ij}}{\Pi_i}\right)^{1-\sigma} \frac{Y_j}{Y_w}\right)^{\frac{1}{1-\sigma}} \tag{11}$$

[15] Anderson and van Wincoop (2003) define world income by $Y_w = \sum_j Y_i$.

Anderson and van Wincoop found that under a system of symmetric transportation cost (i.e. t_{ij}=t_{ji}), the solution to the equation of (9) and (11) is $\Pi_i = P_i$ with:

$$P_j^{1-\sigma} = \sum_i \left(\frac{t_{ij}}{P_i}\right)^{1-\sigma} \frac{Y_i}{Y_w} \qquad (12)$$

Following this, the gravity equation can be expressed as bellow:

$$X_{ij} = \frac{Y_i Y_j}{Y_w} \left(\frac{t_{ij}}{P_i P_j}\right)^{1-\sigma} \qquad (13)$$

Anderson and van Wincoop (2003) refer to the price indices in the last term of equation (13) as "multilateral resistance" variables which depend on bilateral resistance t_{ij}.

Now, the logarithmic transformation of structural gravity equation is as follows:

$$\log X_{ij} = \log Y_i + \log Y_j + (1-\sigma)\log t_{ij} + (\sigma-1)\log P_i + (\sigma-1)\log P_j + K$$

4.1 Model Specification

As we see, in both traditional and structural gravity equation, the trade barriers such as geographical distance, language, border etc., exist but the main difference of this equation with the traditional gravity equation is in the price indices terms. So, we adopted the Anderson and van Wincoop (2003) equation as our preferred model. In our preferred model, the trade barrier is captured through a geographical distance of each province and the term of bilateral resistance term can be interpreted as the difference in the price of feeds, which is proportional to bovine price.

Also, the value of bovine movements among provinces is proportional to the number of heads which transport from one province to another one. Therefore, the number of inter-province heads and stock number of heads in each province can be a suitable proxy for the economic value of the bovines.

Altogether, we can have our gravity equation for cattle movement between provinces:

$$\log interprovince\_heads_{ijt} = \alpha_1 \log heads\_orig_{it} + \alpha_2 \log heads\_dest_{jt} + \alpha_3 \log distance\_prov_{ij} + \alpha_4 \log relprice_{ijt} + \varepsilon_{ijt} + Constant$$

Where i and j are subscripts for provinces and t denotes the time period. $interprovince\_heads_{ijt}$ is our dependent variable which is the number of heads which transfer from province i to province j in time period t. $heads\_orig_{it}$ stansd for the stock of heads in origin province i at timeperiod t. Similarly, $heads\_dest_{jt}$ is the number of heads in destination province j at time priod t. Also, $distance\_prov_{ij}$ denotes the geographical distance between two province i and j which has been computed by the longitudinal and latitudinal coordinates of the provinces[16]. $relprice_{ijt}$ that according to structural gravity model of Anderson and van Wincoop (2003) can be considered as our bilateral resistance term which shows the relative feed price (corn, in our case) of province i over province j at time period t. Finally, ε_{ijt} is the disturbance term.

Now, it is time to estimate our gravity model. There are so many methods for estimation of a gravity model. Each of them has some advantages and disadvantages which in the next part, we try to cover them and explain which method is the best for estimation of our model.

[16] Data on geographical coordinates of the Italian provinces derived from: en.comuni-italian.it/province.htm

5. Material and method

Data on bovine trade movements were obtained from the Italian National Bovine database, which is administered by the Italian National Animal Identification and Registration Database. The database contains the movement of the all Italian population of Bovines between animal premises, providing an extensive image of the locations that bovines have been kept and moved within the country. Additional information was provided for the animal holdings, including the type of holdings (i.e. animal husbandry or farm (AL17), Genetic Material Center (CG18), Collection Center (CR19), Fair/Market (FM20), Pasture (PA21), Point Park (PS22), Animal House (ST23), Stable (SS24) slaughterhouse (SM25),), and their geographic coordinates and also the geographical locations of the municipality where the holdings were located. Each movement case shows the unique identification for the animal, the codes of the holdings of origin and destination nodes, and the date of the movement. Such tracking system allows us to easily follow the path of each bovine and to make the equivalent overall network, and then we can minimize the problems with respect to data accuracy that are found in other tracking systems that do not provide both origin and destination of the movements (Volkova et al. 2010, Christley et al. 2005). Also, we have identification code for the owners and keepers of each node.

In addition, we have the data on the price of corn for some regions mostly northern part of Italy. We have collected this corn prices data for 10 provinces that have more trade among

[17] ALLEVAMENTO
[18] CENTRO MATERIALE GENETICO
[19] STABILIMENTO DI MACELLAZIONE
[20] FIERA/MERCATO
[21] PASCOLO
[22] PUNTO DI SOSTA
[23] STABULARIO
[24] STALLA DI SOSTA
[25] STABILIMENTO DI MACELLAZIONE

themselves and have more data available. These provinces include Alessandria, Bologna, Brescia, Milan, Mantova, Modena, Padova, Turin, Treviso, and Verona. Also, we excluded years 2006, 2014, and 2015, because we do not have a rich dataset of those years. Therefore, our dataset contains 2007 to 2013 and it covers the bilateral transfer of bovines among the 10 provinces from 2007 to 2013.

In addition, data for financial literacy measure is derived from the survey on Household income and Wealth done by Bank of Italy (2010[26]). Actually, Bank of Italy, through the SHIW, gathers detailed data on household consumption, income, and wealth for a representative sample of the Italian population (Banca d'Italia et al., 2010).

They have actually interviewed 7951 households, the households were drawn from the registry office records of 387 municipalities; they comprised 19836 persons, including 13074 income recipients. Then, the required questionnaire have been filled out by the head of the households and financial literacy rates derived from the answers to these questions. The selection of these three questions was based on the work done by Lusardi and Mitchell (2011) for the US Health and Retirement Study (HRS), and they actually capture the understanding of risk. The relevant questions for computing financial literacy are as following:

1- *Which of the following types of mortgage do you think would allow you from the very start to fix the maximum amount and number of installments to be paid before the debt is extinguished?*

   - Floating-rate mortgage 1

   - Fixed-rate mortgage 2

   - Floating-rate mortgage with fixed installments..3

   - Don't know..4

[26] This is the most recent year that financial literacy measure computed by the bank of Italy.

- No answer 5

2- *Imagine leaving 1,000 euros in a current account that pays 1% interest and has no charges. Imagine that inflation is running at 2%. Do you think that if you withdraw the money in a year's time you will be able to buy the same amount of goods as if you spent the 1,000 euros today?*

- Yes 1

- No, I will be able to buy less...2

- No, I will be able to buy more...3

- Don't know...4

- No answer 5

3- *Which of the following investment strategies do you think entails the greatest risk of losing your capital?*

- Investing in the shares of a single company 1

- Investing in the shares of more than one company 2

- Don't know...3

- No answer 4

In addition, the survey provides us the share of correct answers to each question for Italian provinces. Then, to achieve a measure of financial literacy, we mean the share of all correct answers on three questions.

To obtain relevant variables, here we collapsed dataset based on monthly time period and then expanded them to fill out the missing observations with zero values. The reason for breaking down data on monthly base is that as we saw in Figure 1, there are rather significant changes in the price of corn in each month. Then, we merged dataset of movements with data on feed prices. In table 2, we can see our descriptive statistics.

Table 2: summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Interprovince_heads[27]	6,762	5.500438	13.77986	0	242
heads_prov_orig[28]	6,764	241.7155	110.1758	25	834
heads_prov_dest[29]	6,764	477.4766	360.8629	3	2975
Relativeprice[30]	5,881	1.001678	0.039957	0.809927	1.212017
distance_prov (km)[31]	5,887	130.2133	70.24709	32.91862	361.4067
Financial literacy rate of origin nodes[32]	6,764	0.4685699	0.0449434	0.412504	0.5275121

As we discussed in section 3, there is not a commonly accepted method in the literature for estimation of a gravity model. However, Poisson-pseudo maximum likelihood method has become the "workhorse" for estimating structural gravity models. Santos Silva and Teneyro (2006) recommended that the Poisson pseudo-maximum likelihood (PPML) estimator introduced by Gourieroux et al. (1984) possesses all the features that a method needs to become a well-known workhorse for the estimation of gravity equations and, more generally, constant elasticity models, in the trade literature. They mention that in the case of zero value trade flows and logarithmic transformation of the gravity model, ordinary least square (OLS) method and other linear methods give inconsistent estimator. Also, Kareem et al, (2016) summarize three main reasons about the appropriateness of this Poisson estimation method:

1. *"The Poisson model takes into account the observed heterogeneity*

[27] Volume of the heads which have been transferred between two provinces
[28] Stock number of heads in origin node
[29] Stock number of heads in destination node
[30] Relative feed price of origin node over destination node
[31] Distance between two province measured in kilometer
[32] Financial literacy rate of origin provinces

2. *The fixed effects PPML estimation technique gives a natural way to deal with zero valued trade flows because of its multiplicative form.*
3. *The method also avoids the under-prediction of large trade volumes and flows by generating estimates of trade flows and not the log of the trade flows."*

In sum, Santos Silva and Teneyro conclude that the Poisson estimator -in comparison with other linear and nonlinear methods- is the best performing estimator in facing with zero value trade data and logarithmic transformation problem, and gives consistent estimator with the lowest bias.

Also, since we do not have so many zeros in our dataset (about 40% of our dependent variable is zero, so we are not facing with excessive zero value trade which makes a model problematic), so we don't have any major problem using this method (because as we said before, in this method, we estimate the equation with dependent variable in levels rather than its logarithmic version). Actually, the main reason to prefer Poisson Maximum Likelihood method is not the zeros but the heteroscedasticity of trade data. So, even without zeros, PPML is generally preferable.

5.1 Clustered standard errors

Santos & Tenreyro (2006) developed a Poisson pseudo maximum likelihood technique to overcome the problem of convergence of Poisson estimators. They have written a ppml command for stata that bypasses most of these problems. In addition, we cannot use the log-linear model with robust/clustered standard errors because using robust/clustered standard errors will only affect the estimated standard errors, but will have no effect at all on the estimates of the parameters. Therefore, the log-linear model will generally be invalid with or without the robust/clustered standard errors. PPML delivers estimates of the parameters that are consistent under very general conditions; of course, robust/clustered standard errors should also be used with PPML. In addition, ppml command in Stata does not have the option to include province-pair fixed effects; if we really want to include them we can consider Timothy Simcoe's xtpqml Stata command, this command runs Stata's pre-packaged fixed effects Poisson estimator and then computes the robust standard errors suggested by Wooldridge (1999). Another approach which is equivalent with xtpqml and includes province-pairs fixed effect is to use xtpoisson stata command with option "exposure(province-pairs)".

5.2 Estimation results and discussion

As we said in previous sections, we can estimate our gravity equation by the standard method of Poisson Pseudo Maximum Likelihood (PPML) with fixed effects. These methods can be considered if not the best but one of the best estimators for estimating trade data. Because it can overcome the problem of heteroscedasticity and excessive zero trade data. Fortunately, we do not have that much of zeros in our dependent variable and so, the problem of excessive zero value trade does not occur in our model. We have presented the estimated results of our model in table 3.

We remind that the dependent variable is "interprovince_heads" which is the number of bovines that have been traded between two provinces in a time period.

As we see, the first variable, Stock of origin province (log), is one of our mass variables which shows number of heads in origin province in time period t, and the second variable, logdestjt, is our another mass variable which shows the number of heads in destination province, have positive signs and are highly significant which verify our expected results from gravity model. They show that the more is the stock of each province, the more is the probability of transferring the bovines between respective provinces. Actually, it shows that the bigger province in terms of the endowments of the bovines is more likely to trade with each other.

Also, our third variable is the relative price of corn in origin and destination points. This is indeed, the value of corn price in origin province over corn price in destination province. As we see, the effect of this variable is positive and highly significant. Moreover, in table 5, we will see that distance has a negative and (slightly) significant effect which is reasonable and it shows that if two provinces are at a greater distance, the trade costs are higher and so the movement of bovines is smaller.

Table 3[33]. Effects of relative feed prices on the movement of bovines

VARIABLES	(1) Dependent Variable: interprovince_heads	(2) Dependent Variable: interprovince_heads
Stock of origin province (log)	0.607*** (0.0181)	0.598*** (0.0181)
Stock of destination province (log)	0.213*** (0.0128)	0.215*** (0.0128)
Relative feed price (log)	1.977*** (0.303)	-22.90*** (2.962)
Relative feed price*Distance (both in logs)	-	5.490*** (0.651)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: gravity equation results based on fixed-effect (province-pairs) Poisson maximum likelihood estimation method. Standard errors are also clustered by origin-destination nodes.

[33] We also controlled the specifications with time dummies (year and month). The respective results are reported in the tables, as well.

Table 3 continued.

VARIABLES	(1) Dependent Variable: interprovince_heads	(2) Dependent Variable: interprovince_heads
2008.year	-0.0749***	-0.0822***
	(0.0209)	(0.0209)
2009.year	-0.481***	-0.478***
	(0.0227)	(0.0227)
2010.year	-0.0750***	-0.0780***
	(0.0213)	(0.0213)
2011.year	-0.00786	-0.0133
	(0.0207)	(0.0207)
2012.year	-0.0893***	-0.0892***
	(0.0217)	(0.0217)
2013.year	0.148***	0.140***
	(0.0210)	(0.0210)
2.month	0.0145	0.0182
	(0.0287)	(0.0287)
3.month	0.214***	0.219***
	(0.0269)	(0.0269)
4.month	0.0127	0.0203
	(0.0290)	(0.0290)
5.month	-0.0717**	-0.0671**
	(0.0287)	(0.0287)
6.month	-0.0217	-0.0144
	(0.0295)	(0.0295)
7.month	0.209***	0.212***
	(0.0278)	(0.0278)
8.month	0.116***	0.118***
	(0.0290)	(0.0290)
9.month	0.0598**	0.0567**
	(0.0284)	(0.0285)
10.month	-0.141***	-0.132***
	(0.0298)	(0.0298)
11.month	0.0564*	0.0654**
	(0.0290)	(0.0290)
12.month	0.0594**	0.0630**
	(0.0296)	(0.0296)
Log likelihood	-32412.825	-32377.021
Observations	5,877	5,877
Number of origdest_uniquelink	85	85

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Specifically, the dependent variable is a count variable, and as we know that Poisson regression models the log of the expected count as a function of the explanatory variables. This means that we can interpret the Poisson regression coefficients as follows: for a one unit change in the independent variables, the difference in the logs of expected counts is expected to change by the respective regression coefficient, provided to the other independent variables are remained constant. Moreover, since here our independent variables are logged transformations, they can be interpreted as elasticities of the regression. So, for example, table 3 presents that by increasing 10% in relative feed price of origin over destination node, the movements increase by about 19.77% (main effect). Also, we can see the approximate increase in the number of movements due to the 1% increase in relative feed price by calculating the average marginal effects[34]:

Table 4: average marginal effects of feed price

	dydx	Std. Err.	z	P>|z|	[95% Conf. Interval]
Relative feed Price	3.030578	0.3270382	9.27	0.000	2.389595 - 3.671561

Note: the AME is computed based on column (2) of table 3

As we can see in table 4, the AME (Average Marginal Effect) estimation shows that 1% increase in feed price can lead to about 3 additional movements.

Furthermore, we can also compare the importance of distance effects as a proxy for trade cost and price effects as a proxy for resistance term. To do so, we added the interaction term to the equation and estimate the model again. The effect of the interaction term (relative price and distance) is positive and significant. This positive value for the effect of the interaction term would imply that the higher the distance, the greater (more positive) the effect of feed price on movements

[34] For calculating marginal effects, it just need to run command "margins, dydx(logrelativeprice)" in stata.

was. On the other hand, since the distance effect is negative and interaction is positive, it means that the effect of corn price fluctuations are dominant over the distance effects and so they are more important than the cost of transportation. It means that the more the prices in origin locations are and the far distance between two locations is, the farmers are more likely to send their bovines.

On the other side, as we explained in previous parts, we can see the effect of the level of economic understanding of farmers on their decision making process, by looking at the measure of financial literacy which is computed based on a survey done by the bank of Italy and some relevant questions about financial perception of households. The results are shown in table 5.

Table 5: Effects of financial literacy rate of origin nodes on the movement of bovines

VARIABLES	(1) Dependent Variable: interprovince_heads (RE)	(2) Dependent Variable: interprovince_heads (FE)
Stock of origin province (log)	0.607***	0.608***
	(0.0181)	(0.0181)
Stock of destination province (log)	0.214***	0.213***
	(0.0128)	(0.0128)
Distance_km (log)	-0.436*	-
	(0.227)	
Relative feed price (log)	1.902***	2.222***
	(0.303)	(0.311)
Financial literacy rate (log)	3.613***	-
	(1.214)	
Relative feed price* financial literacy rate (both in logs)	-	-4.845***
		(3.356)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: gravity equation results based on Poisson maximum likelihood estimation method. Columns (1) and (2) are estimated by random effects (RE) and fixed effects (FE province-pairs), respectively. Standard errors are clustered by origin-destination nodes (province-pairs unique links)

Table 5 continued.

VARIABLES	(1) Dependent Variable: interprovince_heads	(2) Dependent Variable: interprovince_heads
2008.year	-0.0728***	-0.0762***
	(0.0209)	(0.0209)
2009.year	-0.480***	-0.478***
	(0.0227)	(0.0227)
2010.year	-0.0748***	-0.0696***
	(0.0213)	(0.0214)
2011.year	-0.00767	-0.00487
	(0.0207)	(0.0207)
2012.year	-0.0892***	-0.0851***
	(0.0217)	(0.0217)
2013.year	0.149***	0.152***
	(0.0210)	(0.0210)
2.month	0.0140	0.0140
	(0.0287)	(0.0287)
3.month	0.214***	0.214***
	(0.0269)	(0.0270)
4.month	0.0126	0.00964
	(0.0290)	(0.0290)
5.month	-0.0721**	-0.0751***
	(0.0287)	(0.0287)
6.month	-0.0214	-0.0240
	(0.0295)	(0.0295)
7.month	0.210***	0.208***
	(0.0278)	(0.0278)
8.month	0.117***	0.117***
	(0.0290)	(0.0290)
9.month	0.0602**	0.0613**
	(0.0284)	(0.0284)
10.month	-0.138***	-0.142***
	(0.0297)	(0.0298)
11.month	0.0567*	0.0554*
	(0.0290)	(0.0290)
12.month	0.0594**	0.0597**
	(0.0296)	(0.0296)
Constant interprovince_heads	-1.497	-
	(1.543)	
Constant lnalpha	0.179	-
	(0.137)	
Log likelihood	-	-32407.004
Observations	5,879	5,877
Number of origdest_uniquelink	87	85

Standard errors in parentheses - *** p<0.01, ** p<0.05, * p<0.1

As we can see in column (1) of table 5, financial literacy rate has a positive main (pure) effect on the movement of bovine and it shows that the more is the level of financial literacy of a province; the farmers in that province tend to send their bovines to other places. Actually, the more the farmers are financially literate; they will trade more to each other. Specifically, for example, 10% increase in financial literacy, it increases the movements by about 36%. Also, we can see the approximate change in the number of inter-province movements by looking at the average marginal effects:

Table 6: average marginal effect of financial literacy rate

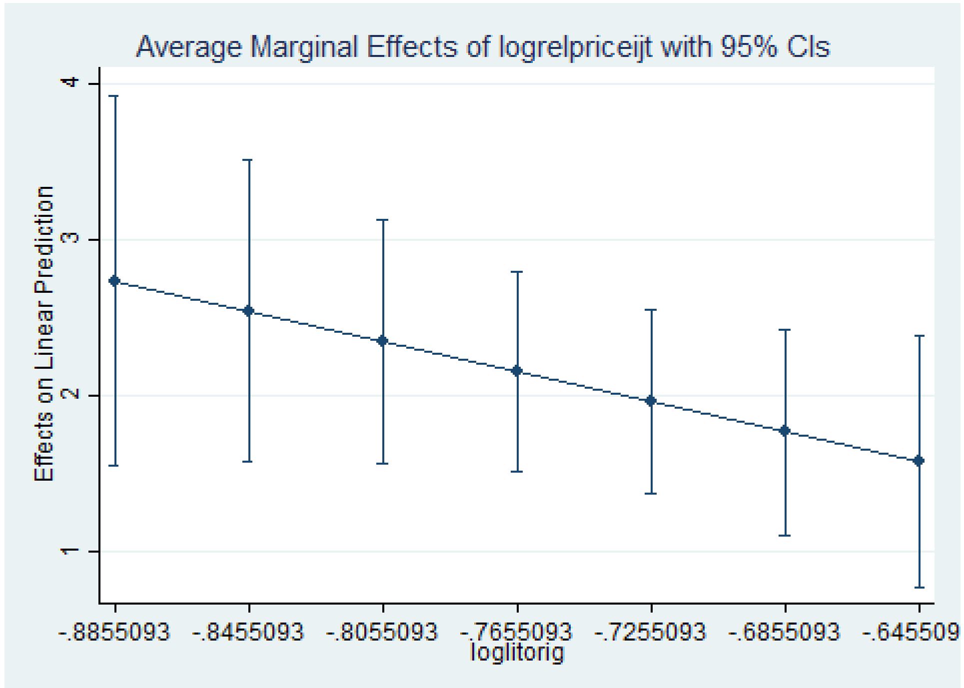
	dydx	Std. Err.	z	P>|z|	[95% Conf. Interval]
literacy rate	1.389291	0.7717803	1.80	0.072	-.123371 - 2.901952

Note: the AME is computed based on column (2) of table 5

Therefore, as we can see in table 6, the AME (Average Marginal Effect) result shows that 1% increase in financial literacy rate can lead to about 2 additional movements.

On the other hand, column 2 of table 5 presents that the interaction of financial literacy and feed price has a negative and significant effect on the bovine movements among provinces. A negative interaction coefficient means that the effect of the combined action of two independent variables is less than the sum of the individual effects. So, the connection between one of the independent variables (IV) and the dependent variable decreases if the other regressor increases. This effect can be better seen by looking at the average marginal effects of relative price in different financial literacy rate (margins graph):

Figure 2: Predictive Margins



As we can see from the margins graph, the slope of the graph is negative and it shows that by increasing the financial literacy rate of origin nodes the marginal effect of feed price decreases. Specifically, y-axis shows the effect of feed price on the movement of bovines at different values of financial literacy rate (log). For example when the logarithm of literacy rate in at its average value (i.e. x-axis is -0.7626), the coefficient of relative price is 1.902 as we see in column (1). On the other word, it shows that the more is financial literacy of provinces, the less is the effect of feed prices on the movement of bovines between different provinces. In fact, a financially literate farmer can take into account more parameter in its decision-making and so for example, he can hedge its future trade and investments and so on.

One point that we also had in mind was that to control the results for the wealth of the provinces, as well. However, since we don't have the GDP of each province, we were thinking to use the unemployement rate of provinces as a proxy for this issue and then control the regressions for the unemployment of origin and destination provinces. We found that the general results remain unchanged. The resulting regressions are reported in the appendix.

5.3 Robustness Check

In this subsection, we are going to explore some tests for the particular pattern of heteroscedasticity assumed by PPML estimator. Although Poisson Maximum Likelihood estimators are consistent even in the case of misspecification of variance function, nevertheless, the scholars can use the following tests to see if a different Poisson estimator would be more suitable and to make a decision whether the use of a nonparametric estimator of the variance is guaranteed or not. Actually, we are going to test the adequacy of the model by using some robustness checks. This means that we should figure out that if the patterns of heteroscedasticity assumed by the Poisson estimator is acceptable or not.

In general, given the estimator that we consider in our study, we have two type of heteroscedastic. First, Constant Coefficient of Variance (CCV) and second, Constant Variance to Mean Ratio (CVMR) type of heteroscedasticity[35]. On one side, we can see CCV heteroscedasticity pattern in the log-linear model, because this model has log-normal errors with constant variance (homoscedasticity assumption). It means that the in CCV, the standard deviation is proportional with its mean. On the other side, in PPML model which is a nonlinear model, we have CVMR type of heteroscedasticity which assumes that variance is proportional with its mean.

As Santos Silva and Teneyro (2006) and later Martinez-Zarzoso (2013) have said, we can test the performance of the models based on MaMu test which Santos Silva and Teneyro name it as a

[35] Refer to Head and Mayer (2014)

Park-type test. Actually, Manning and Mullahy (2001) showed that if the general form of the variance and conditional mean would be:

$$Var(Y|X) = K\ Exp(Y|X)^{\lambda} \qquad (14)$$

the choice of the appropriate PML estimator can be based on a Park-type regression (Park, 1966). So, we can use this test to find out that whether we can accept the pattern of heteroscedasticity which has been assumed by a Poisson estimator or not.

As we see in the general form of the relation of variance and mean, the parameter of the interest is λ that is a non-negative and finite value which determine the pattern of heteroscedasticity of a model. Specifically, as we said the heteroscedasticity pattern in the case of Poisson Pseudo Maximum likelihood is based on a constant variance to mean ratio (CVMR). So it corresponds with λ=1. Thus, as Manning and Mullahy proposed in their 2001 paper, given the equation (14), we can use their MaMu test to analyze the error term and investigate the heteroscedasticity pattern of the estimators. Actually, as Head and Mayer (2014) states, this is a credible method to discern between the log-normal data generating process and the CVMR data generating process.

To do the test, Manning and Mullahy (2001) state that the parameter λ can be determined by the auxiliary equation:

$$ln\left(Y - \hat{Y}\right)^2 = lnK + \lambda ln\hat{Y} + v_i \qquad (15)$$

Which directly derive from the general equation of variance and conditional mean. We should test the null hypothesis of $H_0: \lambda = 1$ against the alternative $H_1: \lambda \neq 1$. The hypothesis will be accepted if the appropriate confidence interval for parameter λ includes 1. The result of the Park test for both of our estimations are shown in tables below.

Table 7: Park test for feed price estimation

	Coefficient	95% Conf. Interval
lnyhat	.8241159	.577514 - 1.070718

Note: Park test for column (2) of table 3

Table 8: Park test for financial literacy rate estimation

	Coefficient	95% Conf. Interval
lnyhat	1.057093	.8757982 - 1.238388

Note: Park test for column (2) of table 5

As we see in tables 4 and 5, the confidence intervals contain 1 and so, we cannot reject the null hypothesis and as a result, our Poisson model passes the test.

6. Conclusion

Exogenous shocks, which may be defined as sudden or unexpected events that have significant influences on an economy or an industry or group of industries, often lead to significant changes in demand, regulation, and cost structure (Gorbenko and Strebulaev, 2010). These shocks can be classified as temporary or permanent shocks, which depend on their duration, can have positive or negative impacts in short-run or long-run.

Over the last ten years, the relative competitiveness of the highly integrated and interdependent Italian cattle industries has been affected by a series of national and international market and policy shocks. These have included changes in the price of the cattle feeds and the introduction of mandatory country of origin labeling (regulated by EU) and etc. As we present some information on the reports on the outlook of Italian cattle industry in the introduction of the chapter, these shocks have had a significant impact on the competitiveness of Italian cattle and livestock, and in the same way, have probably affected the movement of cattle herds among different locations.

In this chapter, we were looking for some of the most important factors that influence on the movement patterns of cattle herds among Italian provinces. As such, We employed a structural gravity model of trade for Italy bovines movement to assess some of this factors. We found that the size and actually the stock of the heads in each province is a significant factor. Specifically, the more are the stock of heads, the more are the number of movements. In fact, it shows that the provinces with more stock of cattle heads are more likely to trade with each other.

One other factor that has an important and influential impact in the cattle trade pattern was distance among trading partners. We saw that the more is the distance between provinces, the less is the trade among them. It can be attributed to transportation costs such as energy, workers, facilities, etc., which hinder the economic agents to trade with each other in far distances.

One other factor that was also our variable of interest was feed price or specifically since the major bovine feed in Italy is corn, it was corn price. We found that the feed price is one of the significant factors that can explain the cause of bovines movements among Italian provinces.

Another important finding of this chapter was the effect of financial literacy rate of farmers on the pattern of bovine movement across Italian provinces. We found that the financial literacy rate has a positive (main) effect on the movement of bovine and it shows that in general, the more is the level of financial literacy of a province; the farmers in that province tend to trade more and send their bovines to other places (provinces). Actually, the more the farmers are financially literate; they will trade more to each other. Also, our estimates revealed that by increasing 10% in relative financial literacy, the bovines movement increase by 5.5%. In addition, AME result shows that 1% increase in financial literacy rate, can lead to 3 additional bovine movements.

In addition, we saw that the interaction of financial literacy and feed price has a negative and significant effect on the bovines' movements among provinces. In fact, this presents that the more the residents of a province are financially literate, the less is the effect of relative feed prices on their decision-making process. In addition, by increasing financial literacy, the farmers can become immune in the exposure of price shocks. Actually, a financially literate farmer can take into account more parameter in its decision-making, and then, for example, he can hedge its future trade and investments to benefit more and so on.

Therefore, as we said to immune the farmers in front of various (feed price) shocks, policy implication could be in the sense that to improve their financial literacy status. By doing so, they can think about, for instance, making some mutual agreements to mitigate their future risks and become not much affected by severe (price) shocks. Actually, farmers can protect themselves against risks through understanding risk mitigations and different types of insurances. In this respect, we can propose some policy implication in the sense that to enhance the economic understanding of farmers in case of price changes. For example, Public futures markets were designed in the 19th century[36] to allow clear, standardized, and efficient hedging of agricultural commodity prices. More specifically, the authorities can schedule some training programs to enable farmers to monitor the market and even try to predict price trends. In fact, these programs can make the farmers stronger in understanding the market and can help them to find the best period for buying/selling their products. Actually, this kind of decisions to buy in advance if the

---

[36] "A survey of financial centres: Capitals of capital". The Economist. 1998-05-07. Retrieved 2011-10-20.

level of prices is low is an intertemporal decision that is related to financial literacy. For instance, buy in a large amount in the time of cheap prices and store them as long as possible. On the other hand, they can try to arrange some agreement and contracts with their partners for future dates in fixed prices. This can enable them in planning and controlling the unexpected volatilities, too.

Moreover, as we know, farmers are facing many risks related to their business. So, understanding risks and importance to hedge the risks can be an influential factor in determining their business behaviors. One of the important sources of risk for any farm and any agricultural economic sector is bovine infectious diseases that can directly affect on the cattle and meat industries. In this respect, governments can devise some strategies to improve the economic and financial literacy of their farmers to decrease the possible losses in case of spreading diseases in their cattle industry. For example, some useful strategies could be investing and taking part in vaccination programs for their animals prior to the outbreak of any kind of disease, or some collaboration with other farms to fight against diseases or using healthy eating systems for bovines, and so on and so forth. Therefore, as we can see the level of financial literacy of farmers can have a significant role in determining their trade decisions. In the next chapter, we will discuss this disease factor and its influence in bovine's movement patterns. Also, we will explore the interaction effect of financial literacy rate and disease status.

Also, one research idea can seems interesting for further studies is that to look at some correlations of feed price and weather situations. We know that the weather conditions are one of the important instrumental variables in determining agricultural economic sectors. Specifically, lack of rain can directly affects on the corn price shocks. Because farmers may harm yield losses due to dry weather and drought conditions (especially corn which needs a lot of water). In fact, dry weather will influence cattle and dairy farmers who grow forage such as hay, corn, corn silage and etc, for their livestock, And as a result they might force to buy their required feed from other places, resulting in higher costs. Also, this weather condition can increase the spread of infectious animal diseases among livestock. At the same time, including changing market prices for corn and other crops, the farmers could face with difficult business status.

In sum, we saw in this chapter that as well as distance and size of the respective economic sector, some other factors such as feed price and financial literacy of the farmers are important and effective in determining the pattern of bovines' movement. Of course, we know that these factors are not the sole determinants of the pattern of bovine trade among different locations and actually, some other exogenous shocks can be affected by this pattern. One important and probably very significant element would be the risk of spreading a disease among cattle in a farm or in a province or country and in general in an international level. In chapter 2, we try to address this subject and analyze the risk of contagious diseases among bovines and the effect of these diseases on bovines' movement.

References

- Anderson, J. E. (1979) "A theoretical foundation for the gravity equation", *American Economic Review 69(1):* 106-116.
- Anderson, J. E. & Marcouiller, D. (2002) "Insecurity And The Pattern Of Trade: An Empirical Investigation", *The Review of Economics and Statistics, MIT Press, 84(2):* 342-352.
- Anderson, J.E., Van Wincoop, E. (2003) "Gravity with gravitas: A solution to the border puzzle", *American Economic Review (93):* 170–192.
- Baltagi, B. H. (2008): "Econometric Analysis of Panel Data", *4th edition, John Wiley & Sons.*
- Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum (2003) "Plants and Productivity in International Trade", *American Economic Review, 93:* 1268-1290.
- Bergstrand, J. H. (1985) "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence", *Review of Economics and Statistics, 67(3):* 474–81.
- Bergstrand, J. H. (1989) "The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade", *Review of Economics and Statistics, 71(1):* 143– 53.
- Burger M., F.G. van Oort and G.M. Linders (2009) "On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-Inflated Estimation", Spatial Economic Analysis, 4(2): 167-90.
- Carmela Aprea, Eveline Wuttke, Klaus Breuer, Noi Keng Koh, Peter Davies, Bettina Greimel-Fuhrmann, Jane S. Lopus. (2016) "International Handbook of Financial Literacy", *Springer Singapore*.

- Carrère, C. (2006) “Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model.”, *European Economic Review, 50(2):* 223- 247.

- Chen, N. and D., Novy (2011) “Gravity, Trade Integration and Heterogeneity across Industries”, *Journal of International Economics, Vol.85 (No.2):* 206-221.

- Deardorff, A. (1998) “Determinants of bilateral trade: Does gravity work in a neoclassical world? In: Frankel, J.A. (Ed.)”, The Regionalization of the World Economy, (Chapter 1).

- Eaton, J., and A. Tamura (1994): "Bilateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment Patterns", *Journal of the Japanese and International Economies, 8(4):* 478-510.

- Eichengreen, B. and D. A. Irwin (1998) “The Role of History in Bilateral Trade Flows”, In: Jeffrey A. Frankel, (Eds), The Regionalization of the World Economy (pp.33-62). Chicago, IL and London: University of Chicago Press.

- Flowerdew, R. and M. Aitkin (1982) “A Method of Fitting the Gravity Model Based on the Poisson Distribution”, *Journal of Regional Science, 22:* 191-202.

- Fornero, Elsa and Monticone, Chiara, (2011) “Financial literacy and pension plan participation in Italy”, *Journal of Pension Economics and Finance, 10:* issue 04, p. 547-564.

- Fort, Margherita & Trucchi, Serena & Manaresi, Francesco (2016) “Adult financial literacy and households's financial assets: the role of bank information policy”, *Economic Policy - vol. 31(88):* 745-782

- Frankel, Jeffrey A. (1997) ”Regional Trading Blocs in the World Economic System”, Institute of International Economics.

- Gómez-Herrera, E. (2013) “Comparing alternative methods to estimate gravity models of bilateral trade”, *Empirical Economics, Springer, 44(3):* 1087-1111.

- Gorbenko, A., and Strebulaev, I. A. (2010) “Temporary versus permanent shocks: Explaining corporate financial policies”, *Review of Financial Studies, 23(7):* 2591–2647.

- Gould, David M. (1994) “Immigrant Links to the Home Country: Empirical Implications for U.S. Bilateral Trade Flows”, *The Review of Economics and Statistics, MIT Press, vol. 76(2):* pages 302-16, May.

- Gourieroux, C., Montfort, A. and A., Trognon (1984) “Pseudo Maximum Likelihood Methods: Applications to Poisson Models”, *Econometrica, 52:* 701-20.

- Head, K., and Mayer, T., (2014) “Gravity Equations: Workhorse, Toolkit, and Cookbook”, *Handbook of International Economics, Elsevier.*

- Hebert, C.J., and B.C. Anderson. (2011) “The Exogenous Effect of Corn Prices on Beef Prices in the U.S. Using R&D as an Instrumental Variable”, *Economics Honors Thesis, Colgate University.*

- Heckman, J. (1979) “Sample Selection Bias as a Specification Error”, *Econometrica, 47:* 153-161.

- Helpman, E. and P. Krugman (1985) “Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy”, *Cambridge, MA: MIT Press.*

- Helpman, E., Melitz, M. and Rubinstein, Y. (2008) “Estimating trade flows: trading partners and trading volumes”, *Quarterly Journal of Economics, (73):* 441–486.

- Jappelli, T. (2010). ‘Economic literacy: an international comparison’, The Economic Journal, 120: F429–51.

- Kareem, Fatima Olanike & Martinez-Zarzoso, Inmaculada & Brümmer, Bernhard, (2016) “Fitting the Gravity Model when Zero Trade Flows are Frequent: a Comparison of Estimation Techniques using Africa's Trade Data”, Discussion Papers 230588, Georg-August-Universitaet Goettingen, GlobalFood, Department of Agricultural Economics and Rural Development.

- Linders, G. J. M. and H. L. F. de Groot (2006) “Estimation of the Gravity Equation in the Presence of Zero Flows”, *Tinbergen Institute Discussion Paper, No. 06-072/3.*

- Manning, W., and J., Mullahy (2001) “Estimating log models: to transform or not to transform?”, *Journal of Health Economics, 20 (4):* 461–494.
- Martin, W. & Pham, C. S. (2008) “Estimating the Gravity Model When Zero Trade Flows are Frequent”, Economics Series 2008_03, Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance.
- Martinez-Zarzoso, I. (2013) “The log of Gravity Revisited”, *Applied Economics, 45(3):* 311-327.
- Melitz, M. J. (2003) “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity”, *Econometrica, (71):* 1695-1725.
- Melitz, M., Ottaviano, G. (2008) “Market Size, Trade, and Productivity”, *Review of Economic Studies, 75 (1):* 295-316.
- OECD (2013a). “Evaluating Financial Education Programmes: Survey, Evidence, Policy Instruments, and Guidance”. http://www.oecd.org/finance/financial-education/G20-Evaluating_Fin_Ed_Programmes_2013.pdf
- Park, R. (1966) “Estimation with Heteroskedastic Error Terms,” Econometrica, 34, 888.
- Rauch, J.E., and V. Trindade (2002) “Ethnic Chinese Networks in International Trade”, *Review of Economics and Statistics 84(1):* 116-130.
- Rose, A. (2004) “Do we really know that the WTO increases trade?”, *American Economic Review, 94 (1):* 98–114.
- Santos Silva, J.M.C., and Tenreyro, S. (2006) “The log of gravity”, *The Review of Economics and Statistics, (88):* 641-658.
- Santos-Silva, J., and S., Tenreyro (2011) “Further Simulation Evidence on the Performance of the Poisson-PML Estimator”, *Economics Letters, 112 (2):* 220–222.
- Sören, P. and B. Bruemmer (2012) “Bimodality & the Performance of PPML”, Institute for Agriceconomics Discussion paper 1202, Georg-August Universität Göttingen, Germany.

- Tinbergen, J. (1962) “Shaping the World Economy: Suggestions for an International Economic Policy”, New York: The Twentieth Century Funds.
- Turkson F.E. (2010) “Logistics and Bilateral Exports in Developing Countries: A Multiplicative Form Estimation of the Logistics Augmented Gravity Equation”, Discussion Papers 11/06, University of Nottingham, CREDIT.
- Xiong, B. and J. Beghin. (2013) “Aflatoxin Redux: Does European Aflatoxin Regulation Hurt Groundnut Exporters from Africa?” *European Review of Agricultural Economics, Foundation for the European Review of Agricultural Economics, vol. 40(5):* 895-895.

Appendix:

Table A: regressions controlled for the wealth

VARIABLES	(1) Dependent Variable: interprovince heads (FE)	(2) Dependent Variable: interprovince heads (RE)
Stock of origin province (log)	0.553***	0.554***
	(0.0186)	(0.0185)
Stock of destination province (log)	0.212***	0.213***
	(0.0128)	(0.0128)
Distance_km (log)	-	-0.440*
		(0.231)
Unemployement rate of origin province (log)	0.418***	0.403***
	(0.0596)	(0.0594)
Unemployement rate of destination province (log)	0.194***	0.190***
	(0.0567)	(0.0565)
Relative feed price (log)	1.856***	1.787***
	(0.301)	(0.301)
Financial literacy rate (log)	-	3.099**
		(1.246)
2008.year	-0.0824***	-0.0797***
	(0.0210)	(0.0210)
2009.year	-0.675***	-0.668***
	(0.0372)	(0.0371)
2010.year	-0.395***	-0.385***
	(0.0505)	(0.0504)
2011.year	-0.302***	-0.293***
	(0.0438)	(0.0437)
2012.year	-0.428***	-0.416***
	(0.0572)	(0.0570)
2013.year	-0.359***	-0.342***
	(0.0760)	(0.0757)
Observations	5,877	5,879
Number of provorigdest_uniquelink	85	87

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: gravity equation results based on Poisson maximum likelihood estimation method. Columns (1) and (2) are estimated by fixed effects (FE province-pairs) and random effects (RE), respectively. Standard errors are clustered by origin-destination nodes (province-pairs unique links). Regressions control for the wealth (index:unemployment rate)

Table A continued (month dummies).

VARIABLES	(1) Dependent Variable: interprovince heads (FE)	(2) Dependent Variable: interprovince heads (RE)
2.month	0.0639**	0.0634**
	(0.0288)	(0.0288)
3.month	0.266***	0.266***
	(0.0271)	(0.0271)
4.month	0.0211	0.0212
	(0.0292)	(0.0292)
5.month	-0.0530*	-0.0533*
	(0.0290)	(0.0290)
6.month	0.0401	0.0403
	(0.0294)	(0.0294)
7.month	0.247***	0.247***
	(0.0280)	(0.0280)
8.month	0.165***	0.165***
	(0.0290)	(0.0290)
9.month	0.0639**	0.0642**
	(0.0288)	(0.0288)
10.month	-0.116***	-0.112***
	(0.0299)	(0.0299)
11.month	0.0909***	0.0912***
	(0.0290)	(0.0290)
12.month	0.0747**	0.0748**
	(0.0299)	(0.0299)
Constant interprovince_heads	-	-2.354
		(1.585)
Constant lnalpha	-	0.207
		(0.137)
Observations	5,877	5,879
Number of provorigdest_uniquelink	85	87

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Chapter II

Impact of disease incidence rate in bovines movement patterns of northern Italian Provinces: a “gravity model” approach

1. Introduction

Animal health and generally animal related products gradually increase their influence on the pattern of agricultural trade. World Trade Organization (WTO) members can use Sanitary and phytosanitary (SPS) measures to avoid the outbreak of various diseases and pests among livestock and plants. In fact, SPS agreement endorsed at the conclusion of the “Uruguay Round of the Multilateral Trade Negotiations” to protect humans, animals, and plants from diseases, pests, or any contaminants.[37] This agreement allows WTO member countries can apply their own regulations and standards to assess the disease risks, provided to their method is based on scientific criteria and must not be inconsistent with any other provision of this Agreement. In general, this agreement is defined to protect human life from the plant- or animal-carried diseases, and to protect animal or plant life from pests, diseases, or disease-causing organisms, and to prevent or limit other damage to a country from the entry, establishment or spread of pests.[38]

WTO members should apply these measures of SPS agreement “to protect human or animal life from risks arising from additives, contaminants, toxins or disease-causing organisms in their food, beverages, feedstuffs; to protect human life from plant or zoonotic diseases (“a disease that can be transmitted from animals to people or, more specifically, a disease that normally exists in

[37] unctad.org/en/PublicationsLibrary/ditctab20122_en.pdf
[38] https://www.wto.org/english/tratop_e/sps_e/spsund_e.htm

animals but that can infect humans"[39]); to protect animal or plant life from pests, diseases, or disease-causing organisms; to protect a country from damage caused by the entry, establishment or spread of pests[40]". Although these measures are mostly applicable in international trading matters, but as some scholars such as Yue and Beghin (2009) have mentioned that, the countries should not have unfavorable food safety standards in their countries, which can result in welfare losses of domestic consumers. Because some of these standard measures are toward issues that directly influence on the animals and domestic consumers. For instance, when a country struggles with an outbreak of a dangerous disease like foot-and-mouth, Bovine Brucellosis, Bovine Tuberculosis, Bluetongue, etc., among its cattle. In the same time, other countries, which are free from this particular type of disease, can use the SPS measures to hinder the transmission of this disease that could be occurred via trade. So, as we see, these disease outbreaks can significantly influence the pattern of trade between trading partners. The intensity of these effects depends on many other issues such as the level of the incidence rate of disease transmission, treatment policies that are adopted by countries, cost of animal movements, and in general domestic market conditions of trading partners. Similarly, these issues are also relevant at a national level where different regions can trade with each other. Thus, analyzing the risk of these animal diseases can have an important effect on the pattern of trade among different locations.

The outbreak of infectious diseases among livestock, and specifically cattle, has brought about much attention in different countries. For instance, some of these disease like Brucellosis, Swine flu, Foot-and-mouth, Avian influenza, and Tuberculosis disease have caused an influential

[39] http://www.medicinenet.com/script/main/art.asp?articlekey=12958
[40] WTO (2010)

problem in the world[41] (World Resource Institute 2005[42], Bengis et al. 2002, National Academies 2005). Also, some studies such as Taylor and et al. (2001) and National Academies (2005) have shown that animal diseases are an important reason for human infection diseases. Also, Delgado et al. (2003) point out that this kind of diseases can impose such influential economic impacts on various agricultural sectors and generally national economy as well as influencing on the farmers. They show that Mad Cow disease in the United Kingdom is a valid example to figure out these harmful effects on the cattle industry and the economy as a whole. Or as Horan and Wolf (2005) have said, the spread of contagious diseases can affect on the products and services that come from these animals. So, all of these concerns, show that we have to consider these animal diseases in more serious and since the outbreak of these diseases can influence on various economic sectors, especially agriculture, in both national and international patterns of trade, we should take into account control and prevention policies in more precise manner. So, as we can anticipate, human and animal transfer and generally environmental changes can influence on the pattern of disease transmission among human and animals.

Therefore, it would be pretty probable that cattle movement among different premises and locations can play an important role in the transmission of a disease in different regions of a respective country like Italy. So, a better understanding of cattle movement patterns can help us to analyze or even forecast the transmission of diseases and relevant economic effects. Of course some works have been done to assess this kind of movements, for instance, Livingstone et al. (2006) show that the transfer of cattle herds from one location to another can increase the

[41] For more details, we can see the online databse presented by World Global Atlas of WHO (World Health Organization). In a single electronic platform, the WHO's Disease Global Atlas is bringing together for analysis standardized data and statistics for infectious diseases at global, country, and regional levels. http://apps.who.int/globalatlas/

[42] Refer to full report provided by World Resource Institue: pdf.wri.org/wrr05_dt_all.pdf

possibility of outbreak of a disease (in their case, bovine tuberculosis) among cattle herds by transmission of disease from infected places to free-infected place and so has a significant impact on the pattern of trade between different regions. So, this kind of movement of cattle, and in general cattle trade, can increase the speed of disease diffusion to much larger scales and so makes it more difficult to deal with disease risks. Also, we can find similar phenomena in the literature. For instance, Hartup et al. (2000) and Schmitt et al. (2002) showed that some human interventions in habitat and feeding way of some wildlife such as bison and elk, can lead to a significant impact on the pattern of disease transmission in a different location.

So, as we see that if we ignore the human impacts on a system of cattle disease transmission, it can lead to ineffective and costly disease control strategies. However, the way that we can detect disease among cattle is also important. Because if we were able to find out the infected cattle, so, we could remove those animals from trade cycle, but as Williams et al. (2002) and Lanfrachi et al. (2003) have also mentioned that many infectious diseases are hard to recognize and they can be identified in final phases of infection. And therefore, it seems that the most control policies are led to unsatisfactory results. For example, Horan and Wolf (2005) say that most bio-economic works have led to imperfect control strategies, mostly because of the lack of technological advancements.

The outbreak of contagious disease around the world, have been influenced a lot on various economic sectors. For example, the spread of FMD disease in Taiwan brought about roughly 387.6 million dollars in the year 1997. This amount was only the direct cost of this outbreak, and the total cost of spreading disease was estimated about 1.6 billion dollars (Yang et, al. 1999). Similarly, Thompson et al. (2002) assessed the economic costs of the outbreak of foot-and-mouth disease in the UK and found out the estimated loss in agricultural sectors and even tourism was close to 3.1 billion pounds. Similar studies have done by other scholars in the US and they also found that scale effects and regionalization have a significant impact on control and management strategies. For example, Viboud et al. (2006) used the seasonal influenza data of US citizens and employed a gravity model to analyze the relevant factors in the outbreak of this disease. They found that population size, distance, and transmission are the important and significant factors in explaining the spread of respective disease.

In general, we can see various factors to influence in an outbreak of infectious cattle diseases between countries and also within a country due to the movement of animals between different locations. Therefore, understanding and recognition of the movements of cattle and the pattern of transferring these animals is very important in addressing control and surveillance strategies to avoid disease transmission among cattle and more importantly the outbreak of diseases in larger scales. This point would be more critical when we are dealing with a situation which resources are scarce and limited.

We can see that the previous studies on the pattern of livestock movements are mostly concentrated on the population size of the premises and spatial/geographical factors and not much on behavioral and motivational aspects of the trade which can be a driving force for the farmers to send their livestock to other places. Here in this chapter, we are going to analyze one of those motivational/behavioral factors, disease incidence rate, that can affect the decision of farmers for sending their cattle to different regions. In fact, by doing this, they want to avoid their cattle from being more infected.

Based on the definition of the Centers for Disease Control and Prevention (CDC), "The basic incidence rate (sometimes called just incidence) is a measure of the frequency with which a disease occurs in a population over a period of time"[43]. Actually, this measure can be considered the best index for comparing the abundance and risk of any diseases among any population. We will give more insight about this measure and related issues in third part of the chapter.

So, in order to assess the pattern of animal trade and the outbreak of infectious diseases through a different location, we need a model to analyze these cattle movements. One important and well-known model in the literature of national and international trade that has been received much attention and success in recent decades is "Gravity model". In fact, these models include the economic size and spatial nature of the trade into a model of the movement of goods between different units, which these units can be in national (like provinces) and international (like countries) levels. It has a strong theoretical and empirical background which has been improved in recent years, as well. The traditional derivation of the model only contains economic masses and the distance between units, but in recent years, there have been some extensions to the model to incorporate more determinants and avoid omitted variable bias in empirical estimations. Head and Mayer (2014), in their survey, call the traditional version of Gravity model as "naive" and the augmented version of gravity model which was originally adopted from Anderson and van Wincoop (2003), as "structural form" of gravity equation. This structural form differs somewhat from a traditional model and it includes some resistance terms into the gravity equation. Such models have received substantial methodological progress in recent years. We will explain more about this model in next parts.

[43] For more detailed information, refer to "Principles of Epidemiology, Second Edition, An Introduction to Applied Epidemiology and Biostatistics"

Also, in general, Gravity models have been extensively employed by epidemiological scholars to model the geographical outbreak of various diseases among various populations such as cattle, pigs, and etc.

As we said in the previous paragraph, here we are going to apply an economic gravity model of Italy cattle trade to investigate disease transmission across cattle heads in the Italy. More specifically, we are attempting to use a structural gravity model of trade and fit it to cattle movement data in Italy using a rich panel dataset of livestock data. Then, we can see the pattern of the cattle trade between origin and destination points to better understand the driving force for this kind of movements.

One of the major advantages of the gravity models in the context of transmission models is that it can be applied to a larger study area which may not cover in the dataset in question. so, by fitting a model to cattle movement in northern part of Italy and extracting the pattern of movements, we would be able to make some predictions about the cattle movement flows of other regions which have somehow these types of movement and similar characteristics.

we can say that in previous decades, the main application areas which gravity models have been used, was the analysis of the demand for goods and services in geographically distributed populations. But in the recent years, the use of gravity models to analyze the role of people movement (like immigration/emigration patterns) in economic activities of the countries, have been increased. In general, gravity models for describing population flows from one place to another, incorporate the distribution of movement between different regions, by using the population size of origin and destination nodes and the geographical distance between those nodes.

One of the early studies that analyzed these type of people movement had done by Zipf in 1946. He actually proposed a theoretical intuition for the people movement from one city to another city. He states that these movements can be explained by some factors and a gravity form of relation among those factors. He shows that there is a multiplicative form of relationships among population size of cities and the distance between respective cities, as P1P2/d, where p is the population and d are the geographical distance. Then, he uses this relationship to describe the pattern of movements in passenger and freight among US cities. We can find this kind of immigration/emigration patterns in both national and subnational levels. Some studies have done on a sub-national scale and use the gravity model to explain the migration trade patterns in a country. For example, D'Ambrosio (2015) uses a gravity model of trade along with a panel dataset of Spanish provinces to find out the contribution of migrants to the export of respective provinces. Analogous to the use of gravity model for explaining the pattern of people movement, there have been many studies have employed the gravity model for describing the movement pattern of livestock among different locations.

All in all, in this chapter, we are going to use a structural gravity model of trade and link it to an epidemiological index (disease incidence rate) to figure out some patterns, incentives and generally driving forces for sending cattle from one place to another. Our finding could lead to enhance some disease control and management strategies. Also, by knowing the disease risk level in different locations, we can improve our analysis and somehow forecast better about our future investments. For example, our finding proposes that performing surveillance endeavors towards provinces that have higher risk levels, can be more effective than selecting provinces that have less exposure to disease risks.

The rest of the chapter is organized as follows. In section 2, we briefly explain the theoretical and empirical background of the structural gravity model, then in the third part, we introduce our descriptive dataset and a brief explanation on epidemiological incidence rate. Then, in section 4, we will present our estimation results and relevant discussions. Finally, in section 5, we conclude the chapter.

2. Theoretical and empirical background

The cattle movement patterns can be described as a subset of the international trade literature and the gravity equation to explain this kind of flows is covered in the theoretical and empirical literature of gravity models as a whole. The gravity model of trade has been extensively employed by economic scholars in determining trade patterns in last fifty years. Of course, there has always been some discussion about the theoretical and empirical foundation of the model in last few decades which has been led to some modification and actually some “augmented” gravity models. for example, Head and Mayer (2014) show in their survey that there have been severe debates in last two decades about the lack of micro-foundation theoretical background to gravity models and then, we have seen a considerable development in the theoretical foundation of the model. One of the most important modifications of the model is done by Anderson and van Winccop (2003), where they introduce a resistance term to the model as a potential obstacle to the trade. their model was called by Head and Mayer (2014) as a “structural form” of gravity model and also, it has been widely used in the literature for explaining trade flows. For example, In this respect, we can refer to work done by Baldwin and Taglioni (2006) that have named three common mistakes in the literature of gravity models, as "gold", "silver" and "bronze-medal mistakes".[44] We will briefly explain this structural gravity model in the next section.[45]

After some debates over the theoretical foundations of gravity model and as a result of it, some changes and modification in this regard, we have been faced with some controversies concerning with the specification and empirical issues of the gravity model. So, several estimation techniques have been introduced to deal with the problems due to empirical issues. The main empirical challenges are in presence of homoscedasticity and zero trade values. Because traditionally the estimated specification was to log-linearize the multiplicative form and then estimate it with ordinary least square (OLS) methods, where we assume homoscedastic error terms (variance of the error is constant over observations). But it has been confirmed that the process of log-linearizing the gravity model and respective OLS estimation technique, is not efficient and actually we are facing with heteroscedasticity in this situation. Also, log-linearizing the gravity equation is troublesome when we have zero trade values because as we know the logarithm of zero is meaningless and indefinite. In this regard, many economic scholars have proposed several alternative estimation techniques in the empirical literature of gravity models which we can not see a commonly accepted method among all of them, nevertheless, it can be seen that the work done by Santos and Silva (2006) is more approved technique and the respective estimators perform better than other estimators. In next sections, we will give a brief introduction to some of the estimation techniques which have been introduced in the empirical literature.

[44] The silver-medal mistake which can be found in some studies employs the average of bilateral flows as a proxy for dependent variable and instead of using the average of the logs, it is associated with the logarithm of the "uni-directional flows" average. On ther other hand, the bronze medal mistake is associated with improper deflation of nominal trade values by the country's aggregate price index, while the inclusion of this parameter can lead to biases through spurious correlations. Because it ignores the global trends in inflation rate.

[45] For more technical detailed information on the structural gravity model, Refer to chapter 2.

2.1 The gravity model

It's been more than 50 years that the gravity models of trade have become a workhorse in analyzing economic and social sciences. The concept of the theory is that the bilateral trade relations between two entity are directly proportional to their economic masses and inversely proportional to the distance between them. This concept was first transformed in a gravity equation of trade by Tinbergen (1962). This equation was not theoretically micro-founded in that time and Head and Mayer (2014), in their review, called this equation as a "naive" gravity equation:

$$X_{ni} = GY_i^a Y_n^b \emptyset_{ni}$$ [46]

Which in the above equation, X_{ni} denotes the amount of trade between two country of n and i[47], Y_i denotes economic mass (i.e. gross production) of country i, Y_n represents the gross production of country n, a and b are parameters that has been involved in the equation and should be estimated[48]. G is a constant and finally $\emptyset_{ni}$ is a term that denotes the trade costs like distance between two countries. it actually join trade costs with their respective elasticities to measure the overall impact on trade flows (Head and Mayer, 2014).

Baldwin and Taglioni (2006) shows that the theoretical foundation of gravity equation has improved in recent decades. One of the earliest studies for strengthening the theoretical background of gravity equation offered by Anderson (1979). Later, Anderson and van Wincoop in 2003 pointed out that there is a problem in the naive version of the gravity model. It lacks a term that can take into account more trade costs into the equation. As we said above Baldwin and Tagliani (2006) refer to this omitted variable problem as the "gold medal mistake" which we can

[46] Notations in this naive version of gravity equation, are based on Head and Mayer (2014).
[47] In their version, n is importing country and i is exporting country.
[48] These two parameters are not necessarily equal to 1.

see in almost all studies before Anderson and van Wincoop (2003). Actually, Anderson and van Wincoop (2003) showed that if we are looking for more accurate estimations, to avoid above mentioned omitted variable bias, this version needs a bilateral resistance term. Head and Mayer (2014) call this category of gravity equations as “structural form” of gravity models.

The Anderson and van Wincoop model is the first “structural” form of gravity equation which can be derived from a general equilibrium framework. This model differs partly from the naive version of the gravity model. It includes a bilateral resistance term that can be considered as a treatment to the omitted variable bias problem. There are some assumptions in the model of Anderson and van Wincoop (2003). First, they consider a Constant elasticity of substitution (CES) utility function for consumers who maximize this utility function subject to a standard budget constraint. Also, they assume that the commodities are differentiated by place of origin and each entity (country) only produce one. Then by using market clearance condition and some further mathematical technicalities, they obtain their gravity equation which includes an extra term more than the naive version of the gravity model. They call This term as “trade resistance” term.[49] Later, Head and Mayer (2014) call this type of gravity equation as “structural form” gravity

$$X_{ij} = \frac{Y_i Y_j}{Y_w}\left(\frac{t_{ij}}{\Pi_i P_j}\right)^{1-\sigma}$$

Where in the above equation, X_{ij} is the volume of trade between two nodes (which can be any units like countries, cities, and specifically in our case provinces) of i and j. Y_i and Y_j are the gross economic masses of nodes i and j (which is usually the GDPs of the nodes i and j). Y_w represents

[49] Refer to chapter two of this thesis for detailed mathematical derivation of structural gravity model

the world income (which can be considered constants across units) and finally the most, important part is $\left(\frac{t_{ij}}{\Pi_i P_j}\right)$ which includes all trade barries and costs in our model.

Then, the common practice in the international trade literature is to log-linearize this equation and estimate it through some econometric methods.

2.2 Estimation methods of gravity models

In addition to the theoretical challenges in the literature of international trade, there are also more challenges about dealing with the empirical analysis of gravity models. Actually, we can see two major challenges regarding this empirical issues. One is the appropriate specification of gravity equation and the second one is that which estimation method is suitable for estimating gravity equation.

To point the first challenge, there is almost a general accepted practice regarding appropriate specification of gravity model to log-linearize the gravity equation and then estimate it. But the challenge about the appropriateness of logarithmic transformation is not so straightforward especially when we have heteroscedasticity and excessive zero values in our dataset. Actually, there is a lot of debates in recent years about the appropriate estimation technique for estimating this logarithmic transformation of gravity equation. Although, we can say that there is not a generally accepted method for estimating gravity equation, but some of them have gained much more popularity among various methods. Here in this part we just point out a brief review of proposed methods and explain which one is better somehow for our case. for a detailed description of different methods, refer to the second chapter of the thesis.

Anderson and van Wincoop (2003) and Feenstra (2004) showed that using country fixed effects is suitable method when we are going to specify resistance terms. But Santos Silva and

Teneyro (2006) pointed out that besides that the gravity equation can be controlled by fixed effects, but the problem of heteroscedasticity is still valid in the model and we should find a method to cope with that. To address this issue, Santos Silva and Teneyro (2006) proposed a pseudo-Poisson maximum likelihood (PPML) method which can be estimated in levels and it can deal with both the problem of zero trade values and heteroscedasticity of the dataset. They also used the monte-Carlo simulation to compare the fitted values of Poisson maximum likelihood estimator and least square estimators and found that the PPML estimator is more robust and relatively well-behaved than other estimates. Of course, the dependent variable in their original work in 2006 was completely positive, and to check that if their estimator is still valid in the case of zero values, They performed a study in 2009 and found that the PML is well-behaved even if we have had a non-negative dependent variable. Also, Westerlund and Wilhelmsson (2009) used a panel data set and monte Carlo simulation to assess the effect of zero trade values on the gravity models and found that the among the estimators, the Poisson fixed effect estimator is the best.

As we said, one issue that has to be addressed in the estimation of gravity equations is the presence of zero trade values. This issue would not be problematic when we have multiple aggregated goods in the analysis, but if we are dealing with one good/commodity, there may be a large number of zero values, and so, we will have to cope with this issue. Among different estimators, Santos Silva and Teneyro (2009;2011) pointed out that the performance of PPML estimator is the highest in the case of excessive zero values. Also, Staub and Winkelmann (2012) show that the PPML estimator is consistent in case of heteroscedasticity and model misspecification. One another advantage of PPML which proposed by Fally (2012) is that estimated fixed effects are consistent with the definition of bilateral resistance terms in a general equilibrium framework.

In sum, we can say that although PPML estimator has received much more attention in recent years, however, there is not a commonly "single" accepted practice to deal with the problems of heteroscedasticity and zero trade values. In fact, it is highly dependent on the nature of the dataset in question. So, we can conclude that the Pseudo Poisson Maximum Likelihood estimator can be our benchmark in analyzing cattle movement pattern.

3. Material

3.1 Data Description

Analogous to the previous chapters, Data on bovine's movements were obtained from the Italian National Bovine Database, which is administered by the Italian National Animal Identification and Registration Database. The dataset contains the movement of the all Italian population of bovines between animal premises, providing an extensive image of the locations that bovines have been kept and moved within the country. Additional information was provided for the animal holdings, including the type of holdings (i.e. animal husbandry or farm (AL[50]), Genetic Material Center (CG[51]), Collection Center (CR[52]), Fair/Market (FM[53]), Pasture (PA[54]), Point Park (PS[55]), Animal House (ST[56]), Stable (SS[57]) slaughterhouse (SM[58]),), and their geographic coordinates and also the geographical locations of the municipality where the holdings were located. Each movement case shows the unique identification for the animal, the codes of the

[50] ALLEVAMENTO
[51] CENTRO MATERIALE GENETICO
[52] STABILIMENTO DI MACELLAZIONE
[53] FIERA/MERCATO
[54] PASCOLO
[55] PUNTO DI SOSTA
[56] STABULARIO
[57] STALLA DI SOSTA
[58] STABILIMENTO DI MACELLAZIONE

holdings of origin and destination nodes, and the date of the movement. Such tracking system allows us to easily follow the path of each bovine and to make the equivalent overall network, and then we can minimize the problems with respect to data accuracy that are found in other tracking systems that do not provide both origin and destination of the movements (Volkova et al. 2010, Christley et al. 2005).

Therefore, here we have the dataset of cattle movements in Italy during the time period of 2006 to 2015. Actually, we excluded years 2006, 2014, and 2015, because we do not have a rich dataset of that year. In fact, we manipulated the dataset and simplify it (based on the availability of the data) for only 10 provinces during 2007 to 2013, which can be seen that the most important with most trade relations occur among them. These provinces include Alessandria, Bologna, Brescia, Milan, Mantova, Modena, Padova, Turin, Treviso, and Verona.[59] In table 1, we can see the definition and descriptive statistics of required variables in our study.

Table 1: descriptive statistics for estimating gravity equation

Variable	Obs	Mean	Std. Dev.	Min	Max
Interprovince_heads[60]	6,762	5.500438	13.77986	0	242
heads_prov_orig[61]	6,764	241.7155	110.1758	25	834
heads_prov_dest[62]	6,764	477.4766	360.8629	3	2975
Incidence_rate[63]	6,764	1.608137	1.816726	0.0139174	27.83705
distance_prov (km)	5,887	130.2133	70.24709	32.91862	361.4067

[59] We used the geographical coordinates of the capitals of provinces to estimate the geodesic distance between provinces. Data on latitude and longitude of the capitals derived from: en.comuni-italian.it/province.htm
[60] Number of heads which has been transferred between twp provinces in a time period (month).
[61] Number of bovines in origin province
[62] Number of bovines in destination province
[63] Incidence rate

Which in above table, “Interprovince_heads” represents the volume of the heads which have been transferred to two provinces, “heads_prov_orig” denotes the stock number of heads in origin node. Likewise, “heads_prov_dest” shows the Stock number of heads in the destination node. “distance_prov” is the geographical distance between two provinces[64], and finally, “origoverdest” represents the relative incidence rate of origin nodes over destination nodes.

The focus of this chapter is to see the effect of incidence rate on the movement of cattle among different provinces. But, first, we should define what an incidence rate is really about. In next part, we will try to go through some epidemiological literature and briefly describe the basic and relevant definitions and issues regarding the outbreak of diseases in a population.

3.2 Morbidity, incidence, and Prevalence

In general, there are several standard ratios and rates in epidemiological literature that can be used to measure and assess the morbidity and frequency of infectious diseases in a given population. Each one, based on the situation and the dataset available, can be used in any epidemiological analysis. Two major morbidity rates that scientists usually use are incidence rate and prevalence. Both of them measure the risk that an animal prone to contracting a disease in a population (Putt et. al. 1987).

Incidence rates can be considered as the most common practice for measuring and analyzing the effects of the contagious diseases in a given population.[65] It simply defined as the rate of new or newly diagnosed disease in a given time period. So, basically, it’s a measure of a disease

[64] We used the “Vincenty” module in Stata to calculate the geodesic distance of the provinces, indexed by their capital geographical coordinates.
[65] http://health.mo.gov/training/epi/

frequency in a population over a period of time. The resulting formula derived from this definition is as follows:

Incidence Rate = (new cases diagnosed during a time period)/ (population at risk during the same time period)

For example, if we want to compare the level of the risk in two population, the higher is the incidence rate of a population, means that population has a higher risk of spreading the disease than other population. Also, it's clear that this rate shows the changes from a healthy situation to a diseased situation and so, we should designate the time period in question. Usually, this time period is considered as yearly for surveillance purposes, but it's not necessarily one calendar year and as long as the time period is specified, we can use any period. In our case, and based on data available, we have used monthly time period.

As we said before, there is another measure for assessing the risk of disease contraction by an animal, which is Prevalence. Actually, Prevalence is the proportion of animals in a population which have a disease at a specified point in time (point prevalence), or over a specified period of time (period prevalence). So, it includes new and pre-existing infected animals, while in incidence rate we just include the newly diagnosed cases. So, the prevalence rate is mostly applied to analyze the burden of chronic diseases such as HIV, malaria and etc. in a population.

So, generally, morbidity relates to the ratio of infected animals in a given population and it contains two main rates, incidence rate, and prevalence. Of course, although both of these rates capture to some extent the risk level of an animal subjected to a specific disease in a given population, but we can say that the incidence rate gives us information about the risk of catching a disease and on the other hand, prevalence presents the how a disease spread out over a large area.

Of course, we should note that the prevalence and incidence rate of a disease may change under different cattle production systems. For example, Prevalence rates tend to be larger in extensive cattle production systems in comparison with more intensive beef and dairy cattle production systems where cattle are restricted in a farm (Tambi et. al. 2006). Here we have assumed that the provinces have the same cattle production system.

Since we don't have the number of the bovines that have died or cured in a time period (a month in our case), so, we can not calculate the prevalence and we use the incidence rate to assess the risk of spreading a disease in Italian cattle industry. Although the relation between prevalence and incidence rate depends on somehow on the nature of the disease in question, but we can say that both of them are moving in the same direction and so, they have a direct correlation together. So, whatever effect of incidence rate has on the movement of bovines among Italian provinces, the same pattern can be derived by the study of prevalence rates.

4. Estimation results

Here we are going to assess the effect of incidence rate on the Italian Bovine movements, specifically northern province. Along with our theoretical model, this incidence rates can be considered as a resistance term for trading cattle among different provinces. So, first, we should create the interest variable which is relative incidence rate of two trading provinces. To do so, we should divide the relative number of infected cattle in origin nodes by the relative number of infected cattle in destination nodes. Then, we have all variables and we just need to run our econometric specification. As we said in previous parts, the ideal estimator for our study is PPML estimator. So, we use the Poisson method to estimate the effect of incidence rate on the movement of cattle in Italian provinces. We can see the results in table 2.

Table 2. Incidence rate and feed price effect on bovines movement

VARIABLES	(1) Dependent Variable: interprovince_heads	(2) Dependent Variable: interprovince_heads
Stock of origin province (log)	0.589***	0.589***
	(0.0189)	(0.0189)
Stock of destination province (log)	0.151***	0.149***
	(0.0143)	(0.0143)
Relative feed price (log)	2.368***	2.464***
	(0.303)	(0.304)
Relative incidence rate (log)	0.0563***	0.0527***
	(0.00964)	(0.00966)
Relative incidence rate*Relative feed price (both in logs)	-	-1.316***
		(0.246)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: gravity equation results based on fixed effects (FE) Poisson maximum likelihood estimation method. Standard errors are clustered by origin-destination nodes (province-pairs unique links)

Table 2 shows the empirical results derived by estimating gravity equation in two version. The first column includes relative corn prices, which analogous with the previous chapter, this variable has a significant and positive sign and it shows that the farmers in provinces with higher feed prices tend to send their bovines to other places. The second column in table 2 shows the result of gravity equation estimation including with the interaction of relative incidence rate and relative feed prices.

Table 2 continued.

VARIABLES	(1) Dependent Variable: interprovince_heads	(2) Dependent Variable: interprovince_heads
2008.year	-0.0702***	-0.0678***
	(0.0210)	(0.0210)
2009.year	-0.464***	-0.468***
	(0.0227)	(0.0227)
2010.year	-0.103***	-0.102***
	(0.0215)	(0.0215)
2011.year	-0.0367*	-0.0411**
	(0.0209)	(0.0209)
2012.year	-0.0232	-0.0316
	(0.0218)	(0.0218)
2013.year	0.171***	0.163***
	(0.0211)	(0.0212)
2.month	0.0515*	0.0534*
	(0.0288)	(0.0288)
3.month	0.199***	0.198***
	(0.0273)	(0.0273)
4.month	0.0279	0.0331
	(0.0291)	(0.0291)
5.month	-0.0705**	-0.0680**
	(0.0290)	(0.0290)
6.month	-0.0166	-0.0145
	(0.0297)	(0.0297)
7.month	0.229***	0.235***
	(0.0280)	(0.0280)
8.month	0.130***	0.135***
	(0.0291)	(0.0292)
9.month	0.0798***	0.0795***
	(0.0286)	(0.0286)
10.month	-0.150***	-0.149***
	(0.0300)	(0.0300)
11.month	0.0787***	0.0789***
	(0.0290)	(0.0290)
12.month	0.144***	0.144***
	(0.0294)	(0.0294)
Log likelihood	-32057.755	-32043.411
Observations	5,877	5,877
Number of origdest_uniquelink	85	85

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

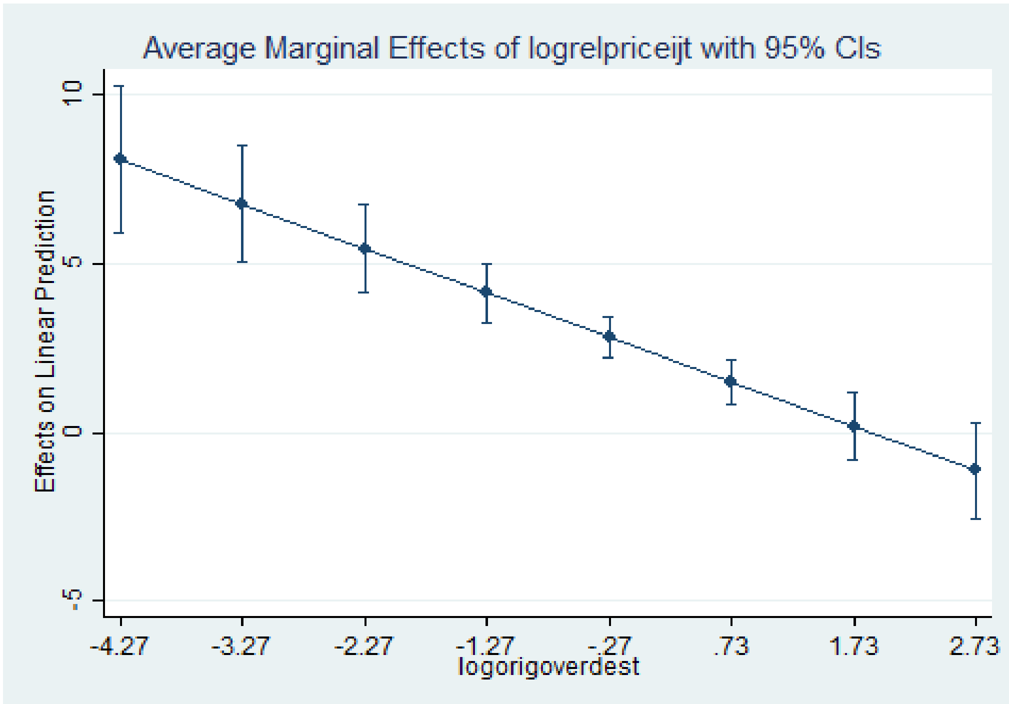
As we can see in table 2, our interest variable, relative incidence rate, has a significant and positive pure effect on the movement of bovines. It shows that the more is the level of the incidence rate of a province; the farmers in that province tend to send their bovines to other places. In other words, the more is the rate of new diseases in a province, the more is the outflow of the bovines. Actually, by doing this, they want to decrease the risk of contracting diseases for other heads and avoid existing bovines to be infected.

Quantitatively, increasing 10% relative incidence rate, it increases the bovines' movement by about 5.7%.

On the other hand, column 2 of table 3 presents that the interaction of incidence rate and feed price has a negative and significant effect on the bovines' movement among different provinces. A negative interaction coefficient means that the effect of the combined action of two independent variables is less than the sum of the individual effects. So, the connection between one of the independent variables and the dependent variable decreases if the other regressor increases. In fact, incidence rate undermines the effects of feed price shocks in the sense that the disease incidence rate parameter can make the feed price less relevant. This effect could be better seen by looking at the marginal effects diagram of relative price in different financial literacy rate (Figure1).

As we can see from the margins graph, increasing the relative incidence rate (origin over destination), the slope decreases and therefore, it shows that the more is relative incidence rate of provinces, the less is the effect of relative feed prices. Numerically, the y-axis of the graph shows the effect of feed price on the movement of bovines at different levels of incidence rate. In fact, in the case of diseases, the farmers do not care that much about the input costs (i.e. feed price) and they tend to send sick bovines to other places.

Figure 1: Predictive Margins of incidence rate on feed prices



4.1 Robustness Check

As we explained in chapter 2, for checking the robustness of the results, we can use a "Park-type" test proposed by Manning and Mullahy (2001). Actually, Manning and Mullahy (2001) showed that if the general form of the variance and conditional mean would be:

$$Var(Y|X) = K\ Exp(Y|X)^{\lambda}$$

To do the test, they state that the parameter λ can be determined by the auxiliary equation:

$$ln\left(Y - \hat{Y}\right)^2 = lnK + \lambda ln\hat{Y} + v_i$$

Which directly derive from the general equation of variance and conditional mean. We should test the null hypothesis of H_0: $\lambda = 1$ against the alternative H_0: $\lambda \neq 1$. The hypothesis will be accepted if the appropriate confidence interval for parameter λ includes 1. The result of the Park test is shown in table below.

Table 3: Park-type test

	Coefficient	95% Conf. Interval
lnyhat	1.07146	0.8686185-1.274302

Note: Park test for column (2) of table 2

As we see in table 4, the confidence interval contains 1 and so, we cannot reject the null hypothesis and as a result, our Poisson model passes the test.

5. Conclusion

Italian national animal identification and registration database have provided us an unprecedented opportunity for research into the pattern of bovines' movement in Italy. The bovines' data can use to merge with data on the infectious disease among Italian cattle to see the effects of the diseases on the pattern of animal trade among Italian provinces.

In this chapter, we employed a structural form of gravity model proposed by Anderson & van Wincoop (2003) which is an augmented version of traditional gravity models and it takes into account some "resistive" terms as an important factor in determining trade patterns between any two entities. Then, we used this model to analyze the epidemiological effect of diseases on the movement pattern of bovines among Italian provinces.

In general, as Rothman and Greenland (1998) state that epidemiological analysis is more informative when the single cases are examined in the context of the entire population which they arose, we tried to use disease bovine case in each holding and each time period and then aggregated them for each province in that time period. In this regard, expressing disease frequency on per head of bovines, that is as incidence rate (or incidence risk), makes it possible to compare the frequency of diseases among different provinces. Of course, we found that most of the movements occur in central and northern regions, specifically, we chose 10 provinces that include most cases of the bovine movement and positive disease test records, and then used this two type of dataset, i.e. movement data and disease record data, and achieved some results. Our results show that relative incidence rate of origin over destination nodes have a positive and influential effect on the pattern of bovines' movement from one province to another.

Furthermore, we tried to mix these results with the effects of feed prices that we argued in the previous chapter, as well. The results showed us that in the case of diseases, feed prices effects are still significant but not so important as we saw in the previous chapter. In other words, we found that incidence rate could undermine the effect of feed prices on the movement of bovines between holdings. This shows that in the case of diseases, farmers tend to send their bovines to other places in order to reduce the risk of contracting diseases among their heads.

In the next chapter, we will go through more details of the movements of bovines among various holdings and will assess the effect of the disease status on the movements of bovines among farms and slaughterhouses. Also, we will see the effect of network characteristics (degree, indegree, and outdegree) and interaction of these features with disease status on the transfer of bovines among Italian premises.

References

- Anderson, J. (1979) “A theoretical foundation for the gravity equation”. *The American Economic Review 69(1)*, 106–116.
- Anderson, J. E., & van Wincoop, E. (2003) “Gravity with gravitas: A solution to the border puzzle”. *The American Economic Review, 93(1)*, 170.
- Bengis RG, Kock RA, Fischer J. (2002) “Infectious animal diseases: the wildlife/livestock interface”. *Rev Sci Tech*. *2002; 21*:53–65.
- Christley RM, Robinson SE, Lysons R et al. (2005) “Network analysis of cattle movement in Great Britain. In: Society for Veterinary Epidemiology and Preventive Medicine”. Proceedings of a meeting held at Nairn, Inverness, Scotland on March 30-April 1, 2005. Mellor DJ, Russell AM, Wood JLN (Eds). 234-244.
- D'Ambrosio, A. (2015) “Migration flows and local systems of production: New comparative evidence on Italy and Spain”. Doctoral thesis, University of Trento.
- Daria Taglioni (2007) "Trade effects of the euro: A comparison of estimators”, with, *Journal of Economic Integration, 22(4)*, December, pp 780–818. Circulated as “Gravity for dummies and dummies for gravity equations” CEPR DP5850, and NBER WP 12516. 2007.
- Delgado, Christopher L., Narrod, Clare A., Tiongco, Marites M. (2003) *“*Project on Livestock Industrialization, Trade and Social-Health-Environment Impacts in Developing Countries” *Policy, Technical, and Environmental Determinants and Implications of the Scaling-Up of Livestock Production in Four Fast-Growing Developing Countries: A Synthesis*, Food, and Agriculture Organization, Rome.
- Feenstra, R. (2004) *“*Advanced International Trade: Theory and Evidence”. Princeton, New Jersey: Princeton University Press.
- Hartup, B.K., G.V. Kollias, and D.H. Ley. (2000) "Mycoplasmal conjunctivitis in songbirds from New York." *Journal of Wildlife Diseases 36(2):*257-264.

- Head, K. Mayer, (2014) “Th. Chapter 3 - Gravity Equations: Workhorse, Toolkit, and Cookbook”, In Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, Editor(s), *Handbook of International Economics, Elsevier, Volume 4:* 131-195.
- Horan, R.D., and C.A. Wolf. (2005) “The economics of managing infectious wildlife disease”. *American Journal of Agricultural Economics 87(3):*537-551.
- Lanfranchi, P., E. Ferroglio, G. Poglayen, andV. Guberti. (2003) “Wildlife Vaccination, Conservation and Public Health.” *Veterinary Research Communications 27*:567–574.
- Livingstone, P.G, T.J. Ryan, N.G. Hancox, K.B. Crews, M.A.J. Bosson, G.J.E. Knowles, and W. McCook. (2006) “Regionalisation: A strategy that will assist with bovine tuberculosis control and facilitates trade”. *Veterinary Microbiology 112(2-4):*291-301.
- Martin Ayong Ayim (2001) “Principles of Infectious Disease Epidemiology”. *Vita Press International*, 2001.
- National Academies. (2005) “Animal Health at the Crossroads: Preventing, Detecting, Diagnosing Animal Diseases”. Washington DC: National Academies Press, 2005.
- Putt S.N.H., Shaw A.P.M., Woods A.J., Tyler L. & James A.D (1987) “Veterinary epidemiology and economics in Africa: A manual for use in the design and appraisal of livestock health policy (ILCA manual)”. *International Livestock Centre for Africa,* 23-24
- Richard Baldwin & Daria Taglioni, (2006) “Gravity for Dummies and Dummies for Gravity Equations”. NBER Working Papers 12516, National Bureau of Economic Research, Inc.
- Rothman K, Greenland S. (1998) “Modern Epidemiology”. Lippincott-Raven, Philadelphia, USA.
- Santos-Silva, J. and S. Tenreyro (2006) “The Log of Gravity”. *Review of Economics and Statistics 88*, 641–58.
- Santos Silva, Joao M C & Tenreyro, Silvana, (2011) “poison: Some convergence issues”. Economics Discussion Papers 3534, University of Essex, Department of Economics.

- Schmitt, S.M., O'Brien, D.J., Bruning-Fann, C.S. and Fitzgerald, S.D. (2002) "Bovine tuberculosis in Michigan wildlife and livestock." *Ann N Y Acad Sci 969*, pp. 262-268.

- Staub, K. E., and R., Winkelmann (2013) “Consistent Estimation of Zero-Inflated Count Models”. *Health Economics, 22(6):* 673-686.

- T. Fally. (2012) “Structural gravity and fixed effects”. Working paper, University of ColoradoBoulder.

- Tambi, N. E., Maina, W. O. & Ndi, C. (2006) “An estimation of the economic impact of contagious bovine pleuropneumonia in Africa”. *Rev Sci Tech 25:* 999–1011.

- Taylor, L.H., Latham, S.M., Woolhouse, M.E. (2001) "Risk factors for human disease emergence. " *Philos Trans R Soc Lond B Biol Sci* 356: 9839.

- Thompson, D., Muriel, P., Russell, D., Osborne, P., Bromley, A., Rowland, M., Creigh-Tyte, S., Brown, C., (2002) “Economic costs of the foot and mouth disease outbreak in the United Kingdom in 2001”. *Revue Scientifique et technique (OIE) 21*: 675–687.

- Tinbergen, J. (1962) “Shaping the World Economy; Suggestions for an International Economic Policy”. New York: Twentieth Century Fund.

- Viboud, C., Bjørnstad, O. N., Smith, D. L. Simonsen, L., Miller, M. A., & Grenfell, B. T. (2006) “Synchrony, waves, and spatial hierarchies in the spread of influenza”. *science, 312(5772):* 447-451.

- Volkova VV, Howey R, Savill NJ, Woolhouse MEJ (2010) “Sheep Movement Networks and the Transmission of Infectious Diseases”. *PLoS ONE 5(6):* e11185.

- Westerlund, J., and F. Wilhelmsson. (2009) “Estimating the Gravity Model without Gravity using Panel Data”. *Applied Economics 41*:1466–4283.

- Williams, E.S., M.W. Miller, T.J. Kreeger, R.H.Kahn, and E.T. Thorne. (2002) "Chronic Wasting Disease of Deer and Elk: A Review with Recommendations for Management.” *Journal of Wildlife Management 66*:551–63.

- Willard Manning and John Mullahy. (2001) "Estimating log models: to transform or not to transform?" *Journal of health economics*, *20(4):* 461-494
- Yang, P.C., Chu, R.M., Chung, W.B., Sung, H.T., (1999) "Epidemiological characteristics and financial costs of the 1997 foot-and-mouth disease epidemic in Taiwan". *Vet. Record 145:* 731–734.
- Yue, C., and J.C. Beghin. (2009) "Tariff Equivalent and Forgone Trade Effects of Prohibitive Technical Barriers to Trade." *American Journal of Agricultural Economics 91(4)*: 930-41.
- Zipf, G. K. (1946) "The P1P2/D hypothesis: On the intercity movement of persons". *American Sociological Review 11:* 677-686.

Chapter III

Bovines movement among farms and slaughterhouses of Italy: A detailed perspective

1. Introduction

The outbreak of infectious diseases, between animals as well as humans, is an important topic often discussed in recent years. The movement of cows and bulls within a country is crucial to the economics of the cattle industry, nevertheless, is subject to the risk of transmission of infectious disease among other geographical zones. For example, the endemic case of foot-and-mouth disease in the United Kingdom in 2001 brought about an emergency in British agriculture and tourism. This disease that spread quickly among animals brought about 2000 cases of the disease in farms across most of the British rural areas. As a consequence, more than 10 million sheep and cattle were killed in that period.

Generally, outbreaks of diseases like foot-and-mouth disease, avian influenza, and swine influenza have led in the recent decades to a growing interest in modeling of infectious diseases. Because such models can be used to explain the transmission of disease and to measure the impact of various control strategies. In fact, these cattle movement records have provided the farmers and vets with precious information on farm disease outbreak and, generally, have helped to overall decreasing disease outbreaks. In addition, these records provide information about the inter-holding trades and their geographical coordinates.

In the wake of the crisis over mad cow disease (or bovine spongiform encephalopathy), the European Union (EU) has adopted new provisions concerning the identification of bovines and the labeling of their meat[66]. The new provisions improve traceability and food security throughout the sector. They enable European consumer confidence to be strengthened and the creation of favorable conditions for breeding bovines for beef and veal production. To ensure traceability of livestock, the European Union requires registration of all movements of cattle and pigs. In addition to traceability in case of an explosion of contagious diseases, the national databases give us an opportunity to analyze the overall structure of the cattle movements, which can be applicable in economic and epidemiological studies. For instance, we can use them for outbreak contingency planning, the design of control and prohibitive programs for epidemic diseases, and risk-based supervision.

One method to look into cattle movements is through network analysis. Network analysis is a prevailing and useful way to make systems of connections visible and it is applicable in many different disciplines, for example in biology, ecology, economics, epidemiology, and sociology, (Albert & Barabási 2002; Danon et al. 2011). In fact, Network analysis is an interdisciplinary research field which origins in both social sciences, through social network analysis, and in the mathematical graph theory and it deals with the relations between nodes and how they are linked to each other. This kind of analysis has come out as a useful technique to explain many real-world phenomena and has also been supported by the development of Geographical Information Systems, which provide a clear interpretation of spatial data (Pfeiffer, 2004). Nowadays, the researchers can use network theory, and some specific software advancements to investigate the complex and large networks.

[66] Identification and labelling of beef and veal:http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV:l12064

Related to livestock disease outbreaks, network analyses have been used to estimate the potential size of epidemics, either directly through descriptive network measurements (Kao et al., 2006) or through modeling outbreaks based on results of network analyses (Dubéet al., 2009). There are several different methods to approach animal movement networks analytically. Often the holdings are treated as 'nodes' and movement of animals as 'links' or contacts. The networks can be analyzed as static and undirected. This means that there is either a link between two holdings, or there is not, and the direction and the frequency of the movements are not considered. Alternatively, the direction can be included, which is often of interest for disease spread (Dubé et al., 2009). Furthermore, the transitory aspect of the movements can be included, and then create dynamic networks by analyzing movements time-period by time-period (Dubé et al., 2008a; Vernon and Keeling, 2009).

One important point is that the outbreak of diseases can change significantly the network structure. Also, it may lead to different policy implications which caused by including exogenous (that is, driven by factors affecting from outside on the trade patterns of premises) and endogenous (that is, based on existing trade patterns among different holdings) factors of network structure. For instance, Marsh and Smith (2000) developed a policy network model and pointed out that although the distinction between endogenous and exogenous factors is not easy to understand, however, the endogenous effects give rise to wider inequalities within societies (networks). An another example, Tomasello et al. (2014) found that endogenous elements are more important than exogenous ones in analyzing R&D network formations. In the case of an outbreak of diseases, this can be considered an endogenous factor, because it may lead to changes in the trade patterns of different holdings and even deletion (sending cattle to the slaughterhouses instead of sending to another farm) and addition (finding new trade partners) of some links. This phenomenon shows that the network may react endogenously to the outbreak of diseases which is very important in analyzing network structure. In previous chapters, we actually discussed some exogenous and endogenous factors that affect the movement of cattle among some Italian provinces.

In this chapter, we are going to assess the endogenous effects of the diseases on the network of bovine movements in Italy. We will see the disease effects on the movements of bovines from farms to another farms and slaughterhouses. Also, another finding of this study is that the disease effects are not the same for all type of movements. That is, the farmers tend to send their sick bovines to a farther distance than closer holdings. This shows somehow an opportunistic behavior which may be adopted by some farms. In addition, in network analysis of the movements, we will see that the outbreak of diseases can cause some changes in the network structure of the movements.

In sum, in this chapter, we are going to have a more detailed look at the pattern of bovines movements among all Italian farms and slaughterhouses. Thus, in part two, we present the complete dataset and the way that we create the relevant variables. Then, in the third part, we will analyze the overall pattern of bovines movement data and the effect of diseases on this movement patterns. Then, in part 4, we will present a general network analysis (degree effects) of the bovine movement across Italy. Finally, we conclude chapter in the fifth part.

2. Material

2.1 Movement data

Data on bovine trade movements were obtained from the Italian National Bovine database, which is administered by the Italian National Animal Identification and Registration Database instituted by Ministry of Health at CSN of Istituto G.Caporale of Teramo. The database contains the movements of the all Italian population of Bovines between animal premises during a time interval 2006 to 2015, providing an extensive image of the locations that Bovines have been kept and moved within the country. Additional information was provided for the animal holdings, including the type of holdings (i.e. animal husbandry or farm (AL[67]), Genetic Material Center (CG[68]), Collection Center (CR[69]), Fair/Market (FM[70]), Pasture (PA[71]), Point Park (PS[72]), Animal House (ST[73]), Stable (SS[74]) and slaughterhouse (SM[75])), and their geographic coordinates and also the geographical locations of the municipality where the holdings were located. Each movement case shows the unique identification for the animal, the codes of the holdings of origin and destination nodes, and the date of the movement. Such tracking system allows us to easily follow the path of each bovine and to make the equivalent overall network, and then we can minimize the problems with respect to data accuracy that are found in other tracking systems that do not provide both origin and destination of the movements (Volkova et al. 2010, Christley et al. 2005). Also, we have identification code for the owners and keepers of each node.

Data on the outbreaks of diseases among cattle are derived from the SIMAN project done by the Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise "Giuseppe Caporale" which is a public health institute with administrative and managerial autonomy. SIMAN is a national system for the notification and management of outbreaks of animal diseases in Italy. Actually, SIMAN is part of the electronic-government process that includes all public administrations of the

[67] ALLEVAMENTO
[68] CENTRO MATERIALE GENETICO
[69] STABILIMENTO DI MACELLAZIONE
[70] FIERA/MERCATO
[71] PASCOLO
[72] PUNTO DI SOSTA
[73] STABULARIO
[74] STALLA DI SOSTA
[75] STABILIMENTO DI MACELLAZIONE

European Union (EU) and relates to the exploitation of data, information and communication technologies for the digital processing of documents in a manner that makes it easier to do administrative affairs on the internet (Colangeli et.al. 2011). The dataset consists of the name of the diseases and health status of bovines in different areas at different point of the time from 2007 to 2013.

Here we are going to examine the movements of bovines from farms to other farms and slaughterhouses over a period of 7 years (by quarter: 28 Time period). It includes 313748 distinct farms and 3602 distinct slaughterhouses. More specifically, we have data on the date of movements and also a number of the bovines which have been transferred between two holdings, that is, from farms to farms (AL to AL) and from farms to slaughterhouses (AL to SM). Then, for any receiver-sender (origin-destination) pair, we tried to collapse dataset in an equal period of the time (quarter) and then aggregate pair-wise movements by quarter. But since we do not have any bovine movement between some nodes in some quarters, we decide to expand the data and fill out (create) the required missing observations. This gives rise to the creation of zero values for the variable "Total number of traded heads" at the time periods which there is not any trade among two given nodes. Also, there might be some changes due to the size/stock of the origin nodes, so, we decided to control the results for stock of origin nodes.

One point that we should note is that the way that we are cleaning the dataset. In fact, the imputed zeros (resulted from the creation of zeros for the nodes which are not connected in some periods), can help us to decrease the biasedness of the results. Because missing data can bring about problems for analyzing data, and so, this imputation process can avoid some of this problems. In other words, when one or more observations are missing for a variable, most statistical packages such as stata ignore any observation of other variables and this might lead to a bias in respective results. Of course, we should note that the number of zeros can be problematic especially when we are using the log-version of the dependent variable “Total number of movements”. Although, we saw in previous chapters that there are some remedies to this problem[76]. Here in this chapter, we use the ordinary least square (OLS) method and check the results with the nonlinear (logit) model, as well.

Also, concerning disease status data, again we collapsed dataset in quarterly-based periods and then aggregated them in each quarter. The infectious diseases status capture through “positive test” records and in the case of our dataset, the disease status in previous period represents by “lagged_disease_status”. The disease first being discovered in a place in a time period and then transmit to other nodes in different areas by the end of the study. Table 1 summarizes some basic properties of the dataset.

Table 1: summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Origin node[77]	10,777,366	736448.9	238437.6	556843	2863695
Destination node[78]	10,777,366	272436.8	419063.5	1	2976688
Origin latitude[79]	10,775,895	43.51578	2.508136	35	50.999
Origin longitude[80]	10,769,221	11.47625	2.73657	6.704167	18.48667
Destination latitude[81]	3,619,536	43.91805	2.262545	35	50.999
Destination longitude[82]	3,613,909	10.93766	2.569462	6.704167	18.48667
Total number of traded heads[83]	10,777,366	3.198363	21.87625	0	4371

[76] Refer to chapter 1, section 3, to see how we can deal with the problem of excessive zeros.
[77] Identification code of origin nodes
[78] Identification code of destination nodes
[79] Latitudinal coordinates of origin nodes
[80] Longitudinal coordinates of origin nodes
[81] Latitudinal coordinates of destination nodes
[82] Longitudinal coordinates of destination nodes
[83] Number of heads which are transferred between holdings in a specific time period

3. Results and discussion

3.1 Contagion effect

The most common way of contagion of a disease through different places is to "buy in" the disease by mistakenly purchasing one or more infected, but apparently healthy, animals to add to existing cattle. The disease would then spread through the cattle. To examine this effect, we considered the effect of disease on the number of heads transferred ("total_traded_heads") from farms to other farms (i.e. AL to AL) and from farms to slaughterhouses (i.e. AL to SM). To do so, we regressed the sickness variables in the previous time-period on the number of heads that has been moved in the current time period. To clarify, variable "lagged_disease_status" is a binary variable that denotes the lagged disease status and it takes one whenever there are a positive disease test and zeroes otherwise. Also, we should note that here we use the linear estimation method. Because in nonlinear estimation, we will face with incidental parameter problem. The reason behind it relies on the nature of our dataset. This is due to the fact that fixed effects estimators of nonlinear panel data models can be highly biased because of this incidental parameter problem. For example in probit/logit models, estimators are consistent. This means that that as the ratio of the number of observations to a number of parameters increases, the parameter estimates will converge toward their true values as standard errors become arbitrarily small. The problem which appears in the case of fixed effects is that the number of parameters becomes larger with the number of observations. Therefore, the parameter estimates can never converge to their true values as the sample size increases. Thus the parameter estimates are seriously unable to be trusted. So, it's better to use the linear fixed effect estimation method. However, since the OLS method in censored data may subject to downward bias, we check the results with non-linear estimation method like fixed-effects logit.

Therefore, here we are interested in predicting the effect of the disease in holdings on the movement of bovines among them. To do so, we should first consider "total_traded_heads" as an outcome variable. In addition, since we are going to analyze bovines movement among origin-destination unique links, we use fixed effect (origin-destination links) estimation with robust standard errors. Table 2 reports results for trade from farms to farms.

Since we are going to see the effect of diseases on the pattern of movements in the current period, we should take into account the disease status of the holdings in previous time period (lagged). So, Column (1) presents estimates in case of disease test in the previous time period (lagged). Also, there might be some changes due to the size/stock of the origin nodes, so, we decided to control it by reestimating models and using "heads over origin stock" as the dependent variable. The respective results are shown in column (2).

As we see in column (1) lagged disease (positive) test variables have a negative and significant effect in the case of bovine movements from farms to another farm (i.e. AL to AL). Also, column (2) shows us that by controlling the origin node size/stocks, the effect of lagged disease test on the movement of bovines remains unchanged.

In addition, as we said above to check the results, we used the nonlinear logit estimation method with a new dependent variable which shows that if there have been any bovines movement between holdings in a time-period or not. We call this binary variable as "transferijt" which takes 1 if there is movement and 0 otherwise. Column (3) confirms the results estimated by linear estimation method and shows that the previous results still hold.

More specifically, to interpret numerically, table 2 shows that at the presence of a positive disease test in a current and previous time period can lead to decrease the number of heads which has been transferred between two farms up to about 2 bovines. We may expect that this reduction in the number of transferred bovines gives rise to an equivalent increase in the number of the movement of heads from farms to slaughterhouses. we will show it in the next table.

Similarly, we can see the effect of disease test (and therefore the existence of sick bovines in a farm) on the movement of bovines from farms to slaughterhouses. Table 3 presents the related estimates. Likewise, column (2) in table 3 shows the estimates when we control for origin node stocks (that is, farms' size). As we see in table 3, the positive diagnosis of diseases on a farm has a positive and significant effect on the movement of bovines from farms to slaughterhouses. This is aligned with our expected conjecture that the generally farms tend to send sick bovines to slaughterhouses and this can hinder the outbreak of diseases among different holdings.

Based on our results in Table 3, the average number of heads that farms tend to send to the slaughterhouses, in the case of positive disease test records, is about 2 heads which are compatible with the results that we had in table 2. Also, as we see there is not much difference in the number of heads which transfer among various holdings (2 heads in average), but we should note that even one sick bovine can cause an outbreak and so it can have an influential effect on the spread of diseases among different locations.

Thus, in general, tables 2 and 3 show that in the case of positive disease test and sick bovines, farms tend to send their bovines to the slaughterhouses than other farms. It means that we will face with an increase in a number of bovines going from farms to the slaughterhouse BUT a decrease in the flows from farms to the other farms. In the other words, farmers are prone to send their bovines to slaughterhouses when there are positive disease tests and similarly, they do not want to send their probable sick bovines to another farm, avoiding the spread of diseases among bovines.

Table 2: contagious bovine disease (movement between farms (AL to AL))

VARIABLES	(1) Dependent Variable: total_traded_heads	(2) Dependent Variable: heads over origin stock	(3) Dependent Variable: transferijt
Number of origin node's bovines	0.0148***	-0.000403***	0.00137***
	(0.00245)	(6.91e-05)	(6.85e-05)
Lagged_disease_status[84]	-1.491***	-0.0215***	-1.041***
	(0.359)	(0.00364)	(0.106)
2009.year	0.0934***	0.00491***	0.126***
	(0.0269)	(0.00136)	(0.00564)
2010.year	0.0564**	0.00565***	0.146***
	(0.0284)	(0.000922)	(0.00613)
2011.year	0.140***	0.0106***	0.264***
	(0.0324)	(0.00119)	(0.00649)
2012.year	0.244***	0.0176***	0.477***
	(0.0372)	(0.00351)	(0.00695)
2013.year	0.870***	0.0294***	1.346***
	(0.0482)	(0.00144)	(0.00806)
2.quarter	-0.0119	0.00447***	0.00866**
	(0.0214)	(0.00164)	(0.00430)
3.quarter	0.0611**	0.00228**	0.00582
	(0.0238)	(0.00116)	(0.00441)
4.quarter	0.431***	0.0102***	0.403***
	(0.0235)	(0.000625)	(0.00438)
Constant	0.267	0.0961***	-
	(0.319)	(0.00845)	
Observations	2,880,066	2,873,541	2,248,746
R-squared	0.682	0.684	-
Number of unique_link	-	-	210,101

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: columns (1) and (2) are estimated by OLS fixed-effects (origin-destination pairs) estimation method and column (3) is estimated by logit fixed-effect (origin-destination pairs) method. Regressions control for stock.

[84] Disease test status in the previous period.

Table 3: contagious bovine disease (movement of farms to slaughterhouses (AL to SM))

VARIABLES	(1) Dependent Variable: total_traded_heads	(2) Dependent Variable: heads over origin stock	(3) Dependent Variable: transferijt
Number of origin node's bovines	0.0154***	-0.000342***	0.000863***
	(0.00182)	(1.63e-05)	(3.79e-05)
Lagged_disease_status	1.157***	0.0409***	0.477***
	(0.113)	(0.00439)	(0.0280)
2009.year	0.180***	0.0153***	0.0932***
	(0.0188)	(0.000377)	(0.00348)
2010.year	0.262***	0.0182***	0.136***
	(0.0208)	(0.000365)	(0.00375)
2011.year	0.405***	0.0209***	0.158***
	(0.0224)	(0.000395)	(0.00391)
2012.year	0.648***	0.0296***	0.317***
	(0.0257)	(0.000474)	(0.00414)
2013.year	1.330***	0.0600***	1.087***
	(0.0352)	(0.000561)	(0.00483)
2.quarter	0.0466***	0.00626***	0.0788***
	(0.0159)	(0.000309)	(0.00284)
3.quarter	0.249***	-0.00525***	-0.0267***
	(0.0166)	(0.000326)	(0.00292)
4.quarter	0.505***	0.0237***	0.405***
	(0.0170)	(0.000332)	(0.00292)
Constant	1.464***	0.0963***	-
	(0.168)	(0.00151)	
Observations	5,409,901	5,353,624	4,827,264
R-squared	0.748	0.373	-
Number of unique_link	-	-	373,650

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: columns (1) and (2) are estimated by OLS fixed-effects (origin-destination pairs) estimation method and column (3) is estimated by logit fixed-effect (origin-destination pairs) method. Regressions control for stock.

In addition, we can check the reliability of our results by looking at the general effect of diseases at the presence and absence of any trade between two different holdings. In this respect, we try to regress the disease status on the binary variable "transferijt" which we explained before. As we said, there are almost two methods for estimating this regression when dealing with binary dependent variables. Linear Probability Models (LPM) and nonlinear probability models (Logit/Probit). The choice between this two models mostly depends on the nature of the data and the purpose of the study (Hippel 2015). Hippel (2015) mentions that there are so many situations that the linear models perform better than non-linear models. In fact, it could be said that the main advantage of linear models is ease of their interpretability. Because in Logistic models we are facing with the log of odds ratio and so the intuition of this term is not so straightforward. Of course, we should note the non-linear models are unavoidable if they fit better on our data. However, in many situations, the linear models fit better than logistic models. Also, Hellevik (2007) shows that on many occasions both linear and nonlinear models give rise to same results that are hardly recognizable except that the non-linear models are difficult to interpret. In this regard, Hippel (2015) proposes a rule of thumb for choosing the proper method between this two models. He states that if the probabilities that we are going to model are very close to 0 and 1, it's better to use Logistic regression and if the probabilities lie almost between 0.2 and 0.8.[85]

Besides the interpretability of the results, another advantage of the linear model which is so essential in our case is the computation speed. Fitting a non-linear model is intrinsically slow because it uses an iterative way of maximum likelihood. This problem is not important when we are dealing with small or medium size datasets. But since our dataset consists of the detailed movements of bovines between many farms (more than three hundred thousands farms), the size of the dataset is so huge and it leads us to use Linear Probability Model.

Therefore, here we use LPM to see the disease effects on the bovines movement (table 4). As we said, In this case, the dependent variable "transferijt" is a binary variable that takes one if there

[85] For more details, refer to: http://statisticalhorizons.com/linear-vs-logistic

has been any movement between two holding i and j in time period t, and 0 otherwise. Therefore, in the case of sickness, we would expect to have a higher probability of trade between farms and slaughterhouse and lower probability of trade between farms and other farms.

Table 4: linear probability model (LPM)

VARIABLES	(1) Dependent Variable: transferijtaltoal	(2) Dependent Variable: transferijtaltosm
Number of origin node’s bovines	0.000251***	0.000121***
	(1.73e-05)	(7.63e-06)
Lagged_disease_status	-0.134***	0.0920***
	(0.0114)	(0.00576)
2009.year	0.0196***	0.0161***
	(0.00109)	(0.000661)
2010.year	0.0208***	0.0228***
	(0.00118)	(0.000710)
2011.year	0.0389***	0.0256***
	(0.00128)	(0.000748)
2012.year	0.0738***	0.0539***
	(0.00143)	(0.000809)
2013.year	0.250***	0.196***
	(0.00181)	(0.000951)
2015.year	0.248***	-
	(0.0941)	
2.quarter	0.000429	0.0141***
	(0.000866)	(0.000547)
3.quarter	-0.000834	-0.00603***
	(0.000877)	(0.000550)
4.quarter	0.0731***	0.0747***
	(0.000915)	(0.000563)
Constant	0.355***	0.398***
	(0.00249)	(0.000937)
Observations	2,880,066	5,409,901
R-squared	0.461	0.343

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Results are estimated by the fixed effects (origin-destination pairs) Linear Probability Model. Regressions control for stock.

Therefore, as we see, the results of LPM are consistent with tables 2 and 3, as well. More specifically, the numerical interpretation of table 4 shows that the at the presence of a positive disease test in a holding decrease the probability of the bovine trade among two farms by 13.4% and increase the probability of sending bovines from farms to slaughterhouses by about 9.2%.

All of these results strengthen the idea that the network changes to reduce the risk of epidemics, that is the positive disease test in a time period, will have a negative effect on the movement of bovines among farms and a positive effect on the flow of bovines from farms to slaughterhouses.

One important point that we saw worth noting in our database is that the owner and keepers of some holdings are the same. It means that for some movement we probably will not see the effects that we had above. Therefore, to investigate the effects of holdings with the same owner and keeper on the movement of bovines, we ran again our specifications for nodes with same owner and keeper and different owner and keeper. Obviously, here we have the trade of bovines from farm to farm. The results are shown in table 5.

As we see in table 5, there are not any significant effects of disease test on the movement of bovines between nodes with the same keeper and same owner (second and fourth columns), because two farms with the same owner/keeper are subjected to the same information and the bovine movement behaviors are not affected by some external agents. While we can see in the third and fifth columns that the presence of a disease affects on the trade of bovines among farms with different keepers and different owners. Therefore, this point shows us that the previous results are valid only if the origin and destination nodes have different owner or keeper.

Table 5: Same and different owners and keepers (farm (AL) to farm (AL))

VARIABLES	Dependent variable: total_traded_heads origin and destination nodes have the same keeper	origin and destination nodes have a different keeper	origin and destination nodes have the same owner	origin and destination nodes have a different owner
Number of origin node's bovines	0.105***	0.034***	0.021	0.041***
	(0.019)	(0.003)	(0.016)	(0.003)
lagged_disease_status	-0.430	-1.066***	-2.550	-1.058***
	(0.829)	(0.138)	(1.962)	(0.138)
2009.year	0.310	0.289***	0.553	0.278***
	(0.623)	(0.0445)	(0.465)	(0.0449)
2010.year	0.0607	0.538***	0.721	0.520***
	(0.726)	(0.0383)	(0.536)	(0.0388)
2011.year	1.485**	0.984***	1.559***	0.977***
	(0.735)	(0.0450)	(0.528)	(0.0457)
2012.year	2.530***	0.942***	2.813***	0.927***
	(0.772)	(0.0525)	(0.567)	(0.0533)
2013.year	7.526***	2.448***	5.056***	2.499***
	(1.007)	(0.0709)	(0.725)	(0.0729)
2.quarter	-0.800	0.125***	-0.260	0.117***
	(0.515)	(0.0339)	(0.363)	(0.0343)
3.quarter	-0.470	0.458***	-1.215***	0.479***
	(0.561)	(0.0441)	(0.404)	(0.0446)
4.quarter	3.100***	0.745***	2.170***	0.759***
	(0.566)	(0.0391)	(0.405)	(0.0396)
Constant	15.66***	1.821***	14.13***	1.814***
	(0.616)	(0.0356)	(0.444)	(0.0360)
Observations	41,859	2,387,795	49,066	2,380,588
R-squared	0.559	0.711	0.687	0.708

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Results are estimated by OLS fixed-effects (origin-destination pairs) estimation method. Regressions control for stock.

Moreover, as we said, we created zero value observations for the situations that we don't have any movement of bovines and so, the dependent variable is only observable in case of presence of the movement and we can not observe the effects of explanatory variables in certain conditions (i.e. absence of any bovine movement). Then, we doubted that our linear estimates might be biased. To address the robustness and reliability of our results, we decided to check them again with censored regression model which normally happens in cases where the dependent variable is only observable in some special cases. As we know, the model which usually uses in case of censored data is the tobit model. Of course, The problem that we are facing with tobit is that it was not possible to run fixed-effects estimates and so we couldn't see the desired effects within groups. However, we also used tobit to estimate the effect of diseases on the movement of bovine from different holdings and finally, we found that the tobit results confirm the general effects of the diseases on the movement of bovines. the respective tobit estimates are reported in the appendix.[86]

3.2 Distance effects

In this part, we would like to assess the effects of distance among holdings on the movement of bovines. Here we have data on the geographical coordinates of the holdings and so, we can use "vincenty" module in stata to compute the respective distances among the holdings. Also, since the data on the geographical coordinates of slaughterhouses are not available, so, the following result and the table of estimation are associated with the bovine movements from farms to other farms (AL to AL).

[86] Since our dependent variables like "total_traded_heads" cannot take negative values and so it is a below-censored or left-censored variable, we use the censored regression model like tobit to estimate the effect of diseases on the movement of bovine from different holdings. To do so, we should first consider total_traded_heads as an outcome variable. Then, to generate a tobit model in Stata, we list the outcome variable followed by the regressors and then specify the lower limit of the outcome variable. The lower limit is specified in parentheses after ll (Lower Limit).

Generally, we would expect that the farms with large distances would not eager to trade with each other. In other words, the farther are the nodes, less likely having trade together. This can be due to the fact that movement among distant nodes needs more transaction cost such as transportation cost etc. Nevertheless, this effect may not be the case when we are facing with sick bovines in different nodes at various distances. Therefore, we regress both distance variable and interaction of distance and disease test variables on the total number of traded heads. The results are shown in table 6.

As we see, in the case of diseases in farms, the signs are not as anticipated as before and distance has a positive and significant effect on the movement of bovines between different. It means that the farms tend to send their bovines to farther farms and slaughterhouses. It could be because the close locations are subject to the same veterinarians or organization for detecting the sickness of bovines and therefore they know about each other situation, while the far nodes have not this kind of information and so, it would be more probable to trade with each other.

Also, the numerical results of the table (6) show that in the case of positive disease test and so, the existence of a disease in a place, one unit increase in distance can lead to 0.0004 increase in the number of heads that can be transferred from one place to another. More specifically, 10000 unit increase in distance (which means 10 km) can increase the amount of traded bovines go up to about 4 heads.

Table 6: distance effects (transfer from farm to farm)

VARIABLES	(1) Dependent Variable:transferijt
Number of origin node's bovines	0.00001***
	(2.96e-06)
lagged_disease_status	-0.255***
	(0.0126)
Distance_km	0.00037***
	(6.93e-06)
Lagged_disease_status*distance_km	0.00041***
	(0.000131)
2008.year	-0.0561***
	(0.00116)
2009.year	-0.0651***
	(0.00111)
2010.year	-0.0752***
	(0.00106)
2011.year	-0.0502***
	(0.000981)
2013.year	0.202***
	(0.00104)
2.quarter	-0.00631***
	(0.000764)
3.quarter	-0.0220***
	(0.000787)
4.quarter	0.0676***
	(0.000758)
Constant	0.460***
	(0.00107)
Observations	2,817,396
Number of unique_link	770,557

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Results are estimated by OLS fixed-effects (origin-destination pairs) estimation method. Also, the dependent variable is "transferijt" which shows that the presence and the absence of movement between two places of i and j at time period t. It takes 1 in the case of movements, and 0 otherwise. Regressions control for stock.

In addition, table 6 reveals somehow an opportunistic attitude by some farmers, in the sense that they tend to get rid of their sick bovines at any cost and even at the expense of the transferring to other farms. This is in comparison with the result of gravity model that we saw in previous chapters which distance had a negative effect on the movements and holdings, in general, tend to a conservative behavior toward their business. But, here in case of the presence of diseases, the farms are acting in a way that they tend to send their sick bovines to farther distances. This shows the kind of behavior that can even exploit variability in movements between farms that have been already in touch.[87]

4. Network analysis

Network analysis provides a useful approach to understanding animal movement patterns as it reveals quantitative aspects of each individual's position and relationship pattern.

Social networks are generally defined as graphs that allow their representation and mathematical analysis. A graph is a set of objects called nodes connected to each other by links called edges. A set of nodes linked to each other recognizes as a component. Two components of a graph are separated if both of them have no node in common. The number of nodes linked to one node is defined as its "degree" measure. The length of a path is the number of edges in it. Two nodes are unreachable if they belong to different components (there is no path that links them).

[87] The "*opportunism"* assumption is explained by Williamson (1985) as *"self-interest seeking with guile"*. It shows the possibility that people may act in a self-interested way with guile. Williamson says that this is due to the fact that people are not "perfectly rational" which he derived from of the another assumption (bounded rationality). In fact, the assumption of opportunism is also connected to people's behavior in the sense that they may not be completely reliable and honest about their purposes. Of course, he concludes that not always all people will act in an opportunistic behavior. Actually, williamson simply assumes that someone may and also has a room to behave in an opportunistic manner at some times and at the same time it is nor easy to understand who is acting opportnistically and who is not in advance. Results are also controlled for stock.

The diameter of a graph is the length of the longest direct path between two nodes in a connected component. The characteristic path length of a network is the shortest path length between two nodes averaged over all pairs of nodes in the network. A graph can be directed if the edge can be followed only in one direction, from x to y. In this case, edges are called arcs. A "loop" is an edge with both ends in the same node.

In our case, we focused on the recorded movements, formalized as a cattle contact network, in which holdings were the nodes in the graph representing a network through which animals were moved. It was assumed that edges were directed (arcs) and there were no loops (a holding could not move animals to itself). An arc was assumed to exist between two holdings (origin and destination) if there was at least one movement recorded between them.

Measures such as degree, closeness and betweenness centrality and density provide insight into various aspects of a network. The network feature for a node can be captured through using the concept of degree centrality. The node that has the most direct links is the most active node in the network and can be called a "hub". Identifying hubs may also reveal the critical nodes of the network, i.e., the most important points for disease outbreak and control.

The first thing is that how we should represent our network. There is much complexity in this kind of contagious networks. So, we have to simplify them a bit. Generally, the global pattern of networks is captured through degree distributions, path lengths, etc. Specifically, in this part, we see the effect of diseases on the network structure of bovines movement through using the commonly discussed characteristics of any network, i.e. indegree, outdegree, and degree.[88]

The degree is a measure of the number of contacts to and from a specific holding. There is no specific direction in the definition of degree, while direction is taken into account, the ingoing and outgoing contacts can be separated; (1) the out-degree is the number of contacts with direction from the premise, thus the number of premises receiving animals directly from a specific premise; (2) the in-degree is the number of contacts with direction to the premise, thus the number of premises from which each premise receive animals directly (Wasserman et al., 1994). The static network was decomposed in seasonal time periods to investigate the pattern and behavior of the network.

The fact that indegree could be important is that we can trace back the diseases by looking at the effect of indegree in the case of the presence of a disease in a given node. On the other side, outdegree characteristic is important because it can shows the extent that nodes are potentially

---

[88] Betweenness centrality is another indicator of a node's centrality in a network. It is equal to the number of shortest paths from all vertices to all others that pass through that node. If a node has been had a high betweenness centrality it means that it has a significant influence on the transfer of bovines through the network, of course, provided to this assumption that bovine movement follows the shortest paths. A node of relatively low degree centrality may play an important role on the network connectivity and so have a high betweenness centrality. Hubs are nodes with high degree and betweenness centrality. Closeness centrality shows the nodes that have the shortest paths to all others, indicating how fast a node can reach to other nodes in the network. Network centrality can reveal much about the overall network structure. One or a few central nodes dominate a very centralized network. If these nodes are removed, the network quickly breaks into unconnected networks. For example, livestock markets are expected to be highly connected to the rest of the network. For this reason, for major disease threats, livestock markets might be immediately closed to slow down infection spread, restricting disease flow through the network. Also, we can think about another property which is important in capturing networks and especially is one which is looking at a local property of the networks. Therefore, in particular, what is going on when we zoom in on given nodes and begin to understand the relationship between different links in the network? This is known as clustering. And in particular, when we begin to think about asking how dense is a network at a local level, we could ask a question of what fraction of the farms which are trading with a given farm, are trading with each other? Therefore, clustering looks at if we have a given node I, and we look at two of I's trade-partner J and K, what is the chance that those two are related to each other. However, since we assume that the decision of farms to whom trade with is independent of their partners, so, this property is not so relevant to our network.

dangerous can spread the disease among holdings. That is these high outdegree nodes can play as a source of the risk for the disease outbreaks.

One important point that we have to notice is that there should be some difference in the coefficients when we are dealing with the ratios rather than level values. In the sense that the number of trading partners of each farm (origin node) change during time and so for example by adding a new trade partner, the overall number of the traded bovines may increase but, at the same time, the number of the trade bovines for another (previous) trading partners may decrease, because the new partner can attract some market share of each node. Therefore, since we are facing with network characteristics of each node, the respective estimated coefficients would be probably sensitive to the use of different dependent variables: "total_traded_heads" and "traded heads over origin stock/size". This status would be more influential especially when we are dealing with the outdegree characteristic of origin nodes because, in that respect, the origin nodes could decide to increase/decrease their trading partners. In the following estimations, we try to execute both specifications and assess the likely differences.

Also, we should note that since we are going to assess the effect of indegree in current movements, so, it's better to use the lagged values of indegree. In this respect, we computed the lagged values of indegree, outdegree, and degree. In addition, we use the number of origin node's bovines to control for the size of the farms. The results of the effects of origin node's indegrees are shown in tables 7 and 8.

As we see in table 7, there is a positive and significant effect of the interaction of indegree and disease status on the movement of bovines from farms to farms. However, by looking at columns 2 and 4 of table 8, we can see that this interaction term is not significant anymore in case of movement from farm to slaughterhouses. It means that the farms that have received more (probably sick) bovines from more suppliers in the previous period, tend to send more bovines to other farms in the current time period. In fact, when there is a positive disease test and at the same time the number of suppliers of a given farm is high, it would be more probable for the bovines of given farm to be contracted by diseases. Therefore, farmers tend to send more bovines to other farms to somehow decrease the risk of an outbreak of diseases in their heads.

Also, as we see from tables 7 and 8, which interaction of indegree and disease status has a significant effect on the bovines movement from farms to farms and insignificant effect on the movements from farms to slaughterhouses. This shows that the kind of behavior which may lead to an outbreak of diseases, in the sense that when a given farm is receiving bovines from more places (i.e. higher indegree), the risk of contracting diseases is also higher and actually the likely source of diseases are higher. This can lead to the transfer of more bovines in the current period to other farms. In other words, this pattern can give rise to the spread of diseases among different farms and locations.

Also, analogous with previous parts, the main effect of disease status (variable "Lagged disease status") is negative/positive on the movements of bovines from farms to farms/slaughterhouses. In addition, to control for origin node size, we added the number of bovines of each origin nodes (column 1 and 2) as a regressor, and Also, we executed again the specifications with "heads over origin stock" as the dependent variable (columns 3 and 4).

Table 7: effects of the indegree of origin nodes on movements from farms to farms (AL to AL)

VARIABLES	(1) Dependent Variable: total_traded_heads	(2) Dependent Variable: total_traded_heads	(3) Dependent Variable: heads over origin stock	(4) Dependent Variable: heads over origin stock	(5) Dependent Variable: heads over origin stock	(6) Dependent Variable: heads over origin stock
Number of origin node's bovines	0.0148***	0.0148***	-	-	-0.000264***	-0.000264***
	(0.00253)	(0.00253)			(5.79e-05)	(5.79e-05)
Lagged disease status	-0.877***	-1.143***	-0.00932***	-0.0138***	-0.00904***	-0.0137***
	(0.272)	(0.347)	(0.00337)	(0.00396)	(0.00341)	(0.00399)
Lagged indegree of origin node	0.00785	0.00785	0.00127	0.00127	0.00136	0.00136
	(0.0144)	(0.0144)	(0.00165)	(0.00165)	(0.00167)	(0.00167)
Interaction of origin lagindegree and disease status	-	0.0229**	-	0.000381**	-	0.000400**
		(0.00947)		(0.000191)		(0.000195)
2009.year	0.273***	0.273***	0.00463***	0.00463***	0.00509***	0.00509***
	(0.0475)	(0.0475)	(0.00122)	(0.00122)	(0.00126)	(0.00126)
2010.year	0.419***	0.419***	0.00856***	0.00856***	0.00937***	0.00937***
	(0.0486)	(0.0486)	(0.00157)	(0.00157)	(0.00168)	(0.00168)
2011.year	0.682***	0.682***	0.0145***	0.0145***	0.0155***	0.0155***
	(0.0587)	(0.0587)	(0.00220)	(0.00220)	(0.00235)	(0.00235)
2012.year	1.002***	1.002***	0.0295***	0.0295***	0.0309***	0.0309***
	(0.0706)	(0.0706)	(0.00999)	(0.00999)	(0.0103)	(0.0103)
2013.year	2.071***	2.071***	0.0396***	0.0396***	0.0412***	0.0412***
	(0.0957)	(0.0957)	(0.00350)	(0.00350)	(0.00373)	(0.00373)
2.quarter	0.0628	0.0628	0.00985**	0.00985**	0.00985**	0.00985**
	(0.0416)	(0.0416)	(0.00425)	(0.00425)	(0.00425)	(0.00425)
3.quarter	0.0660	0.0658	0.000447	0.000445	0.000505	0.000502
	(0.0486)	(0.0486)	(0.000941)	(0.000941)	(0.000946)	(0.000945)
4.quarter	0.678***	0.678***	0.0129***	0.0129***	0.0131***	0.0131***
	(0.0472)	(0.0472)	(0.00104)	(0.00104)	(0.00107)	(0.00107)
Constant	-0.558	-0.558	0.00149	0.00149	0.0403***	0.0403***
	(0.384)	(0.384)	(0.0191)	(0.0191)	(0.0118)	(0.0118)
Observations	851,475	851,475	850,570	850,570	850,570	850,570
R-squared	0.560	0.560	0.098	0.098	0.098	0.098

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Results are estimated by OLS fixed-effects (origin-destination pairs) estimation method. Regressions (i.e. columns 1, 2, 5, and 6) are also controlled by the size of the origin farms Regressions control for stock.

Table 8: effects of the indegree of origin nodes on movements from farms to slaughterhouses (AL to SM)

VARIABLES	(1) Dependent Variable: total_traded_heads	(2) Dependent Variable: total_traded_heads	(3) Dependent Variable: heads over origin stock	(4) Dependent Variable: heads over origin stock	(5) Dependent Variable: heads over origin stock	(6) Dependent Variable: heads over origin stock
Number of origin node's bovines	0.0127***	0.0127***	-	-	-0.000325***	-0.000325***
	(0.00300)	(0.00300)			(2.66e-05)	(2.66e-05)
Lagged disease status	1.183***	1.748***	0.0414***	0.0433***	0.0410***	0.0430***
	(0.328)	(0.499)	(0.00873)	(0.00869)	(0.00869)	(0.00869)
Lagged indegree of origin node	0.00755	0.00755	-0.000357***	-0.000357***	-0.000238***	-0.000238***
	(0.0137)	(0.0137)	(5.17e-05)	(5.17e-05)	(5.23e-05)	(5.23e-05)
Interaction of origin lagindegree and disease status	-	-0.0792	-	-0.000264	-	-0.000279
		(0.0715)		(0.000829)		(0.000829)
2009.year	0.325***	0.325***	0.0146***	0.0146***	0.0148***	0.0148***
	(0.0432)	(0.0432)	(0.000706)	(0.000706)	(0.000708)	(0.000708)
2010.year	0.506***	0.506***	0.0167***	0.0167***	0.0170***	0.0170***
	(0.0485)	(0.0485)	(0.000637)	(0.000637)	(0.000638)	(0.000638)
2011.year	0.783***	0.783***	0.0248***	0.0248***	0.0253***	0.0253***
	(0.0523)	(0.0523)	(0.000718)	(0.000718)	(0.000723)	(0.000723)
2012.year	1.163***	1.163***	0.0356***	0.0356***	0.0360***	0.0360***
	(0.0616)	(0.0616)	(0.000980)	(0.000980)	(0.000997)	(0.000997)
2013.year	2.048***	2.048***	0.0634***	0.0634***	0.0637***	0.0637***
	(0.0905)	(0.0905)	(0.00101)	(0.00101)	(0.00101)	(0.00101)
2.quarter	0.0914**	0.0913**	0.00549***	0.00549***	0.00532***	0.00532***
	(0.0359)	(0.0359)	(0.000533)	(0.000533)	(0.000533)	(0.000533)
3.quarter	0.370***	0.371***	-0.00443***	-0.00443***	-0.00479***	-0.00479***
	(0.0368)	(0.0368)	(0.000667)	(0.000667)	(0.000653)	(0.000653)
4.quarter	0.642***	0.642***	0.0218***	0.0218***	0.0218***	0.0218***
	(0.0372)	(0.0372)	(0.000572)	(0.000572)	(0.000571)	(0.000571)
Constant	2.243***	2.243***	0.0644***	0.0644***	0.0997***	0.0997***
	(0.336)	(0.336)	(0.000695)	(0.000695)	(0.00284)	(0.00284)
Observations	1,827,642	1,827,642	1,813,999	1,813,999	1,813,999	1,813,999
R-squared	0.784	0.784	0.327	0.327	0.328	0.328

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Results are estimated by OLS fixed-effects (origin-destination pairs) estimation method. Regressions (i.e. columns 1, 2, 5, and 6) are also controlled by the size of the origin farms. Regressions control for stock.

The overall results of tables 7 and 8 tell us that the interaction of indegree and disease status of origin nodes has not any significant effect on the movement of bovines only if there is movement from farms to other farms. Actually, it makes sense that on of the reasons for the outbreak of diseases can be due to this phenomena that farms tend to send the more (probably sick) bovines to the other (probably far away) farms, and as a result it can increase somehow the risk of the spreading diseases in different regions. This is compatible with our previous results that showed us some opportunistic behavior adopted by some farmers that existence of a disease in an area (and generally a positive disease status) had positive effects on the movement of bovines from farms to other farms.

On the other side, we can compare the effects of the outdegree characteristic on the movements of bovines between different holdings. One important point that we should take into account is that in each time period outdegree of a node could be correlated with our dependent variable (i.e. total number of traded bovines), in the sense that the more are the total number of bovines, the nodes would probably tend to send their bovines to more trading partner (i.e. more destination node). Therefore, it would be better to use the “lagged” version of outdegree variable. Also, as we said before, total number of traded bovine cannot provide us a good impression of the outdegree effects on the movement of bovines between different holdings, because changing the pattern of trade partners of each farm may change the market share of each destination point and so it would be better to control our results by the stock/size of origin node. Accordingly, here we try to execute the specifications both in level variable (total_traded_heads) and market share variable (heads over origin stock) and interpret each one separately. The results of “origin-node outdegree” effects are shown in tables 9 and 10 below.

As we can see in column (1) and (2) of table 9, the main (pure) effect of the variable “Lagged outdegree of origin node” is negative and significant on the total movements of bovines from farms to farm and not a significant effect on heads over stock. That is the probability of lagged-outdegree is correlated with the number of bovines sent in the previous period. In other words, this means that the more a farm has sent its bovines to other places in the previous period, it would be more likely to send fewer bovines in the current time period. One probable explanation could be due to the fact that a farm has already supplied most of its heads and so, there are not many heads available.

In addition, columns (2) and (4) of table 9 show that the interaction effect of outdegree and disease status is highly significant and positive in case of movement from farms to other farms. This tells us that in the case of diseases, the farms tend to not change significantly the pattern of their trading partners (other farms), and probably they tend to send their (probably sick) bovines to the other farms and this situation can even decrease the number of heads which transfer from farms to slaughterhouses. this confirms the importance of such nodes in the network in the sense that those with higher outdegree can be considered as most dangerous nodes can act as a source of diseases. by looking at column (2) of table 10, we can see this decreasing effect in the number of traded heads from farms to slaughterhouses. (Note: we know that the indegree and outdegree are correlated (i.e. Covariance (indegree,outdegree)>0). In this regard, we also tried to run specifications with indegree and outdegree and their interaction with diseases altogether. We found that in this case, outdegree is the only variable which survives and so it can be considered the dominant factor of the model)

This outdegree result is also in line with the previous parts that the farms may act in a way that sends bovines to other farms to avoid contracting diseases in their heads. This behavior can be one of the reasons for the spreading disease among different areas.

Table 9: origin outdegree effects (movement among farms i.e. AL to AL)

VARIABLES	(1) Dependent Variable: total_traded_heads	(2) Dependent Variable: total_traded_heads	(3) Dependent Variable: heads over origin stock	(4) Dependent Variable: heads over origin stock	(5) Dependent Variable: heads over origin stock	(6) Dependent Variable: heads over origin stock
Number of origin node's bovines	0.0110***	0.0110***	-	-	-0.000406***	-0.000406***
	(0.00220)	(0.00220)			(7.16e-05)	(7.16e-05)
Lagged disease status	-0.892***	-1.204***	-0.0156***	-0.0218***	-0.0153***	-0.0216***
	(0.142)	(0.173)	(0.00280)	(0.00384)	(0.00279)	(0.00383)
Lagged outdegree of origin node	-0.00445**	-0.00445**	0.000431	0.000431	0.000545	0.000545
	(0.00227)	(0.00227)	(0.000618)	(0.000618)	(0.000624)	(0.000624)
Interaction of origin lagoutdegree and disease status	-	0.0466***	-	0.000932***	-	0.000943***
		(0.0112)		(0.000189)		(0.000189)
2009.year	0.285***	0.285***	0.00732***	0.00732***	0.00785***	0.00785***
	(0.0233)	(0.0233)	(0.00102)	(0.00102)	(0.00106)	(0.00106)
2010.year	0.399***	0.399***	0.0104***	0.0104***	0.0115***	0.0115***
	(0.0244)	(0.0244)	(0.000780)	(0.000780)	(0.000780)	(0.000780)
2011.year	0.605***	0.605***	0.0173***	0.0173***	0.0188***	0.0188***
	(0.0284)	(0.0284)	(0.000747)	(0.000747)	(0.000850)	(0.000850)
2012.year	0.873***	0.873***	0.0270***	0.0270***	0.0292***	0.0292***
	(0.0325)	(0.0325)	(0.00295)	(0.00295)	(0.00310)	(0.00310)
2013.year	1.761***	1.761***	0.0430***	0.0430***	0.0458***	0.0458***
	(0.0430)	(0.0430)	(0.00112)	(0.00112)	(0.00137)	(0.00137)
2.quarter	0.0436**	0.0436**	0.00595***	0.00596***	0.00596***	0.00597***
	(0.0187)	(0.0187)	(0.00157)	(0.00157)	(0.00157)	(0.00157)
3.quarter	0.155***	0.155***	0.00429***	0.00429***	0.00444***	0.00444***
	(0.0212)	(0.0212)	(0.00118)	(0.00118)	(0.00120)	(0.00120)
4.quarter	0.565***	0.565***	0.0124***	0.0124***	0.0128***	0.0128***
	(0.0210)	(0.0210)	(0.000581)	(0.000581)	(0.000598)	(0.000598)
Constant	-0.354	-0.353	0.00342	0.00342	0.0582***	0.0582***
	(0.296)	(0.296)	(0.00795)	(0.00795)	(0.0104)	(0.0104)
Observations	2,188,450	2,188,450	2,185,661	2,185,661	2,185,661	2,185,661
R-squared	0.574	0.574	0.636	0.636	0.636	0.636

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Results are estimated by OLS fixed-effects (origin-destination pairs) estimation method. Regressions (i.e. columns 1, 2, 5, and 6) are also controlled by the size of the origin farms. Regressions control for stock.

Table 10: origin outdegree effects (movement from farms to slaughterhouses)

VARIABLES	(1) Dependent Variable: total_traded_heads	(2) Dependent Variable: total_traded_heads	(3) Dependent Variable: heads over origin stock	(4) Dependent Variable: heads over origin stock	(5) Dependent Variable: heads over origin stock	(6) Dependent Variable: heads over origin stock
Number of origin node's bovines	0.0104***	0.0104***	-	-	-0.000202***	-0.000202***
	(0.00213)	(0.00213)			(2.14e-05)	(2.14e-05)
Lagged disease status	1.084***	1.353***	0.0262***	0.0280***	0.0263***	0.0284***
	(0.212)	(0.269)	(0.00533)	(0.00679)	(0.00531)	(0.00678)
Lagged outdegree of origin node	-0.0273***	-0.0273***	-0.000336***	-0.000336***	-0.000256***	-0.000256***
	(0.00442)	(0.00442)	(6.60e-05)	(6.60e-05)	(6.74e-05)	(6.74e-05)
Interaction of origin lagoutdegree and disease status	-	-0.0666**	-	-0.000422	-	-0.000504
		(0.0339)		(0.00157)		(0.00156)
2009.year	0.265***	0.265***	0.00562***	0.00562***	0.00578***	0.00578***
	(0.0295)	(0.0295)	(0.000375)	(0.000375)	(0.000377)	(0.000377)
2010.year	0.407***	0.407***	0.00724***	0.00724***	0.00753***	0.00753***
	(0.0326)	(0.0326)	(0.000273)	(0.000273)	(0.000269)	(0.000269)
2011.year	0.523***	0.523***	0.0101***	0.0101***	0.0106***	0.0106***
	(0.0351)	(0.0351)	(0.000293)	(0.000293)	(0.000297)	(0.000297)
2012.year	0.740***	0.740***	0.0145***	0.0145***	0.0152***	0.0152***
	(0.0386)	(0.0386)	(0.000520)	(0.000520)	(0.000572)	(0.000572)
2013.year	1.219***	1.219***	0.0220***	0.0220***	0.0229***	0.0229***
	(0.0489)	(0.0489)	(0.000446)	(0.000446)	(0.000488)	(0.000487)
2.quarter	-0.0508**	-0.0508**	0.00185***	0.00185***	0.00174***	0.00174***
	(0.0235)	(0.0235)	(0.000229)	(0.000229)	(0.000228)	(0.000228)
3.quarter	0.248***	0.248***	0.00316***	0.00316***	0.00301***	0.00301***
	(0.0242)	(0.0242)	(0.000390)	(0.000390)	(0.000379)	(0.000379)
4.quarter	0.454***	0.454***	0.00779***	0.00779***	0.00785***	0.00785***
	(0.0248)	(0.0248)	(0.000257)	(0.000257)	(0.000258)	(0.000258)
Constant	1.543***	1.543***	0.0216***	0.0216***	0.0488***	0.0488***
	(0.289)	(0.289)	(0.000421)	(0.000421)	(0.00280)	(0.00280)
Observations	2,047,041	2,047,041	2,045,544	2,045,544	2,045,544	2,045,544
R-squared	0.806	0.806	0.294	0.294	0.296	0.296

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Results are estimated by OLS fixed-effects (origin-destination pairs) estimation method. Regressions (i.e. columns 1, 2, 5, and 6) are also controlled by the size of the origin farms. Regressions control for stock.

Moreover, we can look at the effects of total degree of origin nodes on the movements of bovines among farms and slaughterhouses. Of course, as we explained in previous paragraphs, it would be better to use the lagged values of the outdegree and degree of origin nodes. Thus, here we should compute the lagged version of origin degree. Also, we should note that for computing degree, we can not just add the indegree and outdegree of the nodes because one node can receive and send bovine from the same node and so, if we just add the indegree and outdegree of the nodes, we actually counted it twice (which is an incorrect representation of degree), while for the computation of the degree it should be counted once. Accordingly, first, we used the indegree and outdegree values of each node that we had computed in the previous section and created a unique link for each movement. Then tried to delete the repetitive links and keep the unique links once. Then we created the lagged values of origin nodes degrees.

Also, in the previous analysis of outdegree and indegree, we saw that to achieve more reliable results, we should control our dependent variable for the stock of origin nodes because the degree of a node may have a correlation with the stock of the bovines in that node. So, it would be better to use and check results both with variables “total_traded_heads” and “heads over origin stock” as the dependent variable. The results of the effects of origin nodes degree on the movements of the bovines from farms to farms and from farms to slaughterhouses are shown in tables 11 and 12 below.

As we see in table 11, there is not any significant main (pure) effect of origin-node degree on the movements of bovines from farms to other farms. But again, along with the previous results, the interaction of degree and disease status has a positive and significant effect on the movement of bovines from farms to other farms which confirm somehow the opportunistic behavior of the farmers to get rid of their (probably sick) bovines. Also, as we said, this can be lead to an increase in the outbreak of disease among various regions.

Table 11: effect of origin-node degree on bovine's movement from farms to farms

VARIABLES	(1) Dependent Variable: total_traded_heads	(2) Dependent Variable: total_traded_heads	(3) Dependent Variable: heads over origin stock	(4) Dependent Variable: heads over origin stock	(5) Dependent Variable: heads over origin stock	(6) Dependent Variable: heads over origin stock
Number of origin node's bovines	0.0110***	0.0110***	-	-	-0.000414***	-0.000414***
	(0.00221)	(0.00221)			(7.52e-05)	(7.52e-05)
Lagged disease status	-0.890***	-1.122***	-0.0156***	-0.0209***	-0.0153***	-0.0207***
	(0.142)	(0.170)	(0.00280)	(0.00372)	(0.00279)	(0.00371)
Lagged degree of origin node	0.000813	0.000814	0.000628	0.000628	0.000740	0.000740
	(0.00636)	(0.00636)	(0.000857)	(0.000857)	(0.000866)	(0.000866)
Interaction of origin lagdegree and disease status	-	0.0171***	-	0.000392***	-	0.000400***
		(0.00444)		(9.49e-05)		(9.58e-05)
2009.year	0.281***	0.281***	0.00705***	0.00705***	0.00757***	0.00757***
	(0.0241)	(0.0241)	(0.00114)	(0.00114)	(0.00117)	(0.00117)
2010.year	0.394***	0.394***	0.0102***	0.0102***	0.0112***	0.0112***
	(0.0249)	(0.0249)	(0.000933)	(0.000933)	(0.000901)	(0.000901)
2011.year	0.601***	0.601***	0.0172***	0.0172***	0.0187***	0.0187***
	(0.0283)	(0.0283)	(0.000738)	(0.000738)	(0.000832)	(0.000832)
2012.year	0.871***	0.871***	0.0272***	0.0272***	0.0294***	0.0294***
	(0.0327)	(0.0327)	(0.00320)	(0.00320)	(0.00340)	(0.00340)
2013.year	1.764***	1.764***	0.0436***	0.0436***	0.0466***	0.0466***
	(0.0445)	(0.0445)	(0.00174)	(0.00174)	(0.00204)	(0.00204)
2.quarter	0.0437**	0.0437**	0.00598***	0.00598***	0.00599***	0.00599***
	(0.0187)	(0.0187)	(0.00159)	(0.00159)	(0.00159)	(0.00159)
3.quarter	0.156***	0.155***	0.00434***	0.00434***	0.00450***	0.00450***
	(0.0212)	(0.0212)	(0.00119)	(0.00119)	(0.00121)	(0.00121)
4.quarter	0.566***	0.566***	0.0125***	0.0125***	0.0129***	0.0129***
	(0.0210)	(0.0210)	(0.000605)	(0.000605)	(0.000628)	(0.000628)
Constant	-0.414	-0.414	-0.00340	-0.00340	0.0518***	0.0518***
	(0.308)	(0.308)	(0.0169)	(0.0169)	(0.0157)	(0.0157)
Observations	2,188,450	2,188,450	2,185,661	2,185,661	2,185,661	2,185,661
R-squared	0.574	0.574	0.636	0.636	0.636	0.636

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Results are estimated by OLS fixed-effects (origin-destination pairs) estimation method. Regressions (i.e. columns 1, 2, 5, and 6) control for stock. Standard errors are reported in parentheses.

Table 12: effect of origin-node degree on bovine's movement from farms to slaughterhouses

VARIABLES	(1) Dependent Variable: total_traded_heads	(2) Dependent Variable: total_traded_heads	(3) Dependent Variable: heads over origin stock	(4) Dependent Variable: heads over origin stock	(5) Dependent Variable: heads over origin stock	(6) Dependent Variable: heads over origin stock
Number of origin node's bovines	0.0157***	0.0157***	-	-	-0.000334***	-0.000334***
	(0.00184)	(0.00184)			(1.63e-05)	(1.63e-05)
Lagged disease status	1.236***	1.513***	0.0412***	0.0474***	0.0412***	0.0476***
	(0.114)	(0.215)	(0.00444)	(0.00607)	(0.00443)	(0.00605)
Lagged degree of origin node	-0.0191*	-0.0191*	-0.000611***	-0.000611***	-0.000475***	-0.000475***
	(0.0105)	(0.0105)	(4.67e-05)	(4.67e-05)	(4.74e-05)	(4.74e-05)
Interaction of origin lagdegree and disease status	-	-0.0409	-	-0.000911*	-	-0.000944*
		(0.0304)		(0.000489)		(0.000487)
2009.year	0.255***	0.255***	0.0193***	0.0193***	0.0194***	0.0194***
	(0.0196)	(0.0196)	(0.000368)	(0.000368)	(0.000368)	(0.000368)
2010.year	0.374***	0.374***	0.0229***	0.0229***	0.0229***	0.0229***
	(0.0216)	(0.0216)	(0.000356)	(0.000356)	(0.000356)	(0.000356)
2011.year	0.568***	0.568***	0.0318***	0.0318***	0.0320***	0.0320***
	(0.0226)	(0.0226)	(0.000383)	(0.000383)	(0.000384)	(0.000384)
2012.year	0.841***	0.841***	0.0438***	0.0438***	0.0440***	0.0440***
	(0.0257)	(0.0257)	(0.000459)	(0.000459)	(0.000462)	(0.000462)
2013.year	1.537***	1.537***	0.0767***	0.0767***	0.0769***	0.0769***
	(0.0356)	(0.0356)	(0.000546)	(0.000546)	(0.000548)	(0.000548)
2.quarter	0.0571***	0.0571***	0.00770***	0.00770***	0.00753***	0.00753***
	(0.0160)	(0.0160)	(0.000299)	(0.000299)	(0.000299)	(0.000299)
3.quarter	0.276***	0.276***	-0.00204***	-0.00204***	-0.00239***	-0.00239***
	(0.0166)	(0.0166)	(0.000320)	(0.000320)	(0.000317)	(0.000317)
4.quarter	0.540***	0.540***	0.0266***	0.0266***	0.0265***	0.0265***
	(0.0171)	(0.0171)	(0.000328)	(0.000328)	(0.000327)	(0.000327)
Constant	1.451***	1.451***	0.0491***	0.0491***	0.0791***	0.0791***
	(0.190)	(0.190)	(0.000534)	(0.000534)	(0.00151)	(0.00151)
Observations	4,958,576	4,958,576	4,912,573	4,912,573	4,912,573	4,912,573
R-squared	0.751	0.751	0.286	0.286	0.287	0.287

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Results are estimated by OLS fixed-effects (origin-destination unique links) estimation method. Regressions (i.e. columns 1, 2, 5, and 6) control for stock. Standard errors are reported in parentheses.

Also, one thing that we can expect from this pure "insignificant" degree effect and "significant" effect of interaction, is that the central nodes can remain central in course of the time and disease status cannot decrease the number of trade relations from that farm to other farms.

Actually, the trade pattern of a given node with other farms does not decline. That is, even in the case of diseases, the connectivity of a node can be prevalent of the disease status of that node and the number of trading relations remains high, and as we said this can be a reason for the outbreak of a disease. This effect can be also seen by looking at the interaction effect of disease status and degree on the total_traded_heads of the movements from farms to slaughterhouses which is insignificant (columns (1) and (2) of table 12).

In addition, column (4) of table 12 tells us that there is a negative and slightly significant effect of interaction term on the movements from farms to slaughterhouses. This shows that if there is a positive disease status on a farm that has a high level of connections; that farm is to some extent willing to send its bovines to more slaughterhouses and actually it prefer to increase its trading slaughterhouse partners, and so it may decrease (a little bit) the market share of total slaughterhouses.

5. Conclusion

In this chapter, we tried to have a general view on the pattern of bovine movements across Italian farms and slaughterhouses. We found that generally positive disease test which leads to sick bovines has very significant effect on the movement of bovines among different holdings. More specifically, positive disease test and the accordingly existence of sick bovines in a farm in a time period has a negative effect on the movement of bovines from that farm to other farms and a positive effect on the bovine movement from farms to slaughterhouses. Actually, in the case of diseases, more bovines will send to slaughterhouses to avoid the spread of the disease among different holdings.

Then, we notice that some farms have same owner/keeper, and we were wondering to have a different bovine movement pattern in these holdings. So, we decided to assess again our previous results to see to some extent the real effect of the diseases on the movements. Therefore, we found that our previous results are authenticated only if the two trading farm have different owner/keeper. It actually shows that there may be other determinants for the movement of bovine between farms with same owner/keeper. Actually, on the other side, if the effects were driven by movements between farms with the same owners, this could be a rational choice adopted by the owners/keepers in the sense that this shows that owners want to isolate healthy bovines from sick ones.

Then, we analyzed the distance effects on this bovine movement pattern and found that in the case of disease, there is a positive distance effect. It means that in the case of diseases, farms tend to send bovines to far away distances, albeit more transportation costs. This may be due to the fact that the closer farms are subjected to the same disease test systems and perhaps same veterans and so they have similar information about each other status, so, they tend to send their bovines to farther farms. Also, this distance effects showed us that some farms may act in an opportunistic behavior in the sense that to send their sick bovines to larger distances even if it leads to spreading of diseases. This is a conjecture and as we saw in previous chapters, there might be some other factors such as the difference in bovine feed price (specifically, corn) that can motivate farmers to send their bovines to other (probably farther) locations.

Then, we did some network analysis to see the effect of network structure on the bovine movement pattern. More specifically, we saw that the indegree characteristic has a positive main (pure) effect on the bovine movements in general. Actually, it is compatible with the intuition that more a farm receive bovines in a time period, it would be more probable to send more bovines to other places in next period. Also, we saw that only in case of the trade from farms to slaughterhouses, the interaction of indegree and disease status was positive and significant which is reasonably expectable, as well. It says that in the case of diseases, the farms send more bovines to slaughterhouses. Actually, the results on the interaction of indegree and disease tests show us that the farms that have received more (probably sick) bovines in the previous period are more likely to send their bovines to slaughterhouses that farms. This can be seen as a sign that shows us the movement network changes to avoid the spread of diseases among bovines in different farms. This also can help us tracing back the diseases to the nodes that are more exposed to receiving bovines and consequently more subject to contracting diseases.

On the other hand, we found the outdegree is also important only in case of the trade from farms to slaughterhouses. In the sense that the farms that have had a positive disease test records in a period, prefer to extend their slaughterhouse partners and send more bovines to them. So, this means that the market share of previous partners to decrease in the current period. Thie result again is consistent with the fact that holdings trade together in a manner to reduce the risk of an outbreak. Also, knowing the effects of outdegree can help us finding potentially dangerous nodes in the sense that these nodes can be a source of spreading diseases among different premises.

Finally, by looking at the effect of total degree on the movements, we found that only in case of transfer of bovines from farms to slaughterhouses, the degree is significant and the nodes decide about their trading partners in the sense that to decrease the probability of contracting and spreading a disease.

References

- Albert, R., and A.-L. Barabási (2002) “Statistical mechanics of complex networks.” *Reviews of modern physics 74:* 47-97.
- Bigras-Poulin, M., Thompson, R. A., Chriel, M., Mortensen, S. & Greiner, M. (2006) “Network analysis of Danish cattle industry trade patterns as an evaluation of risk potential for disease spread”. *Prev. Vet. Med. 76:* 11–39.
- Carrington, P. J., Scott, J. &Wasserman, S. (2005) “Models and methods in social network analysis”. *Structural analysis in the social sciences, no. 27.* Cambridge, UK: Cambridge University Press.
- Christley RM, Robinson SE, Lysons R, French NP (2005) “Network analysis of cattle movement in Great Britain”. *Proc. Soc. Vet. Epidemiol. Prev. Med*. 234–243.
- Danon, L., A. P. Ford, T. House, P. C. Jewell, M. J. Keeling, G. O. Roberts, J. V. Ross and M. C. Vernon (2011) “Networks and the Epidemiology of Infectious Diseases”. Interdisciplinary Perspectives on Infectious Diseases 2011.
- Dubé, C., C. Ribble, D. Kelton and B. McNab (2008) “Comparing Network Analysis Measures to Determine Potential Epidemic Size of Highly Contagious Exotic Diseases in Fragmented Monthly Networks of Dairy Cattle Movements in Ontario, Canada”. *Transboundary and Emerging Diseases 55:* 382-392.
- Dubé, C., C. Ribble, D. Kelton and B. McNab (2009). “A review of network analysis terminology and its application to foot-and-mouth disease modeling and policy development” *Transboundary and emerging diseases 56:* 73-85.
- Hellevik, O. (2007) “Linear versus logistic regression when the dependent variable is a dichotomy”. *Quality & Quantity*, *43*(1): 59–74. http://doi.org/10.1007/s11135-007-9077-3
- Kao, R. R., L. Danon, D. M. Green and I. Z. Kiss (2006) “Demographic structure and pathogen dynamics on the network of livestock movements in Great Britain”. Proceedings of the Royal Society B: *Biological Sciences 273 (1597):* 1999-2007.

- Marsh, D. and M. J. Smith (2000) “Understanding Policy Networks: towards a Dialectical Approach”. *Political Studies 48:* 4 - 21.
- Paul von Hippel (2009) “Linear vs. Logistic Probability Models: Which is Better, and When?”: http://statisticalhorizons.com/linear-vs-logistic
- Pfeiffer DU (2004) “Geographical information science and spatial analysis in animal health”. *In: Gis and spatial analysis in veterinary science.* Editors: P Durr, A Gatrell. Cabi Publishing, (Oxfordshire): 119-144.
- Mario Vincenzo Tomasello, Nicola Perra, Claudio Juan Tessone, Márton Karsai, Frank Schweitzer (2014) “The role of endogenous and exogenous mechanisms in the formation of R&D networks.” *CoRR* abs/1403.4106
- Volkova VV, Howey R, Savill NJ, Woolhouse MEJ (2010) “Potential for transmission of infections in networks of cattle farms”. *Epidemics 2:* 116–122.
- Vernon, M. C. and M. J. Keeling (2009) “Representing the UK´s cattle herd as static and dynamic networks”. Proceedings of the royal society B 276: 469-476.
- Wasserman, S. & Faust, K. (1994) “Social network analysis”. *Structural analysis in the social sciences, no. 8*. Cambridge, UK: Cambridge University Press.
- Williamson, O.E., (1985) “The Economic Institution of Capitalism: Firms, markets, relational contracting”. the Free Press, New York.

Appendix: Tobit Results

Table A: contagious bovine disease (movement between farms (AL to AL))

VARIABLES	(1) Dependent Variable: total_traded_heads	(3) Dependent Variable: heads over origin stock
Number of origin node's bovines	0.0209***	0.000399***
	(0.000184)	(0.000153)
Lagged disease status	-17.60***	-1.129***
	(0.300)	(0.365)
year_2008	-76.08***	-1.015***
	(0.235)	(0.308)
year_2009	-75.77***	-0.992***
	(0.232)	(0.304)
year_2010	-75.70***	-0.985***
	(0.230)	(0.301)
year_2011	-73.59***	-0.858***
	(0.211)	(0.264)
year_2012	-70.37***	-0.664***
	(0.180)	(0.203)
year_2013	-59.15***	-
	(0.0794)	
quarter_2	-0.0184	0.00196
	(0.174)	(0.00625)
quarter_3	0.123	-0.00366
	(0.177)	(0.00648)
quarter_4	4.774***	0.277***
	(0.131)	(0.0816)
Constant	44.75***	-0.980***
	(0.211)	(0.327)
Log pseudolikelihood	-3685021.3	-1900518.5
Observations	2,188,451	2,185,661

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: coefficients are estimated by tobit model. Standard errors are clusterd by origin-nodes. Standard errors are reported in parentheses. Regressions control for stock.

Table B: contagious bovine disease (movement between farms (AL to SM))

VARIABLES	(1) Dependent variable: total_traded_heads	(2) Dependent Variable: heads over origin stock
Number of origin node's bovines	0.0611***	3.01e-05***
	(0.00272)	(5.99e-06)
Lagged disease status	5.441***	0.0678***
	(0.433)	(0.00791)
year_2008	-13.98***	-0.207***
	(0.567)	(0.00490)
year_2009	-13.77***	-0.199***
	(0.565)	(0.00475)
year_2010	-14.04***	-0.207***
	(0.583)	(0.00499)
year_2011	-12.78***	-0.184***
	(0.529)	(0.00452)
year_2012	-9.898***	-0.141***
	(0.410)	(0.00352)
quarter_2	0.935***	0.0163***
	(0.0611)	(0.000836)
quarter_3	0.273***	-0.0104***
	(0.0460)	(0.000811)
quarter_4	4.880***	0.0766***
	(0.205)	(0.00182)
Constant	-14.56***	-0.0494***
	(0.867)	(0.00558)
Log pseudolikelihood	-11859255	-2960617.5
Observations	4,958,576	4,912,573

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: coefficients are estimated by tobit model. Standard errors are clusterd by origin-nodes. Standard errors are reported in parentheses. Regressions control for stock.

Table C: Same and different owners and keepers (farm (AL) to farm (AL))

	Dependent Variable: total_traded_heads			
VARIABLES	origin and destination nodes have the same keeper	origin and destination nodes have a different keeper	origin and destination nodes have the same owner	origin and destination nodes have a different owner
Number of origin node's bovines	0.0723***	0.0436***	0.0708***	0.0439***
	(0.0104)	(0.0112)	(0.0140)	(0.0112)
Lagged disease status	-66.25	-45.35***	-52.96	-45.63***
	(25.32)	(11.70)	(18.98)	(11.76)
year_2008	-27.61***	-42.02***	-28.09***	-42.16***
	(4.543)	(9.933)	(2.829)	(9.952)
year_2009	-29.10***	-41.25***	-28.27***	-41.41***
	(4.461)	(9.702)	(2.476)	(9.728)
year_2010	-29.73***	-40.82***	-27.07***	-41.03***
	(5.194)	(9.639)	(2.464)	(9.674)
year_2011	-25.34***	-34.78***	-23.65***	-34.92***
	(4.655)	(7.559)	(2.423)	(7.587)
year_2012	-18.67***	-26.98***	-17.06***	-27.09***
	(3.712)	(6.057)	(2.054)	(6.074)
quarter_2	-0.698	0.279	0.466	0.233
	(1.433)	(0.278)	(0.890)	(0.284)
quarter_3	-4.116**	0.121	-3.712***	0.152
	(1.787)	(0.456)	(0.997)	(0.463)
quarter_4	9.978***	11.30***	9.748***	11.34***
	(1.596)	(2.444)	(1.405)	(2.449)
Constant	-5.458**	-38.46***	-7.509**	-38.60***
	(2.520)	(10.02)	(3.753)	(10.03)
Log pseudolikelihood	-142734.7	-4600558.8	-159797.52	-4583388.2
Observations	41,950	2,387,497	49,028	2,380,419

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: coefficients are estimated by tobit model. Standard errors are clusterd by origin-nodes. Standard errors are reported in parentheses. Regressions control for stock.