Who eats for the climate?

A statistical study of socio-demographic factors in relation to the variation in food-related greenhouse gas emissions

Master's thesis in Industrial Ecology

AGNES SÖDERLUND

Department of Space, Earth and Environment
Division of Physical Resource Theory
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2018

REPORT NO. FRT 2018:04

Who eats for the climate?

A statistical study of socio-demographic factors in relation to the variation in food-related greenhouse gas emissions

Master Thesis within the Master's Programme Industrial Ecology

AGNES SÖDERLUND

SUPERVISOR
Fredrik Hedenus

EXAMINER
Jonas Nässén

Department of Space, Earth and Environment, Physical Resource Theory.
Division of Physical Resource Theory
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden, 2018

Who eats for the climate?
A statistical study of socio-demographic factors in relation to the variation in food-related greenhouse gas emissions
AGNES SÖDERLUND

© AGNES SÖDERLUND, 2018.

Technical report no: FRT 2018:04
Department of Space, Earth and Environment, Physical Resource Theory
Division of Physical Resource Theory
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone +46(0)31-772 10 00

Printed by Chalmers Reproservice
Göteborg, Sweden 2018

Who eats for the climate?
A statistical study of socio-demographic factors in relation to the variation in food-related greenhouse gas emissions
Master's thesis in Industrial Ecology
AGNES SÖDERLUND
Department of Space, Earth and Environment
Division of Physical Resource Theory
Chalmers University of Technology

Abstract

Large variations have been found in food-related greenhouse gas emissions from Swedish dietary habits. The food consumption stands for approximately one quarter of the global greenhouse gas emissions, hence is highly contributing to the increased concentrations of greenhouse gases in atmosphere today, triggering climate change. The aim of this study is to investigate if and what socio-demographic factors may explain the variations in diet related emissions among the Swedish population. The study compiles a literature review, which attempts to form the basis for some statistical analysis. Regression models have been elaborated in attempt to describe the connection between the emission variation and socio-demographic factors. A cluster analysis has been carried out to construct diet groups that were evaluated from a socio-demographic point of view and a mean comparison has been performed for conclusions if emission intensive ruminant meat is associated to higher income. The findings show that socio-demographic factors, such as gender and household income, can explain small parts of the variation in food-related greenhouse gas emissions. The conclusion was following that the background factors for whom choosing a more climate intensive diet and who does not is of more complex psychological nature at individual level and cannot be explained by socio-demographic factors only.

Key words: climate change, greenhouse gas emission, agriculture, diet, socio-demographic factors

Det har visat sig vara stora variationer i växthusgasutsläpp från svenskars kostvanor. Den totala matkonsumtionen står idag för en stor del av våra konsumtionsrelaterade utsläpp, vilket bidrar kraftigt till klimatförändringar. Den här studien syftar till att undersöka om, och i så fall vilka, socio-demografiska faktorer som kan tänkas ligga till bakom den stora variationen i växthusgasutsläpp mellan individer i Sverige. Studien innefattar en litteraturstudie som ligger till grund för ett antal statistiska analyser. Regressionsmodeller har tagits fram för att beskriva sambandet mellan variationen i utsläppen och socio-demografiska faktorer. En klusteranalys har utförts för att ta fram dietgrupper som sedan studerats ur socio-demografiska aspekter, samt har medelvärdesjämförelser genomförts för att dra slutsatser om konsumtion av utsläppsintensivt nötkött är kopplat till högre inkomst. Resultaten visar att de socio-demografiska faktorerna kön, samt hushållens inkomst kan förklara små delar av variationen i kostrelaterade utsläpp. Slutsatsen är följaktligen att bakgrunden till vad som avgör vem som äter en klimatintensiv diet eller inte är av mer komplex psykologisk natur som ligger hos individen och inte går att förklara enbart med hjälp av socio-demografiska faktorer.

Acknowledgements

I would like to thank Fredrik Hedenus for being my very helpful supervisor and support through the whole process of my thesis. For discussions and contribution of interesting thoughts and ways to tackle challenges on the way, as well as with expertise and well needed encouragement, making this thesis possible. I would also like to thank David Andersson for support with some statistical issues and as well Jonas Nässén for being my examiner of this thesis. Finally, I would like to express gratitude to my friend Linn Wahlgren and Johan Nyström, your support has meant the world to me when I have been struggling and doubt in myself.

Table of content

List of figures

List of tables

1. Introduction

Climate change is a growing concern worldwide. In 2016 the global temperature in the period of January to September was approximately 1.2°C above the pre-industrial average level for the same period (World Meteorological Organisation (WMO), 2016). Two thirds of the warming have occurred since 1975 (National Aeronautics and Space Administration (NASA), n.d) and much of the increase can be explained by the emissions of greenhouse gases (GHG) from human activities. Rising temperature has large negative effects on ecosystems, food production and water supplies. It causes more extreme weather, increases sea levels and has negative effects on human health (Intergovernmental Panel on Climate Change (IPCC), 2014a). Future changes in the atmospheric concentrations of GHG and temperature are predicted, and scenarios have been projected of a temperature increase between 0.3 and 4.8°C over the next century (IPCC, 2013a). Where on the scale we will end up mainly depends on the level of mitigation work coming to act in the nearest future as well as the climate sensitivity (the climate system response to sustained radiative forcing (IPCC, 2007a). To be able to meet the United Nations (UN) climate target of limiting the global temperature to rise more than 2° Celsius above pre-industrial levels it will not be enough to only focus on mitigation of CO_2 from fossil fuels (Hedenus et al 2014).

Nearly one quarter of the global greenhouse gas emissions (GHGE) can be ascribed to CO_2 due to land use change, CH_4 emissions from ruminants and N_2O from manure handling in the agriculture sector (IPCC, 2014b; European Environment Agency (EEA), 2015; High Level Planet of Experts (HLPE), 2012). This is more than the total emissions from the whole transport area (IPCC, 2014b). In Sweden, food and drinks accounted for approximately 30 % of the emissions from households in 2015 (Naturvårdsverket, 2017a). But relatively little is written about the potential of GHG mitigation in this area (Bryngelsson et al., 2016). However, several studies indicate that our choice of diet has large and various impacts on the total GHG emissions (Bryngelsson et al., 2016, Davis et al., 2010 and Hedenus, et al., 2014). Diets including ruminant meat and dairy have significantly higher total emissions of GHG than plant-based diets. But who is choosing the climate heavy respectively the climate friendly diets have not yet been confirmed through any known study.

Understanding the socio-demographic or other underlying factors behind the choice of our diets could be one way to enable directed and effective climate policy. The most effective way to reduce emissions from the agriculture sector has been suggested to be managed through decreased production of GHG intensive food products, which will be the effect of reduced demand (Wirsenius et al., 2011). Knowing what underlies the choice we make at the grocery store or in restaurants will be helpful for suggesting the most effective measures through political decision-making and policy instrument to limit the consumption of GHG intensive products.

Nässén (2014) has compared GHGE at household level for goods and services by emission data from input and output analyses. The study concluded that higher income generates increased consumption and larger total GHGE. The GHGE in the study is calculated based on money transactions between sectors and direct energy use and emissions in each sector. It is based on primary inputs of fossil fuels and bioenergy and demand of electricity and district heating. It shows a strong connection between income and consumption. But focusing on food consumption will require more information about GHGE based on the biogenic greenhouse gases, such as methane and nitrous oxide that are generated from ruminant meat rearing and manure handling as the emissions from fossil fuel are less important in this sector (Sonesson et al., 2009).

Many previous studies of GHGE connected to food have mainly been related to health and nutrient content (Gazan. et al., 2016; Biesbroek et al, 2014). Scarborough et al. (2012) found that reduced intake of meat and dairy products replaced by vegetables, fruit and cereals prevents death from coronary heart disease, stroke and cancer as well as reduces GHGE. Springmann et al. (2016)

conclude that a lower fraction of meat in the diet will have positive health effects, as well as have climate benefits.

Sjörs et al (2017) have studied the relation between GHGE and nutrition content of Swedish diets. In the study was an individual food-related GHGE estimated based on the food registration in *Riksmaten 2010-11*, a study conducted by the Swedish National Food Agency (Livsmedelsverket; Amcoff et al., 2012) and information from several food-related Life Cycle Assessments (LCA). The individual diets were based on four days diary writing of the participants. One of the outcomes from the study was the revealing of a large variation in the individual food-related GHGE of 0.2 - 6.1 tonnes CO_2 eq. /year.

1.1 Aim

The objective of the study is to investigate if and how strongly socio-demographic factors as gender, age, household income, education and different living situations are contributing to the large variation in individual GHGE associated to food consumption.

1.2 Question formulation

There have been two main research questions set up for this thesis:

RQ1 What socio-demographic background factors can explain the variation in individuals' food-related GHGE?

RQ2 Are there variation in GHGE between different diet patterns. And if so can these be associated to certain socio-demographic background factors?

2. Background

This chapter aims to give a background and an understanding of the purpose of this study. It summarises literature and research in the field of climate change, connects it to agriculture and food production, gives a short summary of consumption patterns and behaviour as well as reviews some findings of previous studies that have related socio-demographics to diets and GHGE.

2.1 Planetary boundaries

In the year of 2009, a group from the Stockholm Resilience Center led by Johan Rockström together with Will Steffen and several other leading researchers defined a safe operating space for humanity with respect to the Earth system. This was to be a tool for decision makers in sustainable global policy making. The report indicated the environmental limits, which within humanity can continue to develop and thrive for generations to come by identifying nine planetary boundaries (Rockström et al., 2009).

An updated version of the report was published in 2015, the same boundaries for the nine systems were confirmed but their state had been updated (Steffen et al., 2015). The nine systems for which the authors found it necessary to define planetary boundaries are the following: 1. Climate change, 2. Change in biosphere integrity (biodiversity loss and species extinction), 3. Stratospheric ozone depletion, 4. Ocean acidification, 5. Biogeochemical flows (phosphorus and nitrogen cycles), 6. Land-system change (for example deforestation), 7. Freshwater use, 8. Atmospheric aerosol loading (microscopic particles in the atmosphere that affect climate and living organisms), 9. Introduction of novel entities (e.g. organic pollutants, radioactive materials, nanomaterial, and micro-plastics). All can be viewed in figure 1.

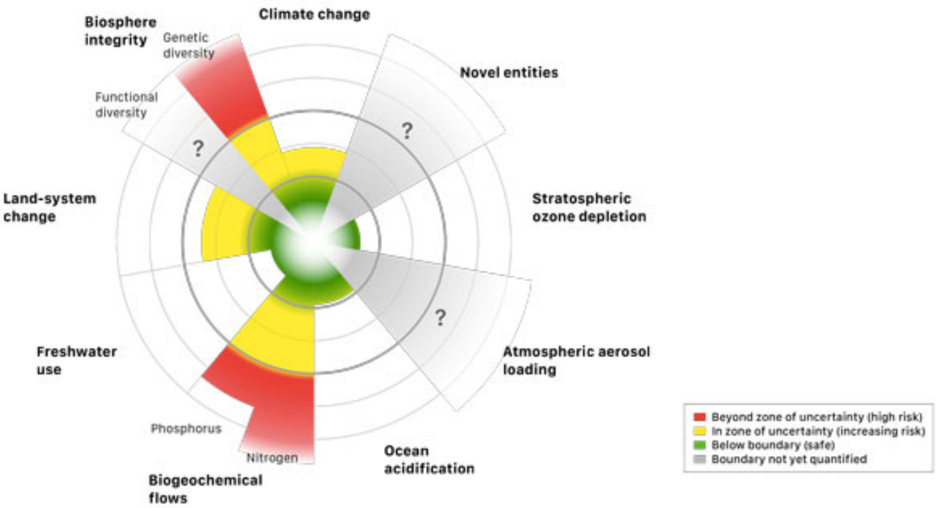


Figure 1 Planetary boundaries (Stockholmresilience, 2015)

As figure 1 confirms, four of the nine boundaries have been exceeded as a result of human activities. These four are climate change, loss of biosphere integrity, land-system change and altered biogeochemical cycles. Climate change is one of three systems that have been identified as, if exceeding its boundary, it will strongly contribute to pushing Earth into a new state and also have a major effect on the remaining system boundaries (Steffen et al., 2015).

2.2 Climate change

Climate on Earth has changed trough history and has never been static. Seven cycles of glacial advance and retreat over the past 650 000 years are indications of fluctuations in climate and temperature on Earth (NASA, 2018a). Changes in climate can be assigned to natural phenomenon as internal dynamics of the climate system such as volcano eruptions (Santer et al. 2014) and ocean currents (Nilsson, 2005) as well as external climate forcings such as small variations in Earths' orbit and solar variation causing variation in radiation and energy received from the sun (NASA, 2018a).

The last 11 700 years, known by geologists as the Holocene epoch, has been relatively stable considering the climate (Steffen et al., 2015). However, there have recently been indications of that this stability is about to break and remarkable changes in the climate have been confirmed by many scientists (IPCC, 2013a). From 1970 until 2015, the global surface temperature has increased with an average rate of 0.17°C per decade. This is more than twice the speed as the 0.07°C per decade increase of the whole observed period 1880-2015 (Dahlman, 2017; NASA/GISS, 2017). In 2017 the average temperature was 0.9°C higher than the average temperature 1951-1980 (NASA, 2018b) and approximately 1.2°C above pre-industrial levels (WMO, 2016). This change since 1880 until today is visualized in figure 2.

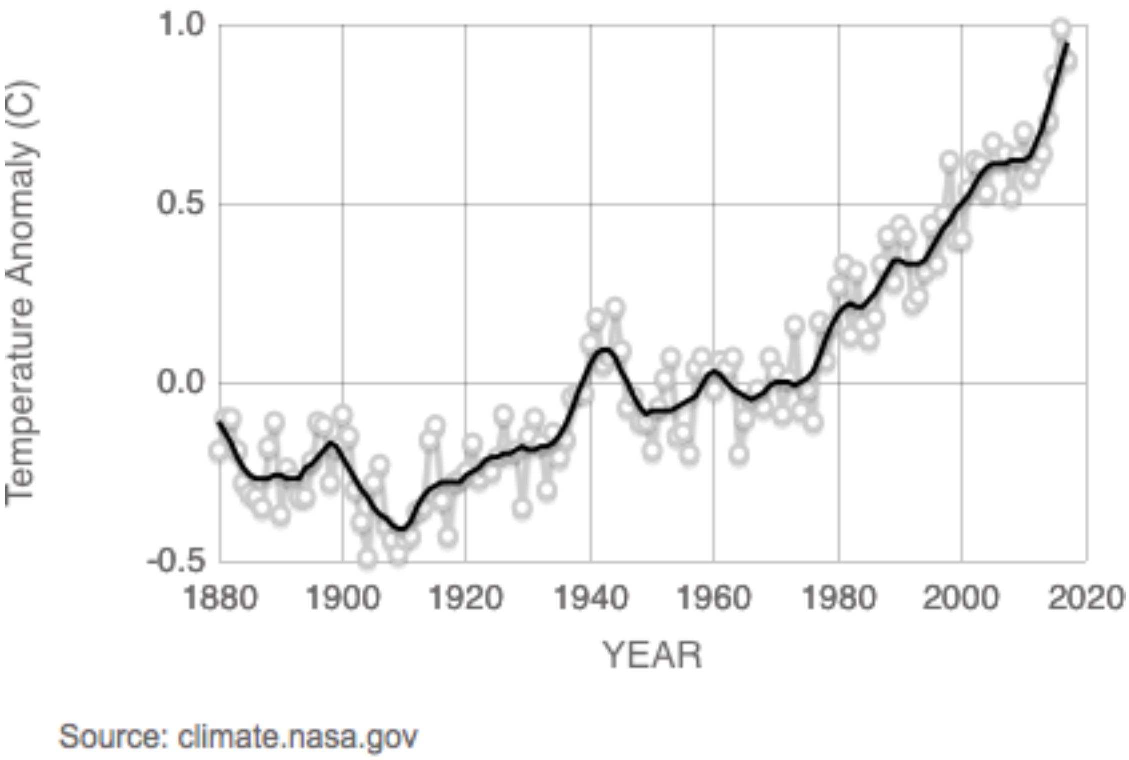


Figure 2 Change in global surface temperature in °C in relation to 1951-1980 average (NASA, 2018b)

The rapid changes can with high certainty be explained by the increased concentration of greenhouse gases (GHG) such as carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O) in the in the atmosphere, caused by human activities (IPCC, 2014b). This change is proposed to be the approach of a new geological epoch where human actions have become the central driver of environmental change on Earth; the Anthropocene (Steffen et al., 2015).

The break point for this change spins back to the first transformation towards an industrial society from an agricultural one that started in the middle of the 18th century (Swedish metrological and hydrological institute (SMHI), 2009). This industrial revolution introduced rapid and intensive use of fossil fuel, technical development, followed by a dramatic population growth, all contributing to increased emissions of CO_2 to the atmosphere (Nationalencyklopedin, n.d.a). Together with the population growth, intensification of food production has resulted in land-use changes such as deforestation that generates carbon dioxide, methane emissions from ruminants and nitrous oxide from manure handling, all to be associated to agriculture (EEA, 2015; HLPE, 2012). An increase from the pre-industrial levels until 2016 of carbon dioxide, methane and nitrous oxide of 45%, 157% respectively 22 % has been registered (WMO, 2017). The level of CO_2 has not been this high in 800 000 years according to the World Meteorological Organization (WMO) Greenhouse Bulletin (2017). Figure 3 shows the concentrations of GHG in the atmosphere from 1850 up until 2012. The recent measurements from 2016 are not in the graph but the concentration of CO_2 in the atmosphere reached 403 ppm, CH_4 1 850 ppb and N_2O 330 ppb that year.

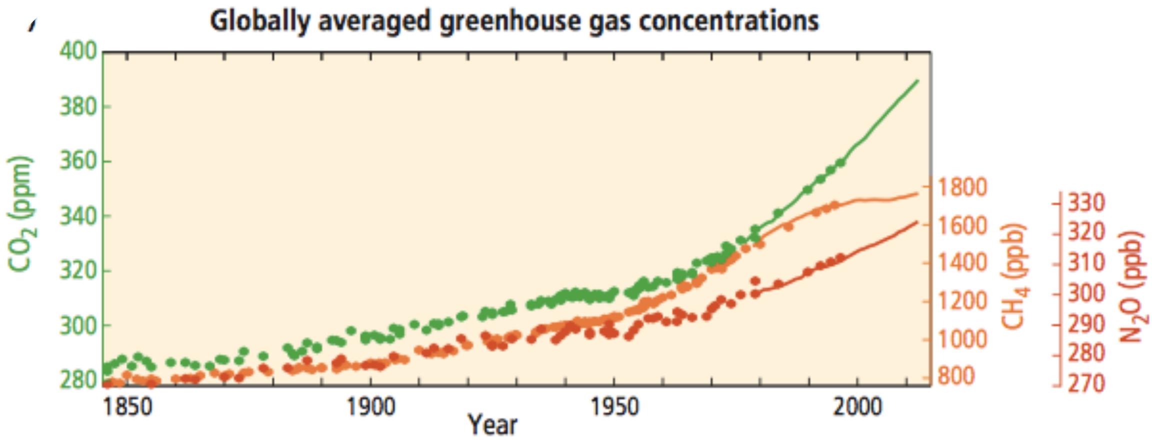


Figure 3 Change in GHGE concentrations in atmosphere 1850-2012 and added measurements from 2016 (IPCC, 2014c)

2.2.1 Physics behind climate change

Climate is driven by the absorption of solar radiation (American Physical Society (APS), 2018). The temperature of the Earth's surface is primarily the result of the absorption and emission of radiation, so-called radiative forcing (Schindell, 2013; NASA, 2018c). Energy from the sun is absorbed by the Earth's surface and radiated back towards space. Greenhouse gases absorb some of the heat and radiate it back towards the surface and create a life-supporting environment, warming the surface of the Earth and the lower atmosphere. Water vapour is the primary greenhouse gas that contributes to the natural greenhouse gas effect (Andersson, 2011). Human activities are indirectly contributing to the increase of water vapour in the atmosphere, as an increased temperature permits the atmosphere to contain more water vapour. It acts as a feedback to the climate. Gases that block the heat from escaping and contributing to greenhouse effect but does not respond physically or chemically to changes in temperature are described as "forcing" climate changes (NASA, 2018c). The gases of importance to "forcing" of climate change in this study includes carbon dioxide, methane and nitrous oxide.

Carbon dioxide (CO_2) is an important but minor component of the atmosphere (NASA, 2018c). It is released through respiration and volcano eruptions and human activities as burning of fossil fuel, deforestation and land use change. Methane (CH_4) is generated both from natural processes in wetlands and marine sediment and from human activities such as decomposition of wastes in landfills, agriculture such as rice cultivation, ruminant digestion and manure management. Nitrous oxide (N_2O) is generated from use of fertilizers, livestock manure, biomass burning as well as fossil fuel combustion.

Global Warming potential

Each greenhouse gas has different capability of absorbing energy, radiative efficiency (Environmental Protection Agency (EPA), 2017). Their lifetime in the atmosphere also varies between a few years to thousands of years. The Global Warming Potential (GWP) has been calculated for each GHG to reflect the effect of the gases in relation to each other, giving them all a certain value of contribution to the warming of the planet. It is a measure of how much energy the emissions of 1 kg gas can absorb over a defined time period in relation to 1 kg CO_2. GWP has the unit of carbon dioxide equivalents (CO_2 eq.) The UNs' Intergovernmental Panel on Climate Change (IPCC) has established GWP for three different time horizons: 20, 100 and 500 years, where 100 years is the most widely used. GWP100 from the latest IPCC report from 2013 (IPCC, 2013b) are: 1 for carbon dioxide, 34 for methane and 296 for nitrous oxide.

2.2.2 Consequences of climate change

The effects of a changed climate have already been noticed on the environment. Glaciers have shrunk, biodiversity has changed, ice on lakes and rivers breaking up earlier and growing seasons have changed (NASA, 2018d). Loss of sea ice, sea levels rise and intense heat waves are all results from the global climate change. But more is about to come with further climate changes. Here are some examples of what have been predicted for the future:

Extreme weathers such as heat waves and drought are caused by increased temperature (National Climate Assessment (NCA), 2014). As the water evaporates from these areas and warmer air can contain more water, the amount of water vapour increases, and heavy downpours are to be expected when the vapour cools (NCA, 2014; Svensktvatten, 2016). Hurricanes may increase as a result of increased sea surface temperature. Extreme heat in itself will result in risks for human health and increased death rates (European Commission (EC), n.d). Air quality changes such as increased formation of ground-level ozone, causing diminished lung function, is affected by heat (NCA, 2014). With higher temperatures the growth of pathogens in water and food intensifies (EC, n.d; NCA, 2014). Water quality, supply and demand are other risk areas. Low flows as well as high can reduce the water quality due to increased precipitation and leaching of fertilizers, nutrients and contaminants (NCA, 2014). Ecosystems will be affected due to climate change by loss of biodiversity. Humanity indirectly depends on biodiversity through ecosystem-services such as fertile soil for food production, access to drinking water and climate regulations (Naturvårdsverket, 2018). Rising CO_2 concentration in the atmosphere increase the absorption of CO_2 in the oceans causing acidification that is harming coral reefs and change the living situation for many spices (National Oceanic and Atmospheric Administration (NOAA), n.d). Warming of the oceans also causing it to expand followed by raised sea levels (NCA, 2014). Agriculture will change through climate change by declining in production of crops and livestock due to weeds, diseases, and insect pests. Erosion and degradation of soils will be more common due to more extreme weathers and intensification of production. Stability and security of food production will be altered. Positive effects might occur due to longer growing seasons but under conditions of higher emission scenarios it will be challenging for the whole sector to enable successful adaption to the changes.

2.2.3 Sustainability goals

In 2015, the 2030 Agenda for Sustainable Development, including 17 sustainable development goals, was adopted by the countries in the UN (UN, n.d.a). The 13th goal is ascribed to climate change, aiming to: Take urgent action to combat climate change and its impacts (UN, n.d.b). Through the Paris Agreement on climate change at COP21 in 2015, 175 countries assigned the first ever universal, legally binding global climate deal; the agreement on work towards limiting the global temperature to rise well below 2°C relative pre-industrial levels and strive for 1.5°C. The agreement is also considering the need for global emissions to peak as soon as possible as well as to undertake rapid reductions thereafter in accordance with the best available science (European Commission (EC), 2018). The implementation of the Paris Agreement is described to be essential for the achievement of the Sustainable Development Goals that provides a roadmap for climate actions that will reduce emissions and build climate resilience (UN, 2016).

2.2.4 Projections of future climate changes

Continued emissions of GHG will increase warming and cause long-lasting, even irreversible, changes in the climate system (IPCC, 2014a). Projections of how the change of climate will evolve up until 2100 have been carried out through different scenarios by IPCC (2013a). The projections vary by emission mitigation depending on socio-economic development as well as climate policy.

In the Fifth Assessment Report of IPCC (2013a) four scenarios as Representative Concentration pathways (RCPs) have been used for simulations of a climate model. Each scenario is identified by its approximate radiative forcing in 2100 in relation to 1750: 2.6 W/m^2 for RCP2.6, 4.5 W/m^2 for RCP4.5, 6.0 W/m^2 for RCP6.0 and 8.5 W/m^2 for RCP8.5. The RCP2.6 assumes rapid and immediate reductions in emissions, RPC4.5 and RPC6.0 are considered two stabilization scenarios. Whereas RPC8.5 is the highest pathway that assumes similar increase of continued emission of GHG as we have today. For an overview of the variation in characteristic emissions of the scenarios, figure 4 is visualizing each scenario annual projected GHGE over time by the black, full lines and corresponding span of projected atmospheric GHG concentrations in ppm CO_2 eq. indicated by the coloured areas covering from 430 ppm CO_2 eq. to just above 1000 ppm CO_2 eq.

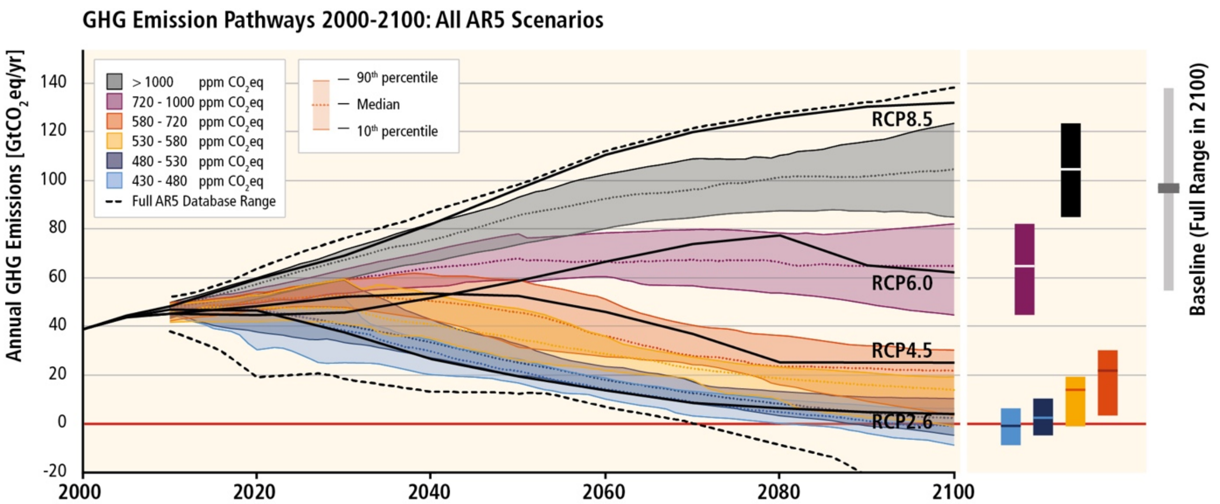


Figure 4 Annual GHGE for four scenarios; RCP 2.6, RCP 4.5, RCP6.0 and RCP8.5 and total atmospheric concentrations for respective scenario up until 2100 (IPCC, 2014b).

RCP8.5 will result in a significantly higher level of GHG concentration (>1000 ppm CO_2 eq.) in the end of 21st century than the other scenarios and corresponds to an annual level of emissions above 120 Gt CO_2 eq. around year 2100. Whereas the RCP2.6 scenario considers net emission near zero, resulting in a GHG concentration level kept below 500 ppm CO_2 eq. by the year 2100.

For the period 2081-2100 the mean surface temperature of the four different scenarios is likely to be in the ranges of 0.3°C to 1.7°C (RSP2.6), 1.1°C to 2.6°C (RCP4.5), 1.4°C to 3.1°C (RCP6.0) and 2.6°C to 4.8°C (RCP8.5) in relation to mean temperature in 1986-2005 (IPCC, 2013a). This is visualized in figure 5 below.

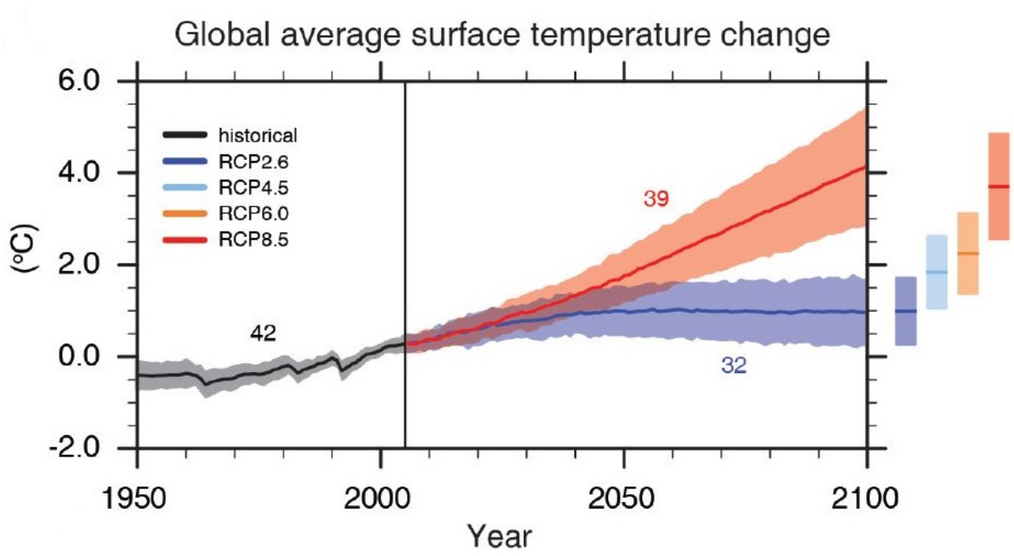


Figure 5 Change in global average surface temperature up until 2100 in °C (Carbonbrief, 2013; IPCC, 2013a)

As figure 5 indicates are the Paris agreement and UNs sustainability goal of keeping the global temperature increase below 2 °C, relative pre-industrial levels, likely to be reached by multiple mitigation pathways (IPCC,2014b). This would demand significant emission reduction over the next decades and emissions close to zero in the end of the century, which requires technological, economic, social and institutional challenges.

2.3 Agriculture and climate

Agriculture is both contributing to climate change and affected by it. However, this section will only focus on the impact agriculture has on climate change.

A growing global population puts pressure on the agricultural system in the manner of secure food production at the same time as minimise the GHGE from the sector (Vetter et al, 2017). Nearly one quarter (IPCC, 2014b) of the global greenhouse gas emissions can be ascribed to the agriculture sector. This is more than the total emissions from the whole transport sector. In Sweden, 66 million tonnes CO_2 eq. was estimated to be generated from household consumption in 2015, approximately 30% was came from food and drinks (Naturvårdsverket, 2017b). This was the same portion as from transport, while 20 % came from accommodation, 12 % to "other" and 9 % from clothes and shoes consumption, as can be seen in figure 6.

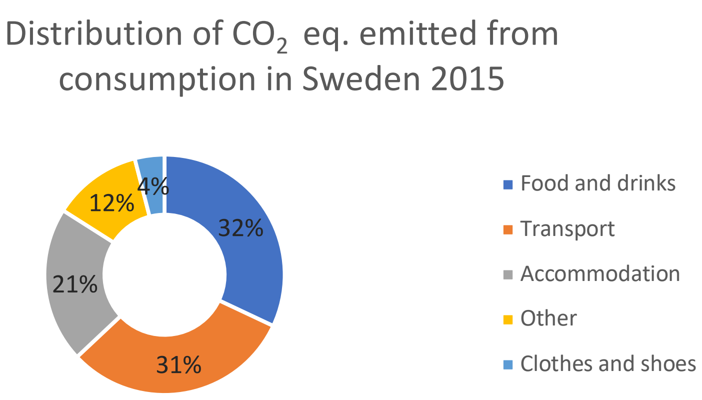


Figure 6 Distribution of consumption areas in the household consumption-based emissions (Naturvårdsverket and Statistiska centralbyrån, 2017)

2.3.1 Product chain

All steps in the food production chain emit GHG but in contrast to many other production chains, the emissions from fossil carbon dioxide are less important (Sonesson et al., 2009). The most important emissions from agriculture are instead emissions of biogenic GHG: methane and nitrous oxide, as well as carbon dioxide from deforestation. Methane and nitrous oxide are two much more potent greenhouse gases than carbon dioxide, as stated in the earlier section of GWP. In fact, agriculture sector is responsible for approximately 60 % of the global non-CO_2 emissions (IPCC, 2007b). Methane is mainly the dominating gas from ruminant meat whereas nitrous oxide is dominating in vegetable production as well as for monogastric animals as pork and poultry. CO_2 from deforestation is one major impact of food production, especially in developing countries (Sonesson et al., 2009). Approximately one third of the GHG allocated to animal production comes from emissions caused by deforestation (Steinfeldt et al. 2006).

One exception, where the major climate impacts of production comes from fossil fuel is for seafood products, whose correlation between energy use and climate impact often is higher due to fossil fuel used in fishing boats. On average, animal-based food generates higher emissions than vegetables. Approximately three quarters of the Swedish food-related GHGE can be assigned to animal-based food (Bryngelsson et al., 2016). But there are variations due to differences in the product chain such as land use, transport, packaging and consumption.

Transport may differ significant assuming less efficient transports as air freight in comparison to sea transport. The difference is on the level of 2 kg CO_2 eq./ tonnes x km for regional air freight compared to 0.01 kg CO_2 eq./ tonnes x km for container ships (Sonesson et al., 2009). Even inefficient distribution to grocery stores and transport between retail store and households have an impact, considering more than 60 % of shopping trips in Sweden where made using cars due to Sonesson et. al (2005). Although, the fact that food shopping often is combined with other activities or shopping implies that food not can carry the whole responsibility for the emissions connected to this transport. However, this makes it difficulties to define the range of actual emissions from this last transport step. Packaging also has an impact on climate that differs due to the choice of material. Plastic made from alternative renewable materials generates less GHGE than plastics produced from crude oil (Posen et al., 2017). At the same time, packaging may reduce food waste by protecting and keeping the food clean and healthy (Sonesson et al, 2009). Food waste is generated from all steps in the product chain and generates approximately 2 million tonnes CO_2 eq./ year which is around 3% of the total annual emissions in Sweden (Jonsson., 2017).

Organic food

Organic food is often considered to be environmental friendly in comparison to conventional food. This is according to avoidance of chemical pesticides and synthetic fertilizers. Considering GHGE from organic, versus conventional diets these emissions are essentially equal, when not considering emissions associated to land use change, due to a study of German consumption data by Treu et al., (2017).

Life Cycle Assessment

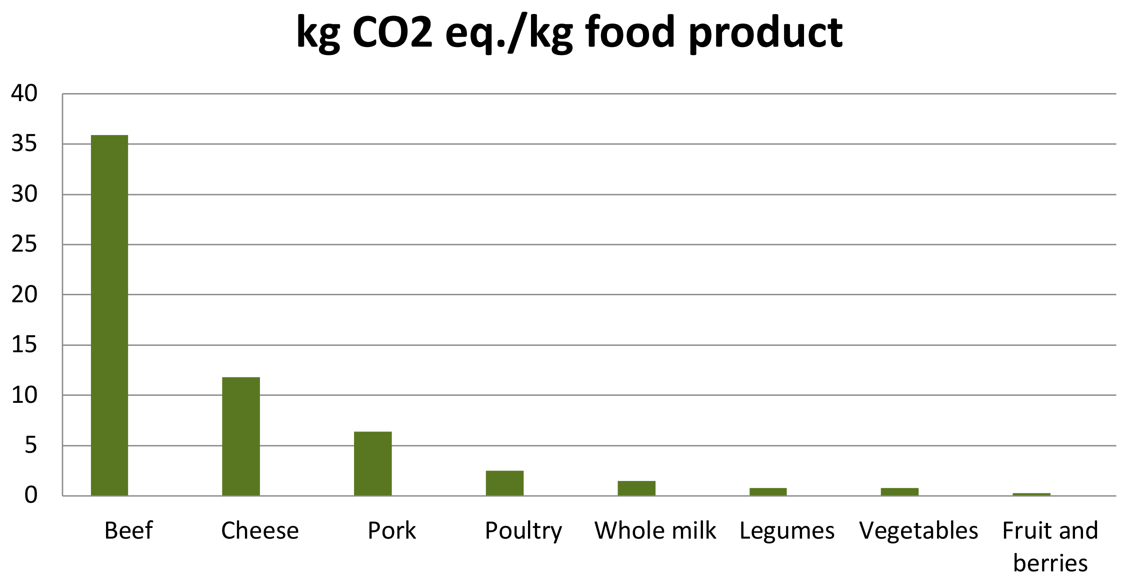
The climate impact of food production is often evaluated through life cycle assessments (LCA). This is a method that evaluates the environmental impacts of a product from "cradle-to-grave", from raw material trough production, distribution, use and disposal (ISO, 2006). It can either focus on one environmental impact e.g. climate impact or many, such as climate impact, eutrophication and acidification. The LCA process contains four main parts:

a) the goal and scope definition - where functional unit, system boundaries and what environmental impact to focus on is decided. Assessing food, 1 kg is often used as functional unit, but it needs to be specified if the product is treated in any way, e.g. 1 kg living cow or 1 kg boneless meat from a cow. The system boundaries decide what flows are to be included in the analysis. If one flow generates more than one product all resources and emissions need to be allocated to the different products, either by economic or physical relations. One example is LCA for milk where it is possible that there also will be meat produced from the same animal.

b) the inventory phase - where data for all included flows are collected.
c) the impact assessment (LCIA)- classification and characterization of the included environmental impacts.
d) the interpretation- where the result is summarized and discussed for conclusions and uncertainties

2.3.2 Food-related GHGE

Due to difference in definition in the goal and scope phase in LCAs, different assessments for the same product will give various results. But to give an indication of the variation among food products, some GHGE data from LCAs used in Sjörs et al. (2017) are compared in figure 7. This GHGE data, among other, is also used in this study. The visualized data is for raw beef[1], pork[1] and poultry[1], whole milk[1,2], dried legumes[1], a mean of vegetables[1,3] and domestic fruit and berries[1]. The data is based on LCAs by Bryngelsson et al. (2016), Flysjö (2012) and Davis et al. (2011) and adjusted to include the same system boundaries. The emissions after the retail phase or related to land use change is not included. The emissions are expressed in kg CO_2 eq./kg food.



Figur 7 Comparison of GHGE in kg CO_2 eq./ kg food product from eight food products: beef, cheese, pork, poultry, whole milk, legumes, vegetables and fruit and berries based on GHGE data used in Sjörs et al. (2017).

As can be seen in figure 7, production of beef emits just above 35 kg CO_2 eq./ kg. This is three times the emissions for cheese on 12 kg CO_2 eq. and five times as much as for pork on 6 kg CO_2 eq./kg. Poultry generates emissions just above 2 kg CO_2 eq./kg and whole milk just below 2 kg CO_2 eq./kg. Legumes and vegetable causes approximately 1 kg CO_2 eq./kg (0.8 $\pm$ 0.6 kg CO_2 eq./kg). Last to the right comes fruit and berries with 0.3 kg CO_2 eq./kg, which is below 1 % of the emissions from beef.

[1] Bryngelsson. et al (2016)
[2] Flysjö (2012)
[3] Davis et al. (2011)

2.3.3 Food-related GHGE and diets

Hence, what we eat result in large variations of GHG emission (Bryngelsson et al., 2016; Davis et al., 2010; Hedenus et al., 2014). For an indication of how different diets can vary in GHGE two diagrams from Bryngelsson (2015) are presented in figure 8. In this study, three made-up diets; “current diet”, “climate smart carnivore” and “vegan” were compared in food content, to the left, and its corresponding total GHGE, to the right. The “current diet” refers to the average diet in Sweden where the protein sources are mainly ruminant meat, other meat and dairy products. The “vegan” does not include any animal-based food and has instead vegetable protein as protein source. The “climate smart carnivore”, who is in-between those diets in the graphs consumes meat, but not from ruminants, and avoids dairy products.

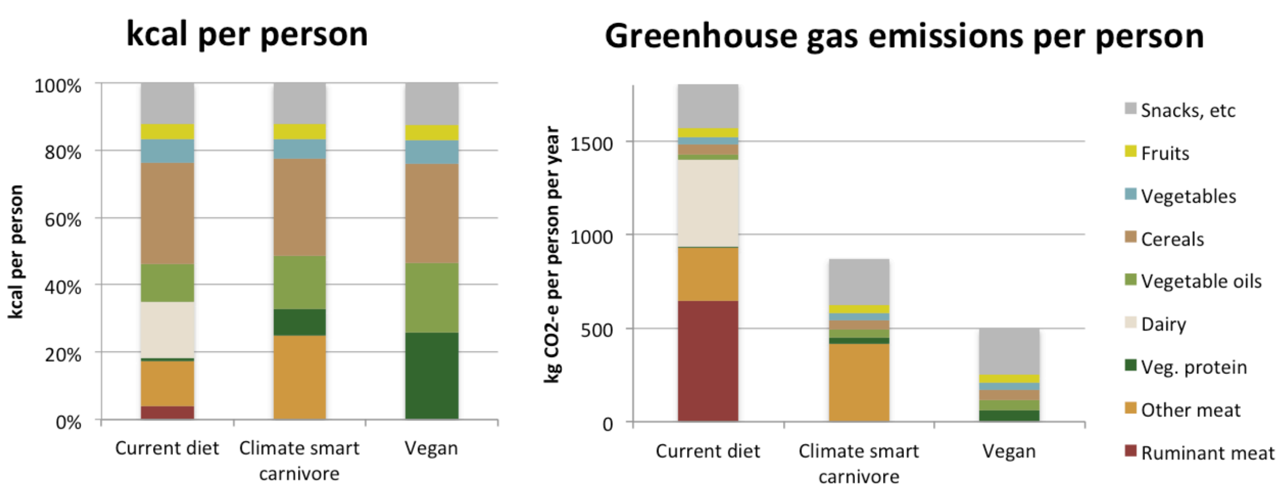


Figure 8 Three proposed diets and their corresponding GHGE (Bryngelsson, 2015)

As can be seen in the right diagram in figure 8, there is a major difference between the emissions from each respective diet. The “current diet” is nearly generating the same emissions as four vegans do in a year and the double of a person avoiding ruminant meat and dairy but still consuming other meat such as poultry.

Scarborough et al. (2014) performed a comparison of diets in the UK which also found that food-related GHGE is approximately twice as high from meat eaters diet than from vegans. In a North American study on GHGE variation between different diet patterns it was concluded that “semivegetarians” (consuming meat more than once a month but less than once a week) had 22 % lower food-related GHGE and “vegetarians” 29 % lower food-related GHGE than “non-vegetarians” (Soret et al., 2014). Davis et al. (2010) looking at different types of meals on a LCA basis, comparing meals of pork and sausage with plant based meals in Sweden and Spain. The GHGE from the Swedish plant-based meal is shown to generate half the GHGE as the Swedish meat based where the plant based in Spain is approximately two thirds of the meat based. Hedenus et al. (2014) also confirms that a diet including less animal-based food causes substantially less GHGE. Even Temme et. al. (2015) confirms the fact that inclusion of meat and dairy products strongly contributes to high total diet related GHG emission.

2.3.4 Food-related GHGE and health

Springmann et al. (2016) suggests that health and climate benefits will both be greater the lower the fraction of animal-sourced in diets are. Whereas Payne et al. (2016) reveals that reduced GHGE from diets are associated with worse health indicators. Both meat and dairy products are sources of important nutrients such as B 12, iron and calcium but at the same time they contain saturated fat and salt which is over consumed by many (Payne et al., 2016). Ruminant meat and charcuteries has also been correlated to cancer (Bouvard et al., 2015). Scarborough et al. (2012) found that reduced intake of meat and dairy products replaced by vegetables, fruit and cereals averts death form coronary heart disease, stroke and cancer and reduces GHGE.

Sjörs et al. (2017) did not just reveal that there was a large variation in GHGE between Swedish individuals, spanning from 0.2-6.1 tonnes CO_2 eq. per person and year, but also that low emissions from diets where associated to high nutrient intake. A group of individuals with the lowest food-related GHGE did show a small but significant compliance to the nutrient recommendations, even though they had a higher intake of added sugar. The study showed that the higher number of recommendation in nutrients achieved through their diet, the lower was the emission of CO_2 eq.

2.4 Consumption behaviour and patterns

The major factors influencing costumers' buying behaviour are cultural, social, personal and psychological (Kasi, n.d). The cultural factors involve a persons' culture, subculture and social class, whereas the social factors are associated to the effects that groups, family, roles and status have on us. The personal variables are defined by the effects of age, lifecycle stage, occupation, economic situation, personality, self-concept and lifestyle, such as interests, opinions and political view and the psychological variable affecting our consumption behaviour through motivation, perception (the ability to organize, identify and interpret from senses) learning, beliefs and attitudes.

2.4.1 Trends of food consumption

The studies of trends in Swedish meat consumption from Jordbruksverket (2016) indicate that consumption of ruminant meat has been about the same since 2000, that the consumption of poultry has nearly doubled whereas pork has decreased slightly. Overall, the meat consumption had a steady increasing trend since the 1980s and increased with approximately 15 % from 2000 to 2015. Growing vegetarian trends in Sweden are notable in different ways. The food retailer Axfood, one of the largest food retailers in Scandinavia, registered a sale increase of vegetarian products by 37 % between 2015 and 2016 and by 25 % the year before (TT, 2017). The increase may be assigned to different reasons such as climate awareness, health trends and animal rights. Market analysis company Nielsen declare that the vegetarian alternatives are the most rapidly growing products in grocery stores (Damberg, 2017). Vegetarian and vegan restaurants have become more common and the variety of vegetarian options in already established restaurants is increasing.

The slow increase in meat prices in Sweden over the last 20 years is explained by the EU entrance in 1995 (Svensktkott, n.d.). The membership caused adaption to the EU's price levels and the concurrence from EU member countries did put a pressure on the prices. In figure 9, the development in ruminant meat prices in Sweden over time is visualised, starting at the price levels in 1958 and goes up to 2011, ruminant minced meat in green and ruminant steak in blue. The graph shows each years' price level of the two products in sek/kg, and the notable drop around 1995 is confirming the effect of the EU entrance.

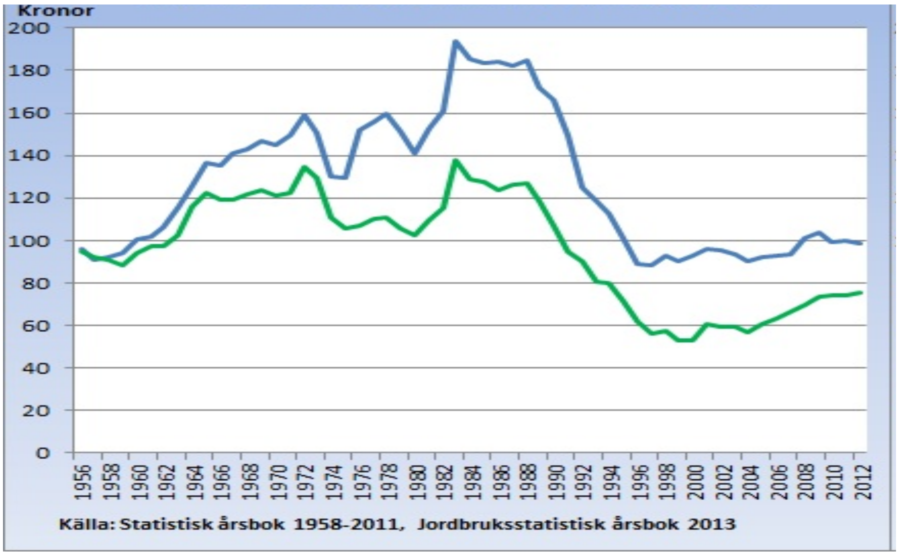


Figure 9 Price development in sek/kg of ruminant meat in Sweden between 1958-2011, ruminant minced meat in green and ruminant steak in blue. (Jordbruksverket, 2013.)

It is not just meat that have has a low price increase rate. The prices of food in general have increased less than the average price development in Sweden over the last years (SCB, n.d.). As the households' income have increased approximately 40 % since 1993, the proportion of expenditures on food has gone from 19 % in 1993 to 14 % in 2017 (Jacobsson, 2011; SCB, n.d.). In 1955, the proportion was 33 % (SCB, 2009).

2.4.2 Socio-demographic factors and diets

Connecting food patterns in the sense of diets with socio-demographic factors are one way to map the food-related GHGE of a population. Hence, finding certain patterns or groups that are generating high respectively low emissions will help in the work of creating optimal directed instruments for reduced climate impact from food. A study on household level for GHGE from goods and services by emission data from input and output analysis confirms that higher income generates increased consumption and larger total GHGE Nässén (2014).

Studies of the relation between socio-demographic variables and diets have revealed that women tend to eat smaller quantities of meat than men according to Treu et al. (2017) and that individuals with higher income tend to spend more money on food in general (SCB, 2004). Globally, higher income has been related to increased quantity of animal food (Johnston et al, 2014) which is in line with Amcoff et al. (2012), who revealed that individuals with higher household income in Sweden have a higher consumption of animal-based products. Clonan et al. (2016), however, revealed that within richer countries people with poorer economy consume more ruminant meat than the people with better economy.

Due to the Swedish Statistiska centralbyrån (SCB, 2004) low-educated individuals are consuming higher quantities of meat and sausages whereas higher consumption of vegetables is connected to higher educational level. Gazan et al. (2016) is another study of French diets relation to socio-economic factors where the result supports that consumption is influenced by socio-economic status and confirm a strong correlation between higher consumption of energy-dense food (e.g. fried and meat products, cereals and potatoes) and low socio-economic status. The association between a higher socio-economic status and "healthy foods", such as whole meal, fruits, vegetables and fish were also confirmed.

Vegetarian and non-meat diets

Several studies have shown that vegetarians are more likely to be young and women in comparison to meat eaters (Allés et al. 2017; Bedford & Barr, 2005). Some studies focusing on socio-demographic and life style characteristics of vegetarians and vegans have shown that vegetarians are more likely to belong to higher socioeconomic categories, such as higher education and economic status, compared to meat eaters (Hoek et al., 2003). While others have revealed the opposite of vegetarians having a lower income compared to meat-eaters (Bedford & Barr, 2005; Allés et al., 2017). Hoek, et al (2003) also confirmed that vegetarians where more likely to live in smaller households in more urbanised areas.

3. Method

The method chapter describes the approach to achieve the objective of this thesis, to answer RQ 1: *What socio-demographic background factors can explain the variation in individuals' food-related GHGE?* And RQ 2: *Are there variation in GHGE from different diet patterns? And if so can these be associated to certain socio-demographic background factors?* The chapter begins with a description of the data that has been used, followed by the hypotheses connected to the research questions. The last section treats the study design of the statistical analyses that have been carried out, including the theory behind these.

3.1 Data

The socio-demographic and GHGE data used in this study are based on *Riksmaten 2010-11* (Amcoff et al., 2012) and Sjörs et al. (2017). The *Statistical Package for the Social Sciences* (SPSS) is the statistic software that has been used throughout the whole study.

3.1.1 Riksmaten 2010-11

Riksmaten 2010-11 is a study carried out by The National Food Agency of Sweden (Livsmedelsverket) where 5 003 individuals aged 18-80 were invited to participate to a mapping of the Swedes' food habits (Amcoff et al. 2012). 1 797 individuals contributed to the study by registering their food and drink consumption for four days complemented by a survey of approximately 50 questions online concerning food habits, living situation, physical activity and work situation. The invited participants represented the population of Sweden according to gender, age and region (Götaland, Svealand and Norrland). Information in the study about the participants such as gender, age, educational level, income and living area was collected from Statistiska Centralbyrån (SCB) population register Registret over totalbefolkningen (RTB), originating from The Swedish Tax Agency's population registration. The final participation of invited individuals did vary across the different socio-demographic groups. Distinctive variations were that invited women participated in a higher extent than men, 41% respectively 31% and only 36% of the invited individuals with lower education contributed to the study whereas 66% among the invited higher educated individuals participated.

3.1.2 Excluded data

To avoid errors due to under-reporting of food intake, the energy intake was compared to the energy demand using the physical activity level (PAL) from *Riksmaten 2010-11*. This was done by the Goldberg cut-off method (Black, 2000) in both *Riksmaten 2010-11* and in Sjörs et al. (2017) to exclude misreporters. By this method 330 under-reporting individuals were excluded from the data and they are not included in any part of this study. 1 467 remained, 627 men and 840 women. Over-reporting was rare and was only registered for one man and two women (Amcoff et al., 2014) These individuals were decided to remain in the study, as they also were included in *Riksmaten 2010-11* and Sjörs et al. (2017).

3.1.3 Diet related GHGE data

The diet-related GHGE used in this study originates from Sjörs et al. (2017). The data is an approximated, individual daily average of GHGE in kg CO_2 eq. based on the food registration in *Riksmaten 2010-11* and validation of several LCA-studies[4] (Sjörs et al., 2017). Global warming potential (GWP) with a 100-year time horizon was used with data from the latest report from the Intergovernmental Panel on Climate Change (IPCC, 2013b); 1 for carbon dioxide, 34 for methane and 296 for nitrous oxide. Data from the food registration in *Riksmaten 2010-11* were used to estimate food-related GHGE by linking all registered different food items to over 100 food groups with approximated values of CO_2 eq. from the LCA-studies (Sjörs et al., 2017). See Appendix for

[4] References under *LCA studies* in the reference list

summary of food groups. The annual average of those approximations is used through this whole study and expressed in tonnes CO_2 eq./year.

LCA-data adjustments
The LCA-data was adjusted to include the same system boundaries (Sjörs et al., 2017). The original data included CO_2 eq. from agriculture and its input. Most of them included emissions to and in the retail phase. Standard emission factors from post-farm processes, such as packaging, distribution and retail, for the ones lacking this information were added. Emissions after the retail phase (e.g. transport to household, cooking and waste management) or related to land-use change were not included. The LCA data were also recalculated for weight changes during food preparation and for avoidable food waste connected to before and after food preparation.

Energy adjustment
In this study, both crude and energy-adjusted GHGE data have been used. The energy-adjusted data is based on the crude GHGE but adjusted to reduce differences in intake of food due to size, physical activity and metabolic effectivity of a person (Johansson, 2014.), while the crude data is the actual food consumption for each individual.

The crude data will evaluate the actual food-related GHGE from diets while the energy-adjusted GHGE is interesting for evaluation of the diets content of GHG intensive food. This will avoid outcome that is based on biological differences e.g. according to of gender and age (Livsmedelsverket, 2017). The crude GHGE is correlated with the energy intake of a person and analysing this will create models that to a high extent defines that high and low-calorie intake generates high respectively low GHGE. Though, this is true, there is the risk of shadowing other important factors for the actual purpose. On the other hand, the energy adjustment is compressing the GHGE data and excludes the authentic variation that still is interesting for this study. Therefore, both GHGE variables were decided to be used and evaluated in the introductory multiple linear regression analyses. The differences in the results will be discussed. The cluster analysis and mean comparison will only be carried out on the energy-adjusted data for analyses of the diets' GHG intensity.

The energy adjustment is based on Willett's (1997) residual method. To retain energy-adjusted GHGE from this method the residuals, from a regression model of the individuals' total energy intake as independent variable and the crude GHGE as dependent variable, are added to a constant (Willett, 1997; Johansson, 2014). The constant is the GHGE responding to the mean energy intake of the sample. Hence, the energy-adjusted GHGE will not be related to the energy intake. See figure 10 for visualization.

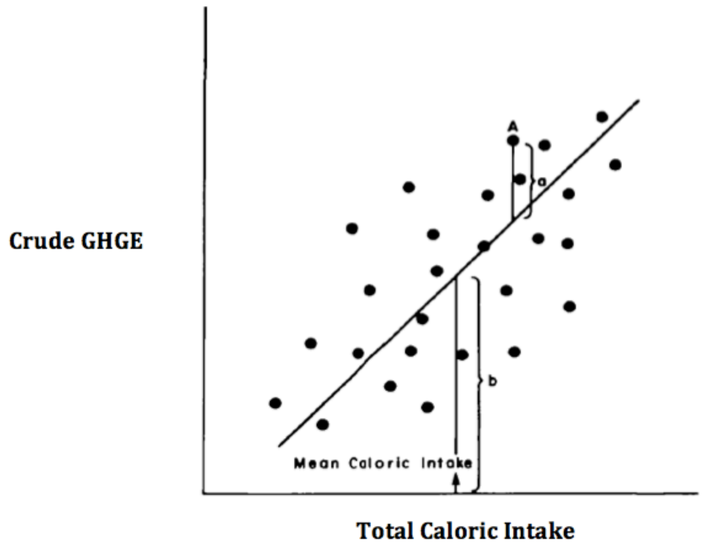


Figure 10 Energy-adjusted GHGE= a+b, where a= residual for subject from regression model with crude GHGE as dependent variable and total energy intake and the independent variable. And b= the expected GHGE for a person with mean energy intake (Willett, 1997).

The distributions of the crude and energy-adjusted GHGE data is visualized in figure 11 and 12.

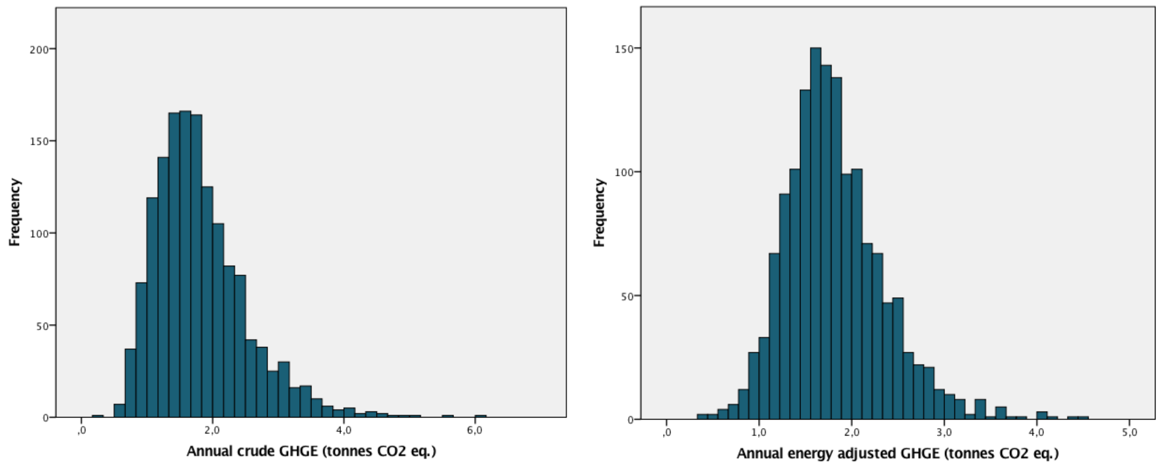


Figure 11 and 12 Distribution of crude GHGE to the left and energy-adjusted GHGE to the right, both in tonnes CO_2/ year.

3.2.4 Socio-demographic variables

Socio-demographic factors are characteristics of a population (Dobronte, 2013). It incorporates factors such as age, gender, migration background, ethnicity, religious affiliation, marital status, household size, living situation, employment and income (Gesis, n.d.). It is a combination of the two close concepts of socio-economic status (SES) and demographics. SES is usually defined to be the relation between social and economic factors within a society (Investopedia, n.d), often connected to factors such as education, income and employment and demography is the study of size, structure and geographical distribution of a population, such as age, gender, marital status, and area of residence (Nationalencyklopedin, n.d.b).

The mean values and distributions of individuals according to the socio-demographic factors used in this study are shown in table 1.

Table1 The socio-demographic variables household income, age mean values and the distribution of gender, education, region, single household, children and student in percent.

Socio-demographic variable		Distribution	Mean
Household income[5]			419 641 sek
Age			49 years
Gender	Male	43%	
	Female	57%	
Education[6]	Low	12%	
	Medium	41%	
	High	47%	
Region[7]	Urban	73%	
	Rural	27%	
Single household		18%	
Children		42%	
Student		8%	

In the household income variable, two individuals with income values of 0 were adjusted to missing value due to the probability of having no household income of any kind is unlikely.

3.2 Hypotheses

The first 1-8 hypothesis (H1-8) are based on the literature review on some previous studies in the field of socio-demographic relations to diet patterns to be able to predict the outcome for RQ 1. The socio-demographic background factors that are assumed to most probable to affect food-related GHGE are gender, household income and education (H1-3). These factors have been found in several studies where they have been associated with certain food patterns. In addition, age, children, single household, region and student (H4-8) have also been evaluated in this study, even though these factors have not been found as frequently occurring in correlations with certain food patterns. These are, however, more common in studies of generalized food patterns such as vegetarian/ vegan verses non-vegetarian.

Drawing conclusions of how socio-demographic factors are associated to the total variation in food-related GHGE from studies generalising eating patterns as "vegetarian" and "non-vegetarian" will involve difficulties mainly according to two reasons. Firstly, considering that the content variation of a meat-eater's and non-vegetarian diets can be large due to what kind of meat and amounts of this that is consumed also that it tells nothing about the inclusion of dairy products. As an example of Bryngelsson (2015) where the ruminant meat eaters' food-related GHGE was more than double compared to the emissions from the "climate smart carnivores" diet. Secondly, the fact that vegetarians make up for a small part of the population, seldom more than 10 %

[5] Household income is the disposable income for each household based on 2008. It is the total income of employment of the household+ capital income+/- transfers of capital – taxes (Amcoff et al. 2012).

[6] Education levels classified as low= elementary school, medium= secondary school and high= university or higher (Amcoff et al., 2012).

[7] Region was classified by urban > 90 000 residents in a radius of 30 kilometres from the municipality centre (the most population dense assembly in the municipality) and rural< 90 000 residents in the ratio of 30 kilometres from the centre (Amcoff et al., 2012).

(Loughnan, 2014), and as a vegetarian diet on average has 50 % lower GHGE than a general average diet including meat (Scarborough et al., 2014) sums up that the overall effect of cut in emissions from food, looking at the total population, are minor. Some of these studies have been source for creating the hypothesis as they still can contribute to some interesting information for predictions, though these hypotheses are more vaguely contributing to the group variation in food-related GHGE.

H1-8 initiates with the hypothetic statement of how the specific variable is expected to influence food-related GHGE, both crude and energy-adjusted. The statements are then followed by the arguments found in literature.

H1 Gender

- Men have higher food-related GHG emissions than women and have a diet with larger share of GHG intensive food types.

This effect is expected to be larger in the crude model, justified by the reason that women have lower energy intake due to biological reasons (Livsmedelsverket, 2017). Women's recommended daily calorie intake is 15-20 % lower than for an average man. Men have also previously been found to consume larger quantities of meat than women (Treu et al. 2017; Räty & Carlsson-Kanyama, 2010) and even higher quantities of red meet (beef, veal, pork and lamb) (Clonan et al., 2016; Rydhagen 2013; Wyness et al., 2011) and processed meat, as well as more often (Clonan et al., 2016). The meat norm is associated to masculinity (Loughnan, 2014) and vegetables, fruits and sweet foods represent the idea of femininity (Roos et al., 2001). In 2007, Naturvårdsverket investigated the knowledge and attitudes about climate change and found that men are less engaged and willing to change their behaviour to reduce their environmental impact (Söderström & Dahlberg, 2007).

H2 Household Income

- Individuals with a higher income have higher food-related GHG emissions and have a diet with higher shares of GHG intensive food types.

Household income has been showed to be the most important variable to explain the total variation in GHGE from consumption (Nässén, 2014), higher income generates higher total GHGE from goods and services, food included. It is also confirmed that higher income implies higher consumption of animal-based food products (Amcoff, et al. 2012). Ruminant meat is costlier than many less GHG intensive food products e.g. chicken, pork and vegetable-based food (Jordbruksverket, 2012).

H3 Education

- Individuals with higher education have lower food-related GHG emissions and have a diet with lower shares of GHG intensive food types.

This hypothesis is assumed for both crude and energy-adjusted GHGE. High educational level has been associated to diets rich in plant-based rather than animal-based food (Lacour et al., 2018). This has also been found to be a factor for climate change mitigation actions (Semenza et al., 2008). Low educated individuals consume higher quantities of meat and sausages whereas higher educated individuals have a higher consumption of vegetables (SCB, 2004; Ricciuto. et al, 2006). There are also studies showing that vegetarians are more likely to be higher educated than meat eaters (Allés et al, 2017) (Hoek et al., 2003).

H4 Age

- Older individuals have lower food-related GHG emissions and have a diet with lower shares of GHG intensive food types.

This effect is expected to be larger in the crude model, justified by the reason that older people have lower energy intake due to biological reasons (Livsmedelsverket, 2017). Diets rich in plant-based in comparison to animal-based food have been associated to older individuals (Lacour et al., 2018).

H5 Children

- Individuals with children have higher food-related GHG emissions and have a diet with higher shares of GHG intensive food types.

This hypothesis is tried for both crude and energy-adjusted GHGE. Women with children have been shown to consume more dairy products and men with children have a higher consumption of ruminant meat, than women and men without children (Al-Adili & Aronsson, 2014). Households with children have also been shown to consume greater quantities of dairy products whereas households with only adults buy more vegetables and fruit (Ricciuto et al, 2006). It has also been found in several studies that it is more likely that an individual is vegetarian if they do not have children (Allés et al, 2017) (Davey et al., 2002) (Hoek et al., 2003).

H6 Single household

- Individuals in single households have lower food-related GHG emissions and have a diet with lower shares of GHG intensive food types.

This hypothesis is assumed for both crude and energy-adjusted GHGE. Effects of influences from other people's diet and food habits are behind this hypothesis. Several women in Macdiarmid et al (2015) explained the reason why they ate meat in the extent they did was due to their husbands' request for meat. They also said that they probably would have eaten less if not being influenced by their husbands. Social influences in food and eating behaviour are well established (Herman et al., 2003). Vegetarians are more often living in single households than meat eaters (Davey et al., 2002).

H7 Region

- Individuals living in urban areas have lower food-related GHG emissions and have a diet with lower shares of GHG intensive food types.

This hypothesis is assumed for both crude and energy-adjusted GHGE. People with diets richer in plant-based rather than anima-based food has been found to live in urban areas rather than rural (Lacour et al., 2018). Even Hoek et al. (2003) found that vegetarians are more likely to live urban. It is more difficult to break the norms in smaller societies outside the cities (Broman, 2017). This can possibly also include that the meat norm is still stronger in the rural than in the urban areas.

H8 Students

- Students lower food-related GHG emissions and have a diet with lower shares of GHG intensive food types.

This hypothesis is assumed for both crude and energy-adjusted GHGE with possibly higher effect in the energy-adjusted model due to assumption of younger individuals having higher energy intake than older (Livsmedelsverket, 2017). Students in general tend to be more interested in changes such as considering environmental issues (Olli et al., 2001) and may therefore be assumed to be more willing to change diets in a climate friendly direction.

H9 This hypothesis is carried out to predict the outcome of RQ 2 and suggests that grouping individuals according to differences in diet content would generate significant differences among those diet groups' food-related GHGE. The reason for this assumption is according to the large variation of GHGE from different food types (Bryngelsson et al., 2016; Davis, 2010; Hedenus et al., 2014). The hypothesis is also that socio-demographic factors could characterise the diet groups so that who eats a certain type of diet, assigned to certain food-related GHGE, could be identified. Diet groups including mainly vegetable-based food and low GHG intensive meat such as chicken, are assumed to have lower food-related GHGE and groups of individuals consuming high quantities of ruminant meat and dairy products are assumed to show higher food-related GHGE. The socio-demographic factors for these certain groups are expected to be in line with the hypothetic statements supported by the corresponding arguments in H1-8.

H10 This hypothesis is connected to both research questions and suggests that individuals that consume any GHG intensive *fine ruminant meat* (*FRM*), including: salted ruminant meat, smoked ruminant meat, sirloin steak, Scotch fillet, veal, roast beef, ruminant steak and brisket of beef, compared to them who does not, generates a significantly higher total food-related GHGE as well as having higher a household income. This is assumed according to the substantially higher GHGE of ruminant meat (Bryngelsson et al., 2016; Davis, 2010; Hedenus et al., 2014). The assumption that individuals who consumes *FRM* have a significant higher household income is due to the marginal higher price compared to vegetable-based food and many other meats such as chicken (Jordbruksverket, 2012).

3.3 Study design

This section explains the statistical tools that have been used to answer the research questions. Three different statistical analyses have been practiced: multiple linear regression analysis, a cluster analysis and a GHGE and income mean comparison between individuals consuming and not consuming GHGE intensive ruminant meat.

3.3.1 Multiple linear regression analysis

The first statistical analysis has been explorative in an attempt to answer the first research question of: What socio-demographic background factors can explain the variation in individuals' food-related GHGE?, connecting to H1-8. The aim was to find suitable multiple linear regression models that describe if and how strongly gender, household income, education, age, student, single household, children and region contribute to the variation of different diets climate impact.

Regression analysis is a research tool frequently used in modelling to find whether variables co-vary and also to quantify how strongly they do so (Dudovskiy, 2018; Sundell, 2009). The multiple linear regression equation in this study consists of the independent socio-demographic variables and the response dependent variable of diet based GHGE for an individual in Sweden (Mendenhall & Sincich, 2003), see equation 1.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \text{ (eq. 1)}$$

Where y is the dependent variable of GHGE, $x_1, x_2 \dots, x_k$ are the independent socio-demographic variables, β_i are the unstandardized regression coefficients that determines the contribution of the independent variable x_i. As an independent variable, x_1, is increased by 1 unit, y is increased by β_1 and ε is the random error.

The dependent GHGE variables are on continuous scales as and the independent variables are continuous or categorical; nominal or ordinal. Nominal means two or more categories. Ordinal variables are similar to nominal variable but are also ranked (Leard Statistics, 2013a). Dummy variables were created for: gender, education, region, single household, children and student to enable these categorical variables in the regression model (Sundell, 2010a). These have the values of 0 or 1 indicating absence or presence of an attribute. In this way it can be treated as a variable on an interval scale and hence be used in regression analysis. Education is a variable with three ranked categories, hence one dummy variable has been created for each attribute. The household income variable is log transformed to normalise the distribution of this variable. Table 2 is summarizing all independent variables that have been used in the study, presented as logarithmic, categorical or continuous. The range of the continuous variables is also shown.

Table 2 The independent variables used in this study classified by type and range or dummy value.

Variable	Type	Range /Dummy value
Household income	log	4 200-3 390 906 kr
Age	Continues	18-80
Gender	Categorical – Nominal (Dummy)	Woman=0 Man=1
Education	Categorical – Ordinal (Three Dummy variables)	Low=1 0 0 Medium= 0 1 0 High= 0 0 1
Region	Categorical – Nominal (Dummy)	0=rural 1= urban
Single household	Categorical – Nominal (Dummy)	0= single household 1= more than one person in household
Children	Categorical – Nominal (Dummy)	0= no children 1= one or more children
Student	Categorical- Nominal (Dummy)	0=not student 1= student

Stepwise regression

Two different models have been analysed; one with crude GHGE and one with energy-adjusted GHGE as dependent variable. The regression analyses are evaluated through stepwise regression on the statistically significance level of α=0.05. Stepwise regression is a typical way to preform regression analysis in explorative purposes and that is why this methodology is practiced in this study (Mendenhall & Sincich, 2003; Löfgren, 2014; Field, 2009). All independent variables of interest are entered to take part of the model. It uses an algorithm to find the best grouping of independent variables. It does multiple regression repeatable times and removes the variables with the lowest correlations, leaving the variables that contributes most to the prediction equation (Mendenhall & Sincich, 2003).

In SPSS the variables are evaluated through the probability of F (p-value) (Horber, 2017). It starts by fitting all included independent variables of the form in eq. 2.

$$E(y) = \beta_0 + \beta_1 x_i \quad \text{(eq. 2)}$$

where x_i is the i^{th} independent variable, i=1,2… k.

SPSS are entering the variable with the smallest p-value, smaller than the entry value of α=0.05, to the model (Horber, 2017; IBM, n.d.). At the next step the variable with the lowest p-value, of the variables not yet included, is added to the model. If the p-value of an already included variable becomes larger than the removal value of α=0.10 when a new variable is included, the variable with increased p-value is removed. This goes on until no more variables receive p-values below the entry value and no more are removed according to the removal value. A criterion for this method is that the data is somewhat normally distributed, which is why household income is log transformed, and that there is no multicollinearity among the independent variables. (School of Geography, University of Leeds, n.d.).

Multicollinearity

Multicollinearity means that there is a strong correlation between two or more independent variables in a multiple regression model (Allen, 1997). The problem with strong correlations between independent variables in a multivariate regression model is the difficulty to separate the effects of the variables on the dependent variable. Multicollinearity can be evaluated by looking at the correlation between the independent variables in the model. In this study it will be evaluated from the Tolerance value in a collinearity test, see equation 3:

$$Tolerance = 1 - R_i^2 \text{ (eq. 3)}$$

R_i^2 is the proportion of variance in the independent variable i, that is associated with another independent variable in the regression model. Tolerance is following the proportion of variance in variable i that is not related to the other variable. A tolerance value of 0.10 is not unusual to use as rule to identify collinearity and sometimes even as high as 0.25 is considered acceptable (Sundell., 2010b, O'Brien, 2007). The lower value the more collinearity between the two independent variables where 1 is the maximum.

$R^2 - adjusted$

The coefficient of determination, R^2, gives information about the goodness of fit of a model (Mendenhall & Sincich, 2003). It is a measure of how well the approximated regression fits the real set of data, a measure of how much of the variance in the data that is explained by the linear regression equation. The value of R^2 is between 0 and 1 where 1 defines perfect fit.

As the number of variables in the multiple linear model grow, R^2 increases even though it has nothing to do with the dependent variable (Frost., 2013; Mendenhall & Sincich, 2003). A model with many independent variables may seem to have a better fit because many variables starts to model the random noise in the data and produces a high R^2. This makes it difficult to make predictions. A way to overcome this is to use the adjusted R^2. It is the coefficient of determination adjusted for the numbers of independent variables in the model. The adjusted R^2 is punishing models that contain many variables with low explanatory power. It reduces the value of R^2 when a variable improves the model by less than is expected by chance and only increase if the added variable improves it more than it would be expected to do by chance.

Following equation 4-7 is how adjusted R^2 is achieved from R^2 (Mendenhall & Sincich, 2003):

$$\mathrm{R}^2 = 1 - \frac{\mathrm{SSE}}{\mathrm{SS_{yy}}} \text{ (eq. 4)}$$

SSE is the sum of squared errors of prediction or, also known as the residual sum of squares:

$$\mathrm{SSE} = \sum(\mathrm{y_i} - \hat{\mathrm{y}}_\mathrm{i})^2 \text{ (eq. 5)}$$

Where $\hat{\mathrm{y}}_\mathrm{i}$ is the predicted value of $\mathrm{y_i}$ for the regression model. And $\mathrm{SS_{yy}}$ is the sample variation of y:

$$\mathrm{SS_{yy}} = \sum(\mathrm{y_i} - \bar{\mathrm{y}})^2 \text{ (eq. 6)}$$

The adjusted R^2 is calculated as follows equation 7.

$$\mathrm{R^2_{adj}} = 1 - \left[\frac{(\mathrm{n}-1)}{\mathrm{n}-(\mathrm{k}+1)}\right]\left(\frac{\mathrm{SSE}}{\mathrm{SS_{yy}}}\right) = 1 - \left[\frac{(\mathrm{n}-1)}{\mathrm{n}-(\mathrm{k}+1)}\right](1 - \mathrm{R}^2) \text{ (eq. 7)}$$

where n is the total sample size and k the number of independent variables in the regression model. $\mathrm{R^2_{adj}}$ is always smaller than R^2.

Effect size

The effect size has been evaluated through the unstandardized coefficients, where the change in the dependent variable is associated with a change in the independent variable of 1 unit. Negative coefficients indicating a decrease and positive an increase. Concerning the dummy variables, the coefficients are the effect of one attribute in comparison to the reference group (the other attribute) (Sundell, 2010a). One other way to look at the effect size of the independent variables in a multiple linear regression is to use the standardized coefficient ($\hat{\beta}$*) (Nieminen et al., 2013). It is a measure that enables to compare variables on different scales of measurements. They show the relative weight of the variable by rescaling the variable to have a mean of 0 and a standard deviation of 1. Acock (2014) argues that the value of $\hat{\beta}$* can be interpreted as $\hat{\beta}$*< 0.2 considered weak effect, 0.2<$\hat{\beta}$*< 0.5 moderate and $\hat{\beta}$*> 0.5 strong. Hence, some sources argue to be careful using this measure of effect size (Fox, 2016; Harrell, 2015). That is why evaluation of effect size will be done in those two different ways.

3.3.2 Cluster analysis

k-means cluster analysis has been used in the second part of this study in order to answer the second research question of: Are there variation in GHGE from different diet patterns. And if so can these be associated to certain socio-demographic background factors? connecting to H9. Clustering has been a way to group individuals by some distinctive characteristics in their food consumption. The reasons to carry out this kind of grouping of individuals in this study was to see if there are patterns in socio-demographics in-between people eating the same sort of diet, following assigned to different intensities of GHGE.

Principal Component Analysis

Principal Component Analysis (PCA) was used as a pre-process in this analysis to simplify the clustering of 23 main food categories in measure of kilocalorie intake into some smaller number of components, for expressive clustering. The 23 main food categories were constructed from over 100 food groups from the original material of *Riksmaten 2010-11*. The 23 main food categories are summarized in table 3. The food groups contributing to each food category are found in Appendix. Some food groups were excluded due to difficulties of defining their character and content. These groups where; soup (no ingredients registered), mayonnaise sauces (no ingredients registered) and other (no specification).

Table 3 Food categories constructed for the PCA analysis.

	Food categories		
1	Ruminant meat	13	Potatoes/root vegetables
2	Pork	14	Fruit and berries
3	Poultry	15	Legumes
4	Fish/seafood	16	Pasta/rice/grain
5	Mixed meat	17	Grain products
6	Sausage	18	Nuts/seeds
7	Pizza/hamburger	19	Vegetarian milk prod.
8	Egg	20	Vegetarian substitute
9	Dairy (excluding cheese and yoghurt)	21	Desserts/sweets
10	Cheese	22	Snacks
11	Yoghurt/sour milk	23	Fat (oil, butter, margarine)
12	Vegetables/mushrooms		

PCA is a mathematical method aiming to find systematic structures in original data and is often used in explorative purposes as a "backstage" tool with the intention to reduce dimensions of a complex data set, whilst keeping as much of the original information as possible (Berglund, 2011). The PCA in this study constructs eleven new components from the 23 food categories and generates a new standardized score value for every new component for each of the 1467 individuals (Berglund, 2011). These score values reflect the variation in energy intake of certain food categories and will be further used for clustering. It was useful for this study due to the number of 23 food category variables had to be reduced (merged) to enable expressive clustering. Below follows a more detailed description of the method.

PCA uses orthogonal transformation to convert correlated variables into uncorrelated new components. It analyses the variation, the variance, of energy intake for each individual in the 23 food category-variable space, according to equation 8.

Variance of a sample: $s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})}{n-1}$ (eq. 8)

Where x_i are the individual data points, $\bar{x}$, the mean of the sample and *n,* the number of data points.

It finds components that catch the variation in the original variable data until all variation is described in the components. The purpose of PCA is to discover if any of the components explain more of the total variance in the variable space than each variable does itself. If so, it is an interesting component and has an eigenvalue, variance, above 1. A PCA always generates as many components as there are variables, in this study 23 food categories, but not all have an eigenvalue above 1. The next step in the PCA will then be to find what variables load at the different components with an eigenvalue above 1. The loading is simply the correlations between the variable and the component and is summarised in a component matrix. See equation 9 for correlation equation.

Correlation: $r(x, y) = \frac{\frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}}{\left(\sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}\right)\left(\sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}}\right)}$ (eq. 9)

Where x_i, y_i are the individual data points, $\bar{x}, \bar{y}$ are the means of the sample and n, the number of data points.

Loadings of 0.3 are used as minimum value for correlation, where perfect correlation has a value of 1. The higher the loading, the more is the variable correlated to the new components. This brings characteristics and information about each component. It is often difficult to interpret the variable structure looking at the loadings in the component matrix. Rotation of the component axes in the space of loading values of the variables is a common method to clarify the structure between them. VARIMAX rotation is used in this study to simplify the interpretation. The component axes are rotated around origin with the axes orthogonal to each other until maximum variance of the loadings is reached. The new loadings are summarised in a rotated component matrix.

PCA geometrics

The first step in the PCA was to standardise the 1467 observations (number of individuals) for the 23 food variables, moving the data to the centre of the 23-dimensional coordinate system of variables, also called mean centring (Dunn, 2016). This is geometrically illustrated for an example of three-dimensions in figure 13 and 14.

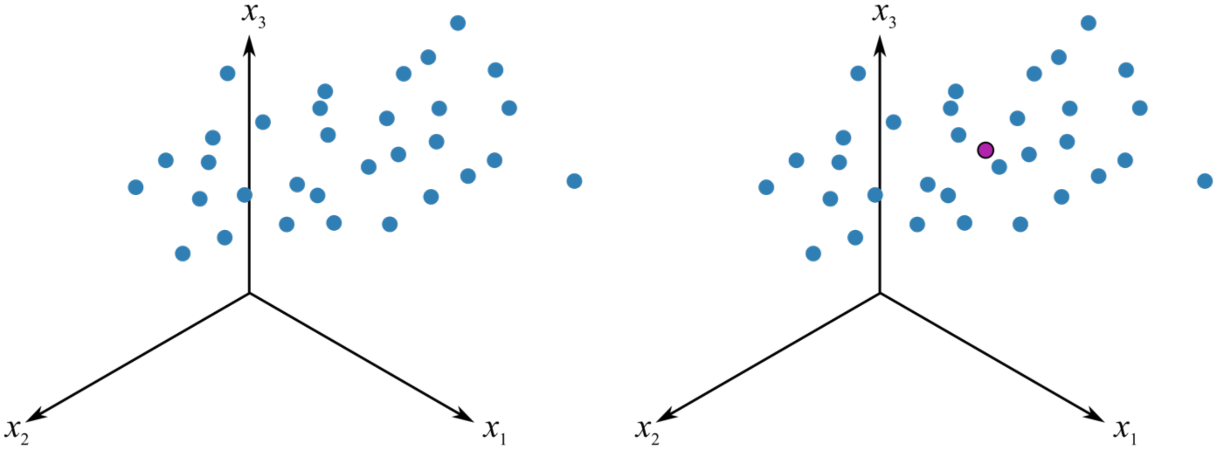


Figure 13 The left coordinate system is the original observation points in a three-dimensional variable space. The right coordinate system illustrates the centre of the data observations (Dunn, 2016).

To identify the first principal component (PC_1), explaining the highest variance, a line is drawn through the data observations that best explains the data. This is a 23-dimensional vector with the minimum residual distance from each observation point, see figure 14 for three-dimensional example.

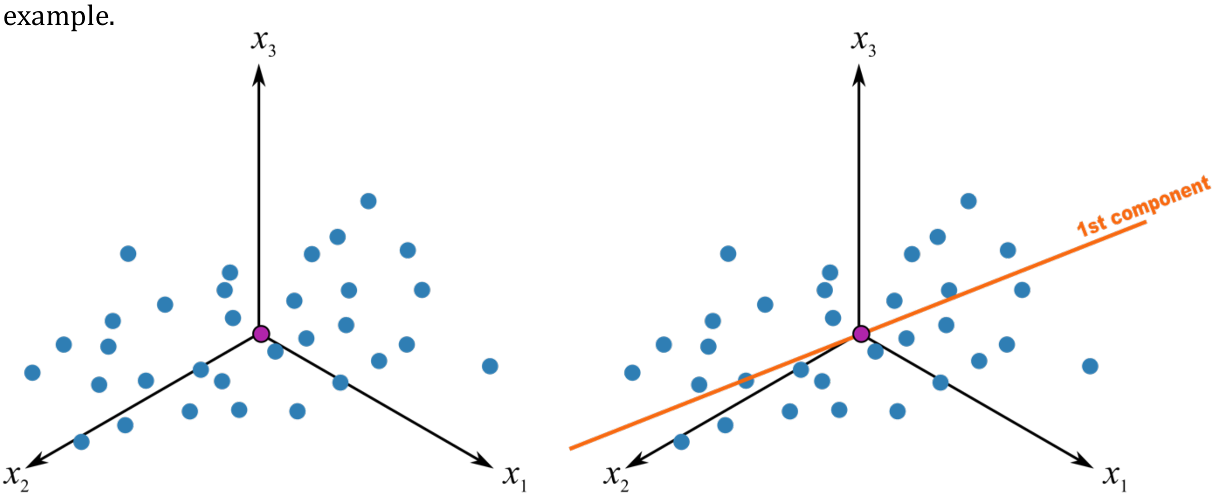


Figure 14 The left coordinate system after data observation centre is moved to the centre of the coordinate system. The right coordinate system with the line explaining PC_1 through the points of data observations and the origin (Dunn, 2016).

The line goes through the origin of the 23-dimensional coordinate system in the direction of maximum variance of the projections on to this line. All 1467 data points are orthogonally projected on the PC_1 vector. The distance from the origin to each projected data point along this line is the score value for each of the 1467 observations for this component. Score values are both positive and negative and observations near average will have score values close to 0.

The second principal component (PC_2) is orthogonal to the first and also goes through origin of the 23-dimensional coordinate system. It is aiming to catch the largest variance that is not caught by PC_1. Orthogonal projection of the 1467 data points to this line is done in the same way as for PC_1, where the distance from origin to each projected point is the score value in the second component, see figure 15 for three-dimensional example.

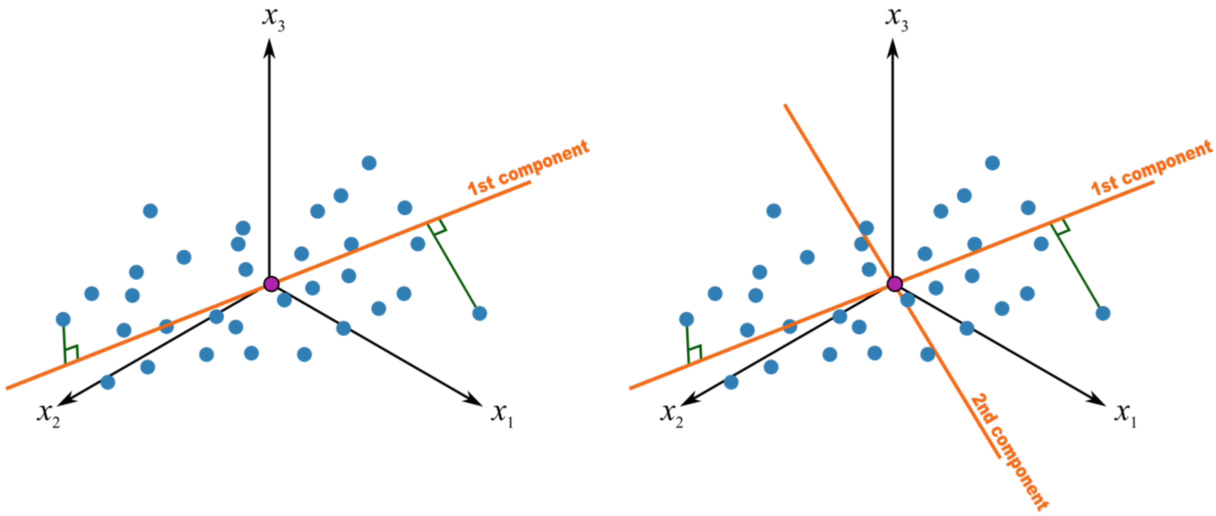


Figure 15 Left is the coordinate system illustration projection of the data observations onto the first principal component. Right is the second component, PC_2, orthogonal to PC_1 (Dunn, 2016).

The two components are creating a plane. With some third or more components the model is defined as a hyper plane, see figure 16. The orthogonal distance from each point on to this plane is the residual distance, or residual error.

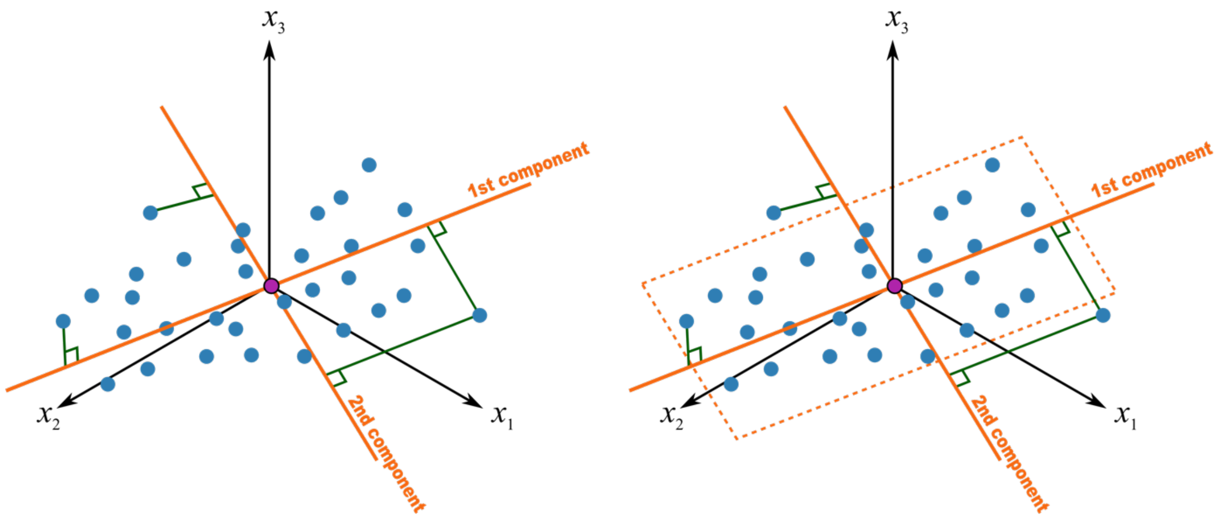


Figure 16 The left coordinate system illustrates projection of observations on to PC_1 and PC_2. Right coordinate system visualizes the plane created by the two components (Dunn, 2016).

According to that the PC matrix is orthogonal and standardized the relationships present in the original data are still present in the new score values. The advantage with the new score values is that it simplifies the process to locate similar observations for further analysing due to a smaller subspace dimension.

***k*-mean clustering**

After finding new components, with eigenvalue above 1, from the 23 food categories, these were clustered by k-mean clustering to identify how the individuals where associated to each new component. This was done by pre-specifying the numbers of clusters to eleven, based on the number of components with eigenvalue above 1 in the PCA.

k-mean clustering is a method presented by MacQueen (1967). It is suitable for large datasets where it is not necessary to calculate the distance measures between all pairs of subjects. The method is aiming to partition the 1467 observations (individuals) into the pre-specified number of eleven clusters to minimize the total within-cluster variance or squared error function, solving equation 10: (MacQueen, 1967; Sayad, n.d.)

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2 \quad \text{(eq. 10)}$$

Where *J* is the objective function, *k* the number of clusters, *n* is the number of observations, x_i the observation *i* and c_j the centroid for cluster *j.*

In k-means clustering it is needed to pre-specify the number of clusters to consider. The idea is to define *k* centres, in this study eleven, spread over the data point space. Then associate each data observation to the nearest centre. After all points are associated to a cluster centre a new centre, centroid, is re-calculated for all clusters, the mean point of all data observations in a cluster. In the next step all data points are associated to their now closest new centroid. This process is repeated until centres are not moved anymore. This creates a separation of the data observations into groups that was saved as a new variable in SPSS for further analysis.

Connecting PCA and k-mean clustering
The subspace of a PCA is spanned by the principal directions of the components and is identical to cluster centroid subspace in k-mean clustering (Xu, 2014; Ding & He, 2004). For example, if the data is distributed between two clusters, the line connecting the two different centroids is the best 1-dimensional projection direction, the first principal component direction (PC_1). Therefore, clustering in the PCA subspace will give a solution close to the global solution (Xu et al, 2014). The score values from PCA can be investigated by clustering to find groups of relationships, as the observations that are similar in the original data observations will cluster together.

GHGE evaluation
After clustering the individuals, which is based on their varieties in diets, the groups were named according to the characteristics of food in each diet group. To avoid impact of biological differences between genders and age, only the energy-adjusted GHGE variable has been used in this analysis. Gender and age are the two variables that affect energy intake due to biological reasons (Livsmedelsverket, 2017). One-way ANOVA and post hoc Tukey 's test was used to evaluate if there were any statistically significant differences, on the level of α=0.05, between the groups GHGE mean.

One-way ANOVA
One-way analysis of variance (ANOVA) is using the F-distribution to find if there is a statistically significant difference between two or more independent group means (Leard Statistics, 2013b). When comparing k independent groups, a null and alternative hypothesis are constructed. Where the null hypothesis is saying that all the groups' means are equal, and the alternative is saying that there is at least two of *k* means that are not (The Pennsylvania State University, n.d.). The F-statistic, also called the variance ratio, is calculated by dividing the between group variance by the within group variance, also expressed as the mean square (MS) ratio (Newsom, 2013, The Pennsylvania State University, n.d.) in equation 11:

$$F = \frac{MS_{Between}}{MS_{Within}} \text{ (eq.11)}$$

The mean squares (MS) between and within are computed by dividing the sum of squares with respectively degrees of freedom as following equation 12 and 13:

$$MS_{Between} = \frac{SS_{Between}}{df_{Between}} \text{ (eq.12)}$$

$$MS_{Within} = \frac{SS_{Within}}{df_{Within}} \text{ (eq.13)}$$

Where $df_{Between}$ is *k-1* and df_{Within} is *n-k*, n is the total sample size, all groups combined. Sum of squares between groups ($SS_{Between}$) are calculated as following equation 14:

$$SS_{Between} = \sum_k n_k (\bar{X}_k - \bar{X}_{GM})^2 \text{ (eq. 14)}$$

Where n_k is the sample size of group *k,* $\bar{X}_k$ is the sample mean of group k and $\bar{X}_{GM}$ is the grand mean for all groups combined. Sum of squares within groups (SS_{Within}), also called error sum of squares, are calculated as following equation 15:

$$SS_{Within} = \sum_k \sum_i (X_{ik} - \bar{X}_k)^2 \text{ (eq. 15)}$$

Where X_{ik} is a single value in group *k*.

The null hypothesis will be rejected if the F test statistic is greater than the critical value obtained in the table of F-values for the F-distribution. In statistical software such as SPSS the p-value is obtained when preforming a one-way ANOVA (The Pennsylvania State University, n.d.). Whilst the F value is a tool to help answering the question if the variance between two means is statistically different, the p-value is determined by the F statistic and is the probability of that the result could have happened by chance (Statistics How To, 2017). The p-value is compared to the chosen significance level α=0.05.

Tukey's HSD test

One-way ANOVA will only tell if there is at least one of the means differing from the others. To evaluate how many and where the difference lies ANOVA is often followed up by a post hoc test. For pairwise comparison among the means of the groups the post hoc test called Tukey's Honest Significant Difference test (Tukey's HSD test) is widely used (Jones, n.d.; Statistics How To, 2016; The Pennsylvania State University, n.d.). The null and alternative hypothesis for the test is set up, where the null hypothesis states that the mean of group *i* and *j* are equal and the alternative states that they are not.

HSD for each pair of means is calculated with formula in equation 16:

$$HSD = \frac{\bar{X}_i - \bar{X}_j}{\sqrt{MS_{Within}/n_k}} \text{ (eq. 16)}$$

Where $\bar{X}_i - \bar{X}_j$ is the difference between the mean of group *i* and *j.*

The value of HSD is compared to the value found in the Tukey's critical value table where the total number of sample k and the degrees of freedom within the group *n-k* is used to locate the critical value. If the HSD value is larger than the critical value, the null hypothesis is rejected proving the two means are significantly different.

Socio-demographic evaluation

After GHGE evaluation of the diet groups, the same socio-demographic factors used as in the regression analysis; gender, household income, age, education, single household, student, children and region were evaluated in the diet groups by comparison to the total average and distribution of the whole sample in terms of the of socio-demographic variables by one sample t-test and Chi^2-test respectively.

One sample t-test

The one sample t-test is used to determine if the means of the continues socio-demographic variables in each diet group is statistically significant different from the total sample mean of the same variables (Libguides, 2018). The null hypothesis is saying that the group mean is equal to the total sample mean whereas the alternative hypothesis is saying that it is not. The test statistic of the test is calculated by equation 17.

$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}} \text{ (eq. 17)}$$

Where $\bar{x}$ is the group mean, μ is the total sample mean, $s_{\bar{x}} = s/\sqrt{n}$, where s is the groups standard deviation and n is the sample size of the group. The calculated t is then compared to the critical t value from the t distribution table with degrees of freedom, $df=n-1$ and the significance level of $\alpha=0.05$. If the t value is larger than the critical t-value, the null hypothesis is rejected, and the group mean statistically significantly differ from the total sample mean.

Chi^2-test

Chi^2 Goodness-of-fit test is used to determine if the diet groups' distributions of the categorical socio-demographic variables differ statistically significant on the level of $\alpha=0.05$ from the total distribution of the sample of all individuals (StatTrek, n.d.). For the test, the null hypothesis is constructed by saying that the data are consistent with a specified distribution whereas the alternative hypothesis saying it is not. Chi^2 test statistic is calculated by equation 18.

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \text{ (eq. 18)}$$

Where O_i is the observed frequency for each attribute of a socio-demographic variable and E_i is the frequency for the total sample. The E_i is the sample size of observations for each attribute in a socio-demographic variable multiplied by the proportion of the attribute in the total sample (MiniTab, 2016). The resulting value of the test statistic, χ^2, are compared to X that follows chi^2-distribution with $k-1$ degrees of freedom, where k is the number of attribute. The p-value is calculated as the probability of that X is larger than the test statistic χ^2. If the p-value is below the significant level of $\alpha=0.05$ the null hypothesis is rejected and the distribution in the diet group is statistically significantly different from the distribution of the total sample.

3.3.3 GHGE intensive ruminant meat analysis

This analysis was carried out to examine the tenth hypothesis (H10), that reflects both research questions: if inclusion of the specific food group of GHG intensive *fine ruminant meat* (*FRM*) in a diet associates to a higher total diet GHGE and higher household income. Individuals that registered any of following *FRM,* such as: salted ruminant meat, smoked ruminant meat, sirloin steak, Scotch fillet, veal, roast beef, ruminant steak and brisket of beef was grouped together. This group was compared with the ones not registered any of the mentioned meat categories according to energy-adjusted GHGE and household income. The groups' mean was compared through independent-samples t-test on significance level of $\alpha=0.05$. It was also studied if the higher intake of these products correlated to increased household income.

Independent samples t-test
The independent samples t-test is a statistical test to determine whether there is a statistically significant difference between the means of two groups (Leard Statistics, 2013c). The t-value is obtained through equation 19 (UC Davis, n.d.).

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\left(\sum X_1^2 - \frac{(\sum X_1)^2}{n_1}\right) + \left(\sum X_2^2 - \frac{(\sum X_2)^2}{n_2}\right)}{n_1 + n_2 - 2} \times \left[\frac{1}{n_1} + \frac{1}{n_2}\right]}} \text{ (eq. 19)}$$

Where $\bar{X}_1 - \bar{X}_2$ is the difference in mean between the two groups, X_1, X_2 is the score in group 1 respectively 2 and n_1 and n_2 is the sample size of respectively group 1 and 2.

The t-value is compared to the critical values for t in the abbreviated table for t-test to find the probability value (p-value). To identify the p-vale associated to the obtained t, the degrees of freedom (*df*) according to equation 20 is needed.

$$df = n_1 - 1 + n_2 - 1 \text{ (eq. 20)}$$

If the t-value is larger than the critical value found in the table at the pre-decided level of α=0.05 the mean of the two groups are significantly different at the level of α=0.05.

Effect size
Validation of the effect size of the differences in GHGE and household income was preformed through calculations of Hedges' g value. Hedges' g value is appropriate to use as an effect size measure as an alternative for Cohens' d, that is widely used to decide effect size, if the sample size of the two groups are different (Socscistatistics, n.d.). It is weighted according to the size of each sample and is calculated according to equation 21 (Ellis, 2009):

$$Hedges'g = \frac{\bar{X}_1 - \bar{X}_2}{SD^*_{pooled}} \text{ (eq. 21)}$$

Where $\bar{X}_1 - \bar{X}_2 = difference\ in\ mean$ is the difference in mean, $SD^*_{pooled} = \sqrt{\frac{(n_1-1)SD_1^2 + SD(n_2-1)_2^2)}{n_1+n_2-2}}$ is the pooled and weighted standard deviation and n_1 and n_2 is the sample size of group 1 respectively 2.

An effect size of 0.2 indicates a small effect, 0.5 a medium and 0.8 a large effect (Lane et al., n.d.).

4. Results

The result chapter is divided into three sections that correspond to each of the preformed analyses. It starts with the linear multiple regression analysis that aimed to create models of how the total composition of diets, according to variation in food-related GHGE, could be explained by socio-demographic variables. The second section are the results from the cluster analysis that clustered individuals, based on characteristic food groups in their total diet composition, to create characteristic diet groups for comparison of food-related GHGE, and to find if those diet groups could be associated to certain socio-demographic features. The chapter ends with the analysis of GHG intensive *FRM* correlates to total food-related GHGE of a diet and household income.

4.1 Explaining food-related GHGE variation by socio-demographic variables

In this section, the outcome for the model explaining the variation in crude food-related GHGE is presented first and the model for the energy-adjusted food-related GHGE as second.

4.1.1 Crude food-related GHGE model

Inclusion of four independent variables: gender, age, household income and student gave the best fit for explanation of the variation in the crude food-related GHG emissions from diets. However, the R^2_{adj} is low, indicating that only 22 % of the variation is explained by this model, see table 4.

Table 4 Modell Summary presents the coefficient of determination for each model achieved from the stepwise regression and the standard error of the estimate with crude annual GHGE as dependent variable.

Model	R	R square	Adjusted R Square	Standard Error of The Estimate
Gender	.431	.186	.185	.635
Gender, Age	.459	.211	.210	.625
Gender, Age, Household income	.468	.219	.217	.622
Gender, Age, Household income, Student	**.471**	**.222**	**.220**	**.621**

The independent variables gender, age, household income and student were all significant on the level of α=0.05, see table 5.

Table 5 Coefficient table presents the coefficients of each independent variable in the model, the standardized coefficient, t-value, the significance value and the collinearity statistics by tolerance.

Variables	Unstandardized Coefficients β	Standardized Coefficients $\hat{\beta}*$	t	Sig.	Tolerance
(Constant)	0.796		2.42	.016	
Gender	.615	.433	18.618	.000	.992
Age	-.008	-.189	-7.344	.000	.809
Household income	0.091	.083	3.580	.000	.987
Student	-.164	-.062	-2.416	.016	.800

The unstandardized β-coefficients in table 5 reveal that, in relation to the average individual food-related GHGE of 1.8 tonnes CO_2 eq./ year, the effect of the attribute man compared to woman will generate 33% higher emissions of GHG from food (0.615 tonnes CO_2 eq./ year). Being a student would decrease the food-related GHGE by 9% (-0.164 tonnes CO_2 eq./ year). It also indicates that a 50 % rise in income will increase the food-related GHGE by 3 % (0.045 tonnes CO_2 eq. / year) and the rise in age by 1 year corresponds to a decrease of approximately 0.4 % (-0.008 tonnes/year). The standardized coefficient $\hat{\beta}*$ also confirms that the strongest effect on the variation in the food-related GHGE is gender with moderate effect, as $\hat{\beta}*$ are in-between 0.2 and 0.5. The other variables have weak effects on the food-related GHGE, as $\hat{\beta}$ are below 0.2. There is no strong correlation (collinearity) between any of the independent variables in the model, indicated by the tolerance value > 0.25 for all variables.

4.1.2 Energy-adjusted food-related GHGE model

Inclusion of four independent variables: gender, age, household income and single household gave the best fit for explanation of the variation in GHG intensity in diets by the energy-adjusted food-related GHGE. It is an even weaker model than the one explaining the food-related GHG emissions with the crude GHGE, as only 4 % of the variation in GHGE is explained, see table 6.

Table 6 Presents the coefficient of determination for each model achieved from the stepwise regression and the standard error of the estimate with energy-adjusted GHGE as dependent variable.

Model	R	R square	Adjusted R Square	Standard Error of The Estimate
Gender	.149	.022	.022	.530
Gender, Household income	.186	.035	.033	.527
Gender, Household income, Age	.198	.039	.037	.526
Gender, Household income, Age, Single household	**.205**	**.042**	**.039**	**.526**

The independent variables gender, age, household income and single household were all significant on the level of α=0.05, se table 7.

Table 7 Coefficient table presents the coefficients of each independent variable in the model, the standardized coefficient, t-value, the significance value and the collinearity statistics by tolerance.

Model	Unstandardized Coefficients β	Standardized Coefficients $\hat{\beta}*$	t	Sig.	Tolerance
(Constant)	0.672		2.451	.014	
Gender	.156	.144	5.585	.000	.991
Household income	0.081	.097	3.598	.000	.903
Age	-.002	-.060	-2.301	.022	.980
Single household	.078	.056	2.077	.038	.892

The unstandardized β-coefficients in table 5 reveal that, in relation to the average individual food-related GHGE of 1.8 tonnes CO_2 eq./ year, the effect of the attribute man compared to woman will generate 9% higher emissions of GHG from food (0.156 tonnes CO_2 eq. /year). Individuals living in single household lowers food-related GHGE by 4% (-0.078 tonnes CO_2 eq.

/year). It also indicates that a 50 % rise in income will increase the food-related GHGE by 2 % (0.04 tonnes CO_2 eq. / year) and the rise in age by 1 year corresponds to a decrease approximately 0,1 % (-0.002 tonnes CO_2 eq./year). The standardized coefficient $\hat{\beta}^*$ confirms that all variables have weak effects on the variation in food-related GHGE as $\hat{\beta}^*$<are below 0.2 for all variables. There is no strong correlation (collinearity) between any of the independent variables in the model, indicated by the tolerance value > 0.25 for all variables.

4.2 Clustering of diet groups

From the PCA analysis, 23 main food groups were reduced into eleven components. The eigenvalue- and rotated component matrixes are found in the Appendix. After clustering of the components from the PCA the similarity in three respectively two clusters according to food composition merged those clusters and resulted in eight diet groups that distinguished according to characteristics in food consumption.

The groups were Vegan, with distinctive consumption of vegetarian meat substitute, vegetarian milk products, nuts and seeds, vegetables and mushrooms. A group named Pescotarian[8] according to that this group had the highest consumption of fish and seafood. Except from that the individuals mainly had higher consumption of cheese, vegetables, mushrooms, nuts, seeds and vegetarian milk products in comparison to the other groups. Then there was a group named Cheese sandwich due to the characteristics for this group was cheese, fat (oil, butter, margarine) and grain products such as bread, even poultry was characteristic for this group. The next group was given the name Spaghetti Carbonara with the characteristics of mixed meat, dairy products (not including cheese and yoghurt), pasta and rice. The fifth group was named Junk food due to the characteristics of pizza, hamburger and snacks. The sixth group was named Full English, short for Full English Breakfast, due to the distinctive consumption of egg, sausage and pork. The Steak eaters was a group named after their characteristic consumption of ruminant meat and dairy products, not cheese and yoghurt included. The last group was named Swedish Svensson; it was the largest group with a mix of food that was not characterised in any special way.

In table 8 the final diet groups with characteristic names, respectively number of individuals and main types of food are presented. Contribution of the 23 main food types to the eight diet groups in table 8 are found in the Appendix.

[8] Pescotarian is someone who includes fish and seafood in their vegetarian diet (Collins English Dictionary, n. d).

Table 8 Characteristic group names for the eight diet groups, number of individuals and main types of food.

Diet Group	Group name	Number of individuals	Main types of food
1	Vegan	12	Vegetarian substitute Vegetarian milk products Nuts/seeds Vegetables/mushrooms
2	Pescotarian	189	Cheese Vegetables/mushrooms Nuts/seeds Vegetarian milk products Fish and seafood
3	Cheese sandwich	179	Cheese Fat (oil, butter, margarine) Grain products (bread etc.) Poultry
4	Spaghetti Carbonara	58	Mixed meat Dairy products (not including cheese and yoghurt) Pasta/Rice/Grain
5	Junk food	115	Pizza/ Hamburger Snacks
6	Full English	178	Pork Egg Sausage
7	Steak eaters	78	Ruminant meat Dairy products (not including cheese and yoghurt)
8	Swedish Svensson	658	Mixed

4.2.2 Food-related GHGE evaluation among diet groups

The one-way ANOVA revealed that there was a statistical significant difference between some of the diet groups at the level of α=0.05 (F= 26.782, p= 0.000). Only two diet groups; Vegan and Steak eaters did show a statistically significant difference in food-related GHGE from the other groups according to post hoc Tukeys' test at the level of α=0.05. The Vegan group had significantly lower GHGE than all other groups (1.16 ± 0.38 tonnes CO_2 eq./ year). Although this turns out to be a small group with less than 1‰ of all individuals included and therefore tells very little about the whole sample. The Steak eaters had significantly higher GHGE than all other groups (2.50 ±0.65 tonnes CO_2 eq./year, p= 0.000). The other six diet groups did not show statistically significant differences in mean values on the level of α=0.05. A boxplot of the groups energy-adjusted individual GHGE is provided in figure 17. The red line marks the total mean GHGE of 1.8 tonnes CO_2 eq./year, person

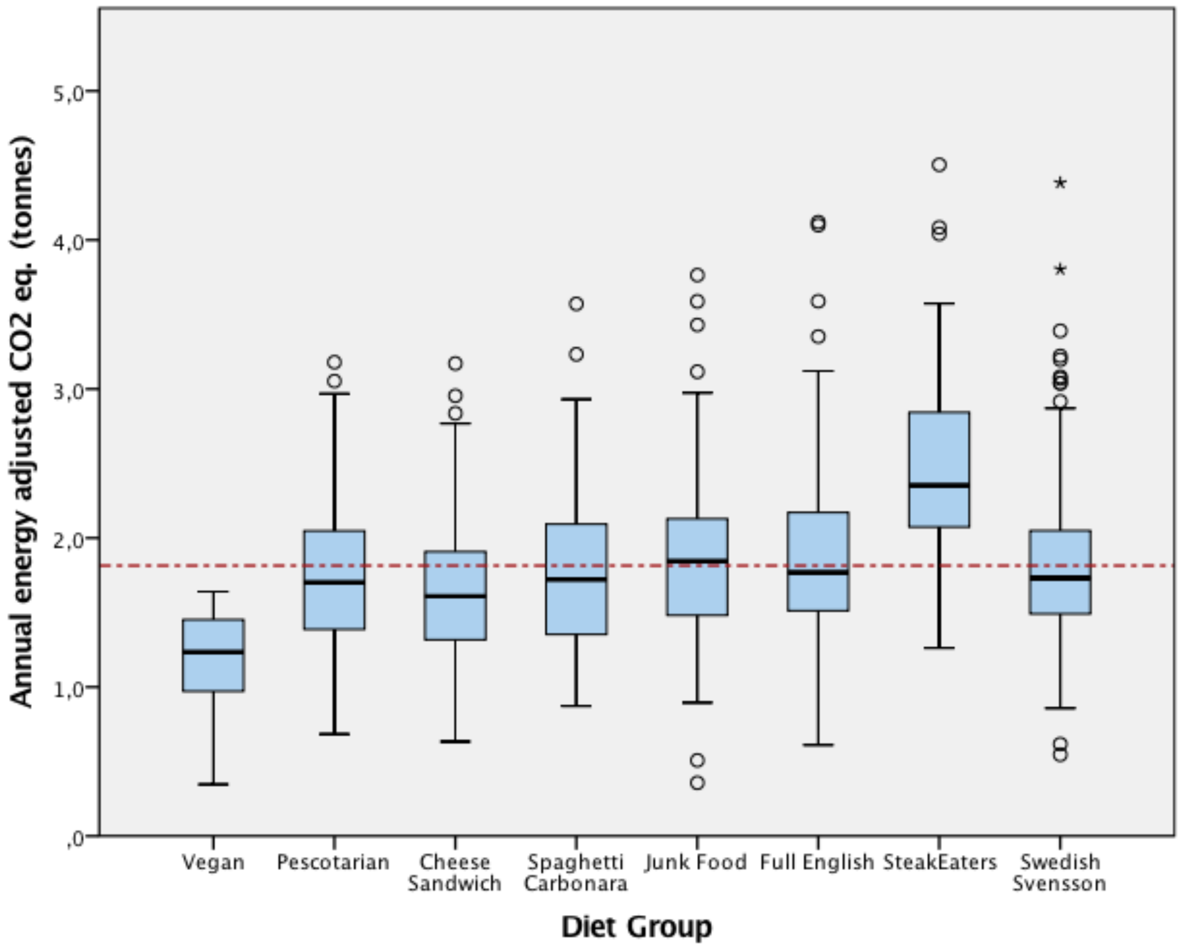


Figure 17 Box-plot of the eight diet groups energy-adjusted GHGE in tonnes CO_2 eq./ year and the mean marked by the red line. The boxes indicate the 25th to 75th percentiles in respective group. As can be seen the Vegan and Steak eater group has the whole box below respectively above the mean value, meaning that at least 75% of the group members have GHGE values that are lower respectively higher than the mean emissions for the whole group. The circles are extreme values and the asterisks the extreme outliers, representing values more than three times the height of the box.

4.2.1 Socio-demographic analysis

There was a prominent difference between the diet groups regarding socio-demographic variables. Below is a summary of the socio-demographic variables for the eight diet groups that statistically significantly differ on the level of α=0.05 from the average mean and distribution of the whole sample. All numbers and percentage distribution are found in Appendix and table 9 and 10 are summaries of the alterations for all eight diet groups according to each demographic variable.

- The Vegan group has the lowest household mean income and the lowest mean age.
- Pescotarian are to the majority females and includes many high-educated individuals, low educated are rare in this group. In this group it is common to live in a single household.
- The Spaghetti Carbonara group has a lower average age than the whole sample
- The Junk food group has low household mean income and tend to be young individuals. It also has a higher share of males than in the total sample.

- Full English are typically males and more likely to live in rural areas than the other groups.
- The Steak eaters are typically males.

The Cheese sandwich group and the Swedish Svensson does not characterise except from higher share of males among the Cheese sandwich group and females in the Swedish Svensson group in comparison to the average shares of the whole sample.

Table 9 summarises the alteration from the total mean for each diet group for the continuous socio-demographic variables.

Table 9 The alteration from total mean of the continuous variables household income and age for each diet group. Red colouring and asterisks indicates statistically significant difference higher than total mean of the sample on the level of α=0.05.

Diet group	Household income sek	Age years
Vegan	-141 789* -34%	-16*
Pescotarian	-16 965 -4%	2
Cheese sandwich	-16 590 -4%	0
Spaghetti Carbonara	-6 935 -2%	-5*
Junk Food	-59 035* -14%	-11*
Full English	15 532 4%	0
Steak eaters	43 994 10%	-1
Swedish Svensson	13 550 3%	1

Table 10 summaries the percentage difference that each group diverges from the total share for each of the socio-demographic dummy variables.

Tabell 10 The percentage difference from total share for the dummy variables gender, education, region, single household, children and student differs from total distribution. Colouring and asterisks indicates statistically significant differences from total distribution on the level of α=0.05: green indicates higher values and red indicates lower.

Diet group	Gender %		Education %			Region %		Single house hold %	Childre n %	Student %
	M	F	Low	Med	High	U	R			
Vegan	-41	31	-	41	-11	26	-70	-6	19	2
Pescotarian	-37*	27*	-57*	-15*	25*	2	-7	50*	-10	-10
Cheese sandwich	31*	-23*	-17	-3	8	2	-7	-23	5	-10
Spaghetti Carbonara	1	-1	4	-	-1	4	-10	-6	26	15
Junk Food	41*	-30*	-	-	-	12	-33	5	14	15
Full English	71*	-53*	39	5	-13	-15*	42*	-25	14	-23
Steak eaters	50*	-37*	-13	2	-	-	-	-28	21	3
Swedish Svensson	-30*	22*	13	2	-5	-	-	5	-10	15

M= male, F=female, U=urban and R=rural

4.3 Ruminant meat in diet and household income

Including any type of GHG intensive *fine ruminant meat* (*FRM*) such as: salted ruminant meat, smoked ruminant meat, sirloin steak, Scotch fillet, veal, roast beef, ruminant steak and brisket of beef in diet revealed to be an indicator on high total diet GHGE. 15 % (n= 219) of the individuals included some sort GHGE intensive *FRM* in their diet. This group has a statistically significantly higher energy-adjusted food-related GHGE mean at a level of α=0.05 (2.199± 0.569 tonnes CO_2 eq./year) than the ones not including any *FRM* (1.747± 0.500 tonnes CO_2 eq./year). The difference between the groups of individuals is substantially large according to Hedges' g value (g=0.88). See figure 18 for boxplots visualizing the difference between the groups of food-related GHGE in tonnes CO_2 eq./year.

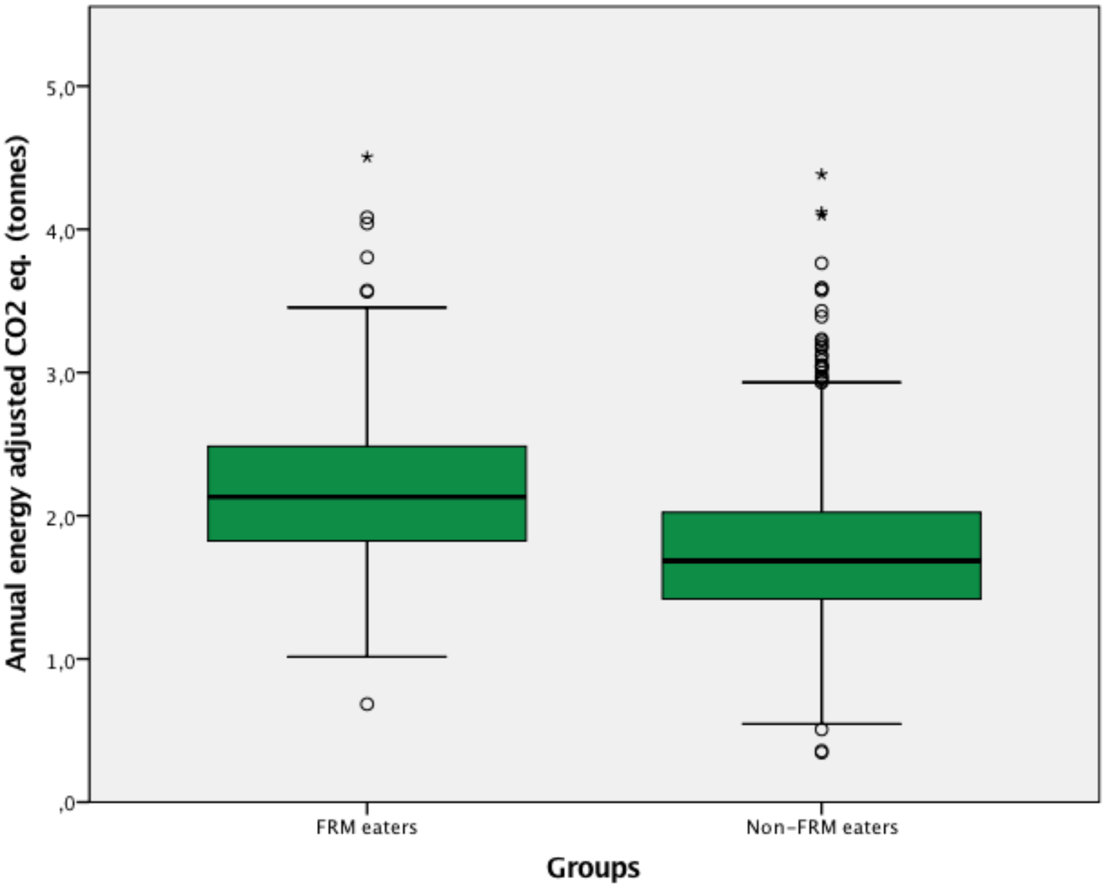


Figure 18 Boxplots of ruminant meat eaters and non-ruminant meat eaters comparing energy-adjusted GHGE in tonnes/year. The circles are extreme values and the asterisks the extreme outliers, representing values more than three times the height of the box

The ruminant meat eaters also have a statistical significantly higher mean income) at a level of α=0.05 (472 037± 274 805 sek/year compared to the others (410 424 ± 243 816 sek/ year) though the difference between the groups mean according to Hedges' g was small (g=0.25). The ruminant meat eaters' household income is approximately 13 % above the mean income of the sample and the non-ruminant meat eaters 2 % below. There was no linear correlation between increased household income and the amount of *FRM* consumed.

5. Discussion

In this chapter, the notable findings from the analyses will be discussed according to the research questions and previous research. The methods and data used will be discussed and it will end with some thoughts about further research and development from the findings.

5.1 Findings

RQ1 *What background factors may explain the variation in GHGE from food?*

The results of the multivariate linear regression analysis indicate that gender is a factor that explains part of the variation in food-related GHG emissions and GHG intensive food in diets. Both models in the regression analysis have a low value on their coefficient of determination, indicating that the models only explain a small part of the variation in GHGE (4 % with the energy-adjusted data and 22% with the crude).

Confirmation of men having higher food-related GHG emissions than women is in line with Treu (2017) and Räty & Carlsson-Kanyama (2010), who claim that women eat smaller quantities of GHGE intensive meat than men. This result also confirms the hypothesis of that men have higher food-related GHG emissions than women. The norm of meat association to masculinity (Loughnan, 2014) and that vegetables and fruits represent the idea of femininity (Roos et al., 2001) is in line with the outcome and hypothesis of men having more GHG intensive food in their diets. The fact that differences in calorie intake is related to gender of biological reasons (Livsmedelsverket, 2017) explains why this variable has a substantially stronger effect in food-related GHG emissions (the crude model) than on the composition of GHG intensive food in diets (the energy-adjusted model); 0.615 tonnes per year is the difference between men and women in the crude model compared to 0.156 tonnes difference per year in the energy-adjusted. This is a difference of 35 respectively 10 % in relation to the average food-related GHGE of approximately 1.8 tonnes CO_2 eq./year for both variables. Hence, the biological differences are more important for the variation in GHGE from diets than other differences between the genders related to content of diet. The same conclusions can be drawn from the age variable where large age differences are shown to affect the variation in food-related GHGE but where the biological differences have major stronger impact on the food-related GHGE than age affects the content of GHG intensive food in a diet. 30 years increase in the crude model will lower GHGE by approximately 12% in relation to the average emissions and the same increase in the energy-adjusted model lower emissions by 3%.

The other socio-demographic variables from the regression models indicated to have a minor or none effect on the variation in food-related GHG emissions and in GHG intensive food in diets. Though, as suggested in the hypothesis H9, the mean comparison between the *FRM* eaters and the non-*FRM* eaters that showed a large significant difference in GHG intensity in diets between the groups also indicated a small but significant difference in household income, as those who includes any *FRM* in their diet had slightly higher mean income than those who did not. Although, there was no significant association found between income and intake of those ruminant meat products. The result is kind of in line with the economic effect in the regression models where household income tended to have a weak, but significant, contribution to the variation in GHGE in both regression models. The results showed that a 50 % increase in household income only would contribute to approximately 2-3 % increase in emissions due to both models.

Though the effect of income is small, the result is in contrast to Clonan (2016) who suggests that people with lower household income in richer countries consume more red meat than people with higher household income. One reason why the effect of income is small and only explains parts of the variation in food-related GHGE might be related to the low price development in meat as well as food in general in comparison to the income development (Jacobsson, 2011; SCB, n.d). This generates a more widespread ability for individuals with different levels of income to consume

GHG intensive meat. Although, the fact remains that individuals with higher income tend to spend more money on food (SCB, 2004) but then not necessary GHGE intensive meat. There is for example a suggestion that richer people are consuming more "healthy" food (Gazan, 2016), which may not essentially equal GHG intensive products.

RQ2 *Are there variation in GHGE from different diet patterns? And if so can these be associated to socio-demographic background factors?*

This question was mainly treated in the cluster analysis where two diet groups had statistically significant different food-related GHGE. The Vegan group, a group of mainly vegetable-based diets, was contributing to the lower GHGE and the Steak eaters, that was characterised as high consumers of ruminant meat and dairy products, was contributing to the higher GHGE. Even though both these groups were small and covered less than 1‰ and 5 % respectively of the total number of individuals in the study, they do give some information about the extremes in the large variation of GHGE.

The characteristics of men in the groups with high GHGE where in line with the result from the regression analysis carried out in the first part of this study, and so with Treu (2017) and Räty & Carlsson-Kanyama (2010) as discussed above. The characteristics of a high share of young individuals in the Vegan group confirm Allés et al. (2017) and Bedford & Barr (2005). These same studies also indicate that vegetarians in relation to meat-eaters have lower income, which this cluster analysis also implies. Higher income was not statistically significant for the Steak eater group, which is in opposition with the global perspective suggested in (Johnston et al, 2014) and also the findings in Amcoff et al. (2012). It is not in line with the economical tendencies as registered in the household income mean comparison between consumers of *FRM* and non-*FRM* eaters and the small positive effect in the regression analysis in this study. It was not statistically significantly confirmed that higher education relates to the lower GHGE diet, as proposed by SCB (2004) and Hoek et al. (2003). The cluster analysis did neither support the idea of Hoek et al. (2003) that vegetarians are more likely to live in urbanised areas. The small size of the Vegan cluster is a likely reason for not showing significant difference according to the total average distribution, as eleven of twelve individuals in the Vegan group registered that they live in urban areas.

5.2 Limitations and error sources

5003 individuals were invited to participate in *Riksmaten 2010-11* but only 1 797 (36%) completed the food registration and was included in this study (Amcoff et al. 2014). As a matter of fact, certain socio-demographic groups had a low representation in the final sample. For example, 66% of the invited individuals with higher education participated, whereas the frequency among the invited individuals with low education was only 36%. This is worth highlighting before drawing the final conclusions. This could mean that information about low educated individuals is missing in the analysed data.

In this study the total GHGE from a persons' diet was approximated for one year based on 4 days diary writing. This could give a wrong reflection of actual eating patterns due to knowledge of being documented might generate a different food consumption during these days. One issue with relying a study on people filling in and registering their own habits and consumption is also the tendency to answer what they think the questioner would like to hear. *Riksmaten 2010-11* is a study made by Swedish National Food Agency (Amcoff et al., 2012) on food habits and there are reasons to believe that people might want to look good and healthy in their answers. Though, the effect of this assumption has probably not been a major issue for the purpose of this study. Hypothetically people would rather cut down on declaring food types considered unhealthy for the majority of people such as sweets, snacks, pizza rather than the GHGE intensive ruminant meat and dairy products.

This study only carried out modelling of linear relations in the regression analysis. More complex non-linear relations have not been explored.

Cluster analysis was used in this study to find high consumption of certain food categories among the individuals to be able to create characteristic diets. For the clustering, PCA was used as a "backstage" tool. The clustering was performed on the individual "factor scores" constructed for each component received from the PCA. It is worth mentioning that this is not an established widely used method but has been an explorative way in this study to construct groups of individuals with similar eating patterns. Since the factor scores are based on the sum of the products of the factor loadings multiplied with the reported caloric intake of the specific food category, the clustering will group individuals with high intake of a certain food category together. Why clustering was not performed on the 23 food categories directly was mainly according to two reasons. Firstly, the number of clusters needs to be pre- specified and for that, some indication of the number of clusters was received from the PCA. Secondly, 23 variables showed to be too many variables for relevant and informative clustering. Hence, instead of excluding food categories by hand, the PCA was a helpful tool to instead combine the food categories to new, fewer components. To ensure the reliability of this explorative method, the mean calorie intake in the resulting clusters where checked for all 23 food categories

Exclusion of outliers in data was done for 330 under-reporters. Due to the low number of identified over-reporters, three individuals, and inclusion of these in previous studies (Sjörs, 2017; Amcoff, et al, 2012), as well as adjustment for energy intake, they were left remaining. This could though be questioned for the crude GHGE data. But according to that energy-adjusted GHGE data was used in all analyses, extreme values corresponding to high intake was adjusted and not considered an issue.

5.3 Future research and development

To my knowledge, this is the first study of how socio-demographic factors are connected to variation in food-related GHGE in Sweden. The result is therefore of interest to further research in this field. This study concludes that socio-demographic factors such as gender, and household income can only explain a very small share of the variation in individual' food-related GHGE. The question of what, or if any other features may explain the large variation remain. The reasons behind the choices of what we eat are possibly more connected to taste and habits as well as psychological aspects of the individual human being. Which has not been treated in this study but is known to influence consumption behaviour (Kasi, n.d). Recommendations for further research in this field are interdisciplinary studies that include more complex behavioural theories and research linked to psychology. Hence, this is a first indication of that the variation in food-related GHGE possibly have other underlying explanations and can not only be understood by socio-demographic variables.

The food data used in the thesis was collected eight years ago, between 2010 and 2011. It would be interesting to carry out a similar study with more present data to get an idea of change over time and the present status. More present data could hypothetically generate another outcome from the analyses due to changed patterns in food consumption and market supply. Especially in relation to the increased popularity of vegetarian products (Damberg, 2017) (TT, 2017) as well as changes in the consumed types of meat types (Jordbruksverket, 2017). Increased awareness in society in the area of how food consumption contributes to the climate change, as well as health and animal rights may all be reasons for the growing popularity and supply of vegetarian food (TT, 2017). These changes may possibly be related to psychological factors rather than socio-demographic. Although, the changed trends may be more frequent in certain socio-demographic groups.

6. Conclusions

The purpose of this study was to analyse the role of socio-demographic factors in relation to the variation in Swedes' food-related GHGE. Two main research questions have been answered. Firstly; *What background factors that may explain the large variation in GHGE from food?* Secondly; *Are there variation in GHGE from different diet patterns and if these patterns may be associated to certain socio-demographic factors?*

The conclusions of this thesis are the following:

- Part of the variation in food-related GHG emissions and of GHG intensive food in a diet can be connected to gender.

- Large age differences have effect on part of the variation in food-related GHG emissions.

- Though, most of the variation cannot be explained by socio-demographic factors with statistical significance.

- It has been confirmed through cluster analysis that two smaller groups of characteristic diet patterns could be distinguished to have significantly lower and higher GHG-intensities. Lower GHG intensive diets, characterized by vegetable-based food products, was associated with a group that distinguished by being younger individuals with lower household income. Higher GHG intensive diets, characterised by ruminant meat and dairy products, were shown to be associate with a group with a male majority. Though, these groups were small, incorporating less than 1‰ and 5 % of the individuals respectively.

- It was confirmed that individuals including any GHG intensive *fine ruminant meat* (*FRM*), in terms of salted ruminant meat, smoked ruminant meat, sirloin steak, Scotch fillet, veal, roast beef, ruminant steak and brisket of beef, had diets with a higher GHG-intensity than those that did not consume any of these ruminant meat categories.

- It was also confirmed that those who included any of these *FRM* types had a slightly higher household income than those who did not.

Summing it all up, the overall conclusion closing this thesis is that the large variation in food-related GHGE could not be explained through the socio-demographic factors of gender, household income, age, student, children, single household and region. Though, gender, age and household income, to a low extent, have been confirmed to have a small effect on parts of the variation in the food-related GHG emissions and GHG intensive diets among the Swedish population. The explanations behind the large variations in food-related GHGE are instead suggested to rely on more complex psychological aspects within the individual, involving ethics, believes, opinions, politics and trends, as well as taste and habits. The findings are suggested to be used for further development involving interdisciplinary research that includes psychological aspects of behaviour theories.

7. References

Acock, A. C. (2014). *A Gentle Introduction to Stata* (4th ed.). Texas: Stata Press.

Al-Adili, S & Aronsson, D. (2014) *Vuxnas matvanor I hushåll med eller utan barn, En kvantitativ studie avseende livsmedelsval och näringsintag samt följsamhet av Livsmedelsverkets kostråd bland vuxna I hushåll med eller utan barn* (Master thesis, Uppsala University) Available at: https://www.diva-portal.org/smash/get/diva2:785269/FULLTEXT01.pdf

Allés, B, Baudry, J., Méjean, C., Touvier, M., Péneau, S., Hercberg, S., Kesse-Guyot, E. (2017) Comparison of Sociodemographic and Nutritional Charachteristics between Self-Reported Vegetarians, Vegans, and meat-Eaters from the NutriNet-Santé Study. *Nutrients 2017, 9,* 102. doi: 10.3390/nu9091023

Amcoff, E., Edberg, A., Enghardt Barbieri, H., Lindroos, A. K., Nälsén, C., Pearson, M., Warensjö Lemming, E. (2012) *Riksmaten-Vuxna 2010-11, Livsmedels- och näringsintag bland vuxna I Sverige.* Uppsala: Livsmedelsverket. (*Riksmaten-Vuxna 2010-11)*

Amcoff, E., Edberg, A., Enghardt Barbieri, H., Lindroos, A. K., Nälsén, C., Pearson, M., Warensjö Lemming, E. (2014) *Riksmaten-Vuxna 2010-11, Livsmedels- och näringsintag bland vuxna I Sverige.- metodrapport, Report nr: 8, 2014.*

Bedford, J. L. & Barr, S. I., (2005). Diets and selected lifestyle practices of self-defined adult vegetarians from a population-based sample suggest they are "health conscious". *International Journal of Behavioural Nutrition and Physical Activity 2005, 2* (4). doi:10.1186/1479-5868-2-4

Berglund, T. (2011) *Dimensionsanalys: Principal Component Analysis*. Sociologiska institutionen, Göteborgs universitet 2011

Biesbroek, S., Bueno-de-Mesquita H. B., Peeters, P. HM., Verschuren M. VM., van der Schouw, Y. T., Kramer, G. FH., … Temme, E. HM. (2014) Reducing our environmental footprint and improving our health: greenhouse gas emission and land use of usual diet and mortality in EPIC-NL: a prospective cohort study, *Environmental Health 2014,* 13:27, doi: 10.1186/1476-069X-13-27

Black, A.E. (2000). Critical evaluation of energy intake using the Goldberg cut-off for energy intake:basal metabolic rate. A practical guide to its calculation, use and limitations. *Int J Obes Relat Metab Disord 24,* 1119.

Bouvard, V., Loomis, D., Guyton, K. Z., Grosse, Y., Ghissassi, F. E., Benbrahim-Tallaa, L, … Straif, K. (2015). Carcinogenicity of consumption of red and processed meat. *The Lancet, Ocology, 16* (16), 1599-1600, doi: 10.1016/S1470-2045(15)00444-1

Bryngelsson, D., Wirsenius, S., Hedenus, F., Sonnesson, U. (2016) How can the EU climate targets be met? A combined analysis of technological and demand-side changes in food and agriculture. *Food Policy*, *59*, 152-64

Clonan, A., Roberts, K. E. & Holdsworth, M. (2016). Socio-economic and demographic drivers of red and processed meat consumption: implications for health and environmental sustainability. *The Proceedings of the Nutrition Society, 75*(3) 367-373. doi: 10.1017/S0029665116000100

Davey, G. K., Spencer, E. A., Appleby, P. N., Allen, N. E., Knox, K. H., Key, T. J. (2002). EPIC-Oxford: lifestyle characteristics and nutrient intakes in cohort of 33 883 meat-eaters and 31 546 non-meat-eaters in the UK. *Public Health Nutrition;* 6(3), 259-268. doi: 10.1079/PHN2002430

Davis, J., Sonnesson, U., Baumgartner, D. U. & Nemecek, T. (2010) Environmental Impact of Four Meals with Different Protein Sources: Case Studies in Spain and Sweden. *Food Research International 43,* 1874-1884, doi**:** 10.1016/j.foodres.2009.08.017

Davis, J., Wallman, M., Sund, V., Emanuelsson, A., Cederberg, C., Sonesson, U. (2011) *Emissions of greenhouse gases from production of horticultural products. Analysis of 17 products cultivated in Sweden,* SIK- The Swedish Institute for Food and Biotechnology, Gothenburg, Sweden, SR 828, ISBN 978-91-72-90-301-2

Ding, C. & He, X. (2004) K-means Clustering via Principal Component Analysis, *Proc. of Int'l Conf. Machine Learning (ICML 2004)*, 225–232. Available at: http://ranger.uta.edu/~chqding/papers/KmeansPCA1.pdf

Field, A (2009*) Discovering Statistics using SPSS*, Third edition, SAGE.

Flysjö, A. (2012) *Greenhouse gas emissions in milk and dairy product chains: Improving the carbon footprint of dairy products* (PhD thesis). Aarhus University, Denmark

Fox, J. (2016). *Applied Regression Analysis and Generalized Linear Models* (3rd ed.). Los Angeles: Sage Publications. P 102

Gazan, R., Béchaux, C., Crépet, A., Sirot, V., Drouillet-Pinard, P., Dubuisson, C. and Havard, S. (2016). Dietary patterns in the French adult population: a study from the second French national cross-sectional dietary survey (INCA2) (2006-2007) *British Journal of Nutrition 116*, 200-3015

Harrell, F. E. (2015). Regression Modelling Strategies (2nd ed.). *Cham: Springer*. P. 103-104.

Hedenus, F., Wirsenius, S. & Johansson, D. J. A (2014) The importance of reduced meat and dairy consumption for meeting stringent climate change targets, *Climatic* Change, 124 (1), 79-91

Herman, C. P., Roth, D. A., & Polivy, J. (2003). Effects of the presence of others on food intake: A normative interpretation. *Psychological Bulletin, 129*(6), 873-886

High Level Planet of Experts (HLPE) (2012). *Food security and climate change.* A report by the High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security, Rome 2012.

Hoek, A. C., Luning, P. A., Stafleu, A., de Graaf, C. (2003) Food-related lifestyle and health attitudes of Dutch vegetarians, non-vegetarian consumers of meat substitutes, and meat consumers. *Appetite, 42*(3), 265-272

Intergovernmental Panel on Climate Change (IPCC) (2007a). *"2.3 Climate sensitivity and feedbacks". In Pachauri, R.K; Reisinger, A. Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Geneva, Switzerland: Intergovernmental Panel on Climate Change. 2010-07-03.*

Intergovernmental Panel on Climate Change (IPCC) (2007b) *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel* *on Climate Change* [B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer (eds)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., XXX pp. 409

Intergovernmental Panel on Climate Change (IPCC) (2013a) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA,

Intergovernmental Panel on Climate Change (IPCC) (2014a) Summary for policymakers. In: *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1-32.

Intergovernmental Panel on Climate Change (IPCC) (2014b) Summary for Policymakers. In: *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Intergovernmental Panel on Climate Change (IPCC), (2014c): *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Johansson, G. (2014). *Nutritionsepidemiologi: Kostundersökningar: ändamål, design, biologiska markörer, felkällor och evaluering av kvalitén*. Halmstad: Halmstad University Press. Available at: http://urn.kb.se/resolve?urn=urn:nbn:se:hh:diva-25431

Johnston, J. L., Fanzo, J. C. & Cogill, B. (2014). Understanding sustainable diets: A descriptive analysis of the determinants and processes that influence diets and their impact on health, food security, and environmental sustainability, *Advances in Nutrition,5(*4) 418-429

Lacour, C., Seconda, L., Allés, B., Hercberg, S, Langevin, B., Pointereau, P…. Kesse-Guyot, E. (2018) Environmental Impacts of Plant-Based Diets: How does Organic Food Consumption Contribute to Environmental Sustainability? *Frontiers in Nutrition 5* (8) doi: 10.3389/fnut.2018.00008

Loughnan, S., Bastian, B. & Haslam, N. (2014) The psychology of eating animals. *Sage Journals 23*(2)104-108

Macdiarmid, J. I., Douglas, F. & Campbell (2015). Eating like there's no tomorrow: Public awareness of the environmental impact of food and reluctance to eat less meat as part of a sustainable diet, *Appetite, 96*, 487-493

MacQueen, J. B. (1967). *Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. **1**.* University of California Press. 281–297

Mendenhall, W. & Sincich, T., (2003) *A second course in statistics Regression Analysis*. Sixth Edition. Upper Saddle River, N.J.:Pearson Education Inc.

Naturvårdsverket (2017b) *Fördjupad analys av svensk klimatstatistik 2017*, rapport 6782. Available at: http://www.naturvardsverket.se/Documents/publikationer6400/978-91-620-6782-3.pdf?pid=21185 (accessed 2018-01-30)

Nieminen, P., Lehtiniemi, H., Vähäkangas, K., Huusko, A., Rautio, A., (2013) Standardised regression coefficient as an effect size index in summarising findings in epidemiological studies. *Epidemiology Biostatics and Public Health-2013*. *10,(*4). doi: 10.2427/8854

Nässén, J. (2014) Determinants of greenhouse gas emissions from Swedish private consumption: Time-series and cross-sectional analyses, *Energy, 66*, 98-106

O'Brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors, *Quality & Quantity (2007) 41,* 673-690. doi: 10.1007/s11135-006-9018-6

Olli, E., Grendstad, G., & Wollebaek, D. (2001). Correlates of environmental behaviours: Bringing back social context. *Environment and Behaviour, 33*(2), 181-208

Payne, C. L. R., Scarborough, P., & Cobiac, L. (2016). Do low-carbon-emission diets lead to higher nutritional quality and positive health outcomes? A systematic review of the literature. *Public Health Nutrition*, *19*(14), 2654–2661. Cambridge University Press.

Posen, D., Jaramillo, P., Landis, A. E., and Griffin, M. W. (2017) Greenhouse gas mitigation for U.S plastics production: energy first, feedstocks later, *Environmental Research Letters, 12*(3)

Ricciuto, L, Tarasuk, V & Yatchew, A. (2006). Socio-demographic influences on food purchasing among Canadian households, *European Journal of Clinical Nutrition 60*, 778-790, doi: 10.1038/sj.ejcn.1602382

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S. III., Lambin, E. … Foley, J. (2009) Planetary Boundaries: Exploring the safe operating space for humanity. *Ecology and Society 12*(2),32

Roos, G., Prättälä, R., & Koski, K. (2001). Men, masculinity and food: Interviews with Finnish carpenters and engineers. *Appetite, 37*(1), 47-56

Rydhagen, B. (2013). *Genus och miljö: genusaspekter på miljö och hållbar utveckling*. (1. uppl.) Lund: Student litterature

Räty, R. & Carlsson-Kanyama, A. (2010) Energy consumption by gender in some European countries. *Energy Policy* 38(1), 646-649

Santer, B. D., Céline Bonfils, C., Painter, J. F., Zelinka, M. D., Mears, C., Solomon, S., … Wentz, F. J. (2014). Volcanic contribution to decadal changes in tropospheric temperature, *Nature Geoscience 7,* 185-189. doi: 10.1038/ngeo2098

Scarborough P, Allender S, Clarke D, Wickramasinghe K, Rayner M (2012) Modelling the health impact of environmentally sustainable dietary scenarios in the UK. *European Journal of Clinical Nutrition 66*, 710–715. doi: 10.1038/ejcn.2012.34

Scarborough, P., Appleby, P.N., Mizdrak, A., Briggs, A. D. M., Travis, R. C., Bradbury, K. E. & Key, T. J. (2014). Dietary greenhouse gas emissions of meat-eaters, fish-eaters, vegetarians and vegans in the UK, *Climatic Change 125*(2), 179-192. DOI 10.1007/s10584-014-1169-1

Semenza JC, Hall DE, Wilson DJ, Bontempo BD, Sailor DJ, et al. (2008). Public perception of climate change voluntary mitigation and barriers to behaviour change. *American Journal of Preventive Medicine 35*, 479–487. doi: 10.1016/j.amepre.2008.08.020

Sjörs, C., Hedenus, F., Sjölander, A., Tillander, A. & Bälter, K. (2017). Adherence to dietary recommendations for Swedish adults across categories of greenhouse gas emissions from food, *Public Health Nutrition*. doi:10.1017/S1368980017002300

Sonesson, U., Antesson, F., Davis, J. & Sjödén, P-O. (2005) Home Transports and Wastage – Environmentally Relevant Household Activities in the Life Cycle of Food*, Ambio,* 34(*4-5) 368-372*

Sonesson, U. Davis, J., Ziegler, F. (2009). *Food Production and Emissions of Greenhouse Gases, An overview of the climate impact of different product groups*. SIK- the Swedish Institute for Food and Biotechnology

Soret, S., Mejia, A., Batech, M., Jaceldo-Siegel, K., Harwatt, H. & Sabaté, J. (2014) Climate change mitigation and health effects of varied dietary patterns in real-life settings throughout North America. *The American Journal of Clinical Nutrition, 100*(1), 490-495.
doi: 10.3945/ajcn.113.071589

Springmann, M., Godfray, H. C. J., Rayner, M., Scarborough, P. (2016). Analysis and valuation of the health and climate change co-benefits of dietary change, *Proceedings of the National Academy of Sciences 2016, 113*(15), 4146-4151. doi: 10.1073/pnas.1523119113

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., … Sörlin, S. (2015). Planetary boundaries: Guiding human development on a changing planet, *Science, 347*(6223). DOI: 10.1126/science.1259855

Steinfeld H, Gerber P, Wassenaar T, Castel V, Rosales M, de Haan C. (2006). *Livestock´s Long Shadow. Environmental issues and options*. FAO, Rome

Söderström, M & Dahlberg, S. (2007). *Genusperspektiv på allmänhetens kunskaper och attityder till klimatförändringen (tidigare växthuseffekten).* Stockholm: Naturvårdsverket.

Temme, E. H., Toxopeus, I. B, Kramer, G. F., Brosens, M. C., Drijvers, J. M., Tyszler, M. & Ocké, M. C. (2015) Greenhouse gas emission of diets in the Netherlands and associations with food, energy and macronutrient intakes, *Public health nutrition, 18*(13), 2433-2445.

Treu, H., Nordborg, M., Cederberg, C., Heuer, T., Clapein, E., Hoffmann, H., Berndes, G. (2017). Carbon footprints and land use of conventional and organic diets in Germany. *Journal of Cleaner Production 161,* 127-142. https://doi.org/10.1016/j.jclepro.2017.05.041

Vetter, S. H., Sapkota, T. B., Hillier, J., Stirling, C. M., Macdiarmid, J. I., Aleksandrowicz, L., … Smith, P. (2017). Greenhouse gas emissions from agricultural food production to supply Indian diets: Implications for climate change mitigation. *Agriculture, Ecosystems & Environment*, *237*, 234–241. doi: 10.1016/j.agee.2016.12.024

Willett, W. C., Howe, G. R., Kushi, L. (1997). Adjustment for total energy intake in epidemiologic studies*. The American Journal of Clinical Nutrition 1997; 65*, 1220S-8S

Wirsenius, S., Hedenus, F. & Mohlin, K. (2011) Greenhouse gas taxes on animal food products: rationale, tax scheme and climate mitigation effects, *Climate Change*, *108,* 159-184

Wyness, L., Weishselbaum, E., O'Connor, A., Williams, E.B, Benelam, B., Riley, H. & Stanner, S. (2011) Res meat in the diet: an update. *Nutrition Bulletin 36*, 34-77

Xu, Q., Ding, C., Liu, J., Lou, B. (2014) PCA-guided serch for K-means, *Pattern Recognition Letters 54*(2015) 50-55. Available at: https://doi.org/10.1016/j.patrec.2014.11.017

Web sources

Allen M., P. (1997). *Understanding Regression Analysis,* New York, Plenum Press. Available at: https://link.springer.com/content/pdf/10.1007%2Fb102242.pdf (accessed 2017-11-23)

American Physical Society (APS) (2018) *Policy & Advocacy, Climate Change.* Available at: https://www.aps.org/policy/reports/popa-reports/energy/climate.cfm (accessed 2018-02-03)

Andersson, L. (2011) *Växthusgaser*. Available at: https://tellus.science.gu.se/fokus_arktis/globala_uppvarmningen/uppvarmningens_orsaker/vaxthusgaser (accessed 2018-02-28)

Broman, E. (2017) *Normer och traditioner lever kvar*, Bohuslänningen 7 mar 2017. Available at: http://www.bohuslaningen.se/nyheter/normer-och-traditioner-lever-kvar-1.4186435 (accessed 2018-02-13)

Bryngelsson (2015) Chalmers University of Technology. (2015, April 1). Eggs and chicken instead of beef reap major climate gains. *ScienceDaily*. Available at: www.sciencedaily.com/releases/2015/04/150401084157.htm (accessed 2018-01-30)

Carbonbrief (2013). *Change in global average surface temperature up until 2100 in °C.* Available at: https://www.carbonbrief.org/ipcc-six-graphs-that-explain-how-the-climate-is-changing (accessed 2018-02-02)

Collins English Dictionary (n. d) *Definition of pescatarian.* Available at: https://www.collinsdictionary.com/dictionary/english/pescatarian (accessed 2018-01-08)

Dahlman, L.A. (2017) *Climate change: Global temperature.* Available at: https://www.climate.gov/news-features/understanding-climate/climate-change-global-temperature (accessed 2018-02-03)

Damberg, J. (2017, 11 August) Vegetarisk mat ökar allra mest- bland köttätare. *ETC*. Available at: https://www.etc.se/kultur-noje/vegetarisk-mat-okar-allra-mest-bland-kottatare (accessed 2018-02-07)

Dobronte, A. (2013). *The importance of socio-demographics in online surveys*. Available at: https://www.checkmarket.com/blog/socio-demographics-online-surveys/ (accessed 2018-02-04)

Dudovskiy, J. (2018) *The Ultimate Guide to Writing a Dissertation in Business Studies: A Step-by-Step Assistance, Regression Analysis*. Available at: https://research-methodology.net/research-methods/quantitative-research/regression-analysis/ (accessed 2018-03-06)

Dunn, K. (2016) *Principal Component Analysis (PCA).* Available at: https://learnche.org/pid/latent-variable-modelling/principal-component-analysis/index (accessed 2017-12-27)

Ellis, P.D. (2009), *Effect size equations.* Available at: http://www.polyu.edu.hk/mm/effectsizefaqs/effect_size_equations2.html (accessed 2018-01-11)

European Commission (EC) (2018). *Paris Agreement.* Available at: https://ec.europa.eu/clima/policies/international/negotiations/paris_en (accessed 2018-02-03)

European Commission EC (n.d). *Klimatförändringens konsekvenser.* Available at: https://ec.europa.eu/clima/change/consequences_sv (accessed 2018-02-28)

European Environment Agency (EEA) (2015) "*Agriculture and Climate change*". Available at: https://www.eea.europa.eu/downloads/a898650f58a641589eb0ad2cd92b55be/1481900146/agriculture-and-climate-change.pdf (accessed 2017-11-24)

Frost, J. (2013). *Multiple Regression Analysis: Use Adjusted R-Squared and Predicted R-Squared to Include the correct Number of Variables.* The Minitab Blog. Available at: http://blog.minitab.com/blog/adventures-in-statistics-2/multiple-regession-analysis-use-adjusted-r-squared-and-predicted-r-squared-to-include-the-correct-number-of-variables. (accessed 2017-11-16)

Environmental Protection Agency (EPA) (2017) *Understanding Global Warming Potentials*. Available at: https://www.epa.gov/ghgemissions/understanding-global-warming-potentials (accessed 2018-01-29)

Gesis (n.d.) *Socio demographic characteristics*. Available at: https://www.gesis.org/en/gesis-survey-guidelines/instruments/survey-instruments/socio-demographic-variables/ (accessed 2018-02-04)

Horber, E. (2017) *Regression methods.* Available at: http://www.unige.ch/ses/sococ/cl/spss/cmd/regression.methods.html (accessed 2018-02-05)

IBM (n.d) *Linear Regression Options.* Available at: http://127.0.0.1:55093/help/index.jsp?topic=%2Fcom.ibm.spss.statistics.help%2Fspss%2Fbase%2Fidh_regr_opt.htm (accessed 2018-02-05)

Intergovernmental Panel on Climate Change (IPCC) (2013), Climate change 2013: The Physical Science Basis. Summary for Policy makers. Working group I Contribution to the IPCC Fifth Assessment Report. Available at: http://www.climatechange2013.org/report/ (accessed 2018-03-02)

Investopedia (n.d) *Social Economics*. Available at: https://www.investopedia.com/terms/s/social-economics.asp (accessed 2018-01-31)

ISO (2006*) ISO 14040:2006(en) Environmental management- Life cycle assessment- Principal and framework.* Available at: https://www.iso.org/obp/ui/#iso:std:iso:14040:ed-2:v1:en (accessed 2018-01-11)

Jacobsson, I. (2011) *Dyr mat- en myt,* Råd & Rön. Available at: https://www.radron.se/artiklar/Dyr-maten-myt/ (accessed 2018-03-05)

Jones, J. (n.d.) *Stats: Scheffe' and Tukey Tests*. Available at: https://people.richland.edu/james/lecture/m170/ch13-dif.html (accessed 2018-01-05)

Jonsson, C. (2017) *Matavfallet behöver minska*, Naturvårdsverket. Available at: http://www.naturvardsverket.se/Sa-mar-miljon/Mark/Avfall/Matavfall/ (accessed 2018-02-16)

Jordbruksverket (2012) 19. *Priser på livsmedel*, Jordbruksstatistik årsbok 2011. Available at: https://www.jordbruksverket.se/download/18.4b2051c513030542a92800014485/1370043485805/Kap+19+Priser+på+livsmedel.pdf (accessed 2018-03-03)

Jordbruksverket (2013) *Konsumtion, produktion och prisutveckling för nötkött 1956-2012.* Available at: https://jordbruketisiffror.wordpress.com/2013/07/29/konsumtion-produktion-och-prisutveckling-for-notkott-1956-2012/ (accessed 2018-02-01)

Jordbruksverket (2016) *Totalkonsumtion (förbrukning av kött) åren 1960-2016.* Available at: https://jordbruketisiffror.wordpress.com/2016/02/28/totalkonsumtion-forbrukning-av-kott-aren-1960-2015/ (accessed 2018-01-30)

Lane, D. M., Hebl, M., Guerra, R., Osherson, D. & Scott, D. (n.d) *Introduction to statistics, Difference between two means*. Available at: http://onlinestatbook.com/2/effect_size/two_means.html (accessed 2018-01-14)

Leard Statistics (2013a) *Multiple Regression Analysis using SPSS Statistics*. Available at: https://statistics.laerd.com/spss-tutorials/multiple-regression-using-spss-statistics.php (accessed 2017-10-10)

Leard Statistics (2013b) *One-way ANOVA in SPSS Statistics.* Available at: https://statistics.laerd.com/spss-tutorials/one-way-anova-using-spss-statistics.php (accessed 2018-01-05)

Leard Statistics (2013c) *Independent t-test for two samples*. Available at: https://statistics.laerd.com/statistical-guides/independent-t-test-statistical-guide.php (accessed 2018-01-20)

Libguides (2018) *SPSS Tutorials: One Sample t Test.* Available at: https://libguides.library.kent.edu/SPSS/OneSampletTest (accessed 2018-03-16)

Livsmedelsverket (2017) *Energi, kalorier.* Available at: https://www.livsmedelsverket.se/livsmedel-och-innehall/naringsamne/energi-kalorier (accessed 2017-11-25)

Löfgren, H. (2014) *Grundläggande statistiska metoder för analys av kvantitativa data*. Available at: http://muep.mau.se/bitstream/handle/2043/18962/HLStatbok15.pdf?sequence=2&isAllowed=y (accessed 2017-11-22)

Kasi, A. K. (n.d.) MBA-tutorials, *Factors affecting consumer behaviour.* Available at: http://www.mba-tutorials.com/marketing/621-factors-affecting-consumer-behavior.html (accessed 2018-02-26)

MiniTab (2016) Methods and formulas for chi-square Goodness-of-fit test. Available at: http://support.minitab.com/en-us/minitab-express/1/help-and-how-to/basicstatistics/tables/chi-square-goodness-of-fit-test/methods-and-formulas/methods-and-formulas/ (accessed 2018-03-16)

Naturvårdsverket (2017a) *Konsumtionsbaserade utsläpp av växthusgaser per område.* Available at: http://www.naturvardsverket.se/Sa-mar-miljon/Statistik-A-O/Vaxthusgaser-konsumtionsbaserade-utslapp-per-omrade/ (accessed 2017-12-09)

Naturvårdsverket (2018). *Ekosystemtjänster- när grönt är mer än pynt.* Available at: http://www.naturvardsverket.se/ekosystemtjanster (accessed 2018-02-28)

Nilsson, L.G. (2015), Havets strömmar kontrollerar jordens klimat, *Allt on Vetenskap*, nr 8, 2005, available at: http://www.alltomvetenskap.se/nyheter/havets-strommar-kontrollerar-jordens-klimat (accessed 2018-02-03)

National Aeronautics and Space Administration (NASA) (2018a). *Climate change: How do we know?* Available at: https://climate.nasa.gov/evidence/ (accessed 2018-01-29)

National Aeronautics and Space Administration (NASA) (2018b) *Global Temperature.* Available at: https://climate.nasa.gov/vital-signs/global-temperature/ (accessed 2018-02-01)

National Aeronautics and Space Administration (NASA) (2018c). *A blanket around the Earth*, https://climate.nasa.gov/causes/ (accessed 2018-02-03)

National Aeronautics and Space Administration (NASA) (2018d). *The consequences of climate change*, NASA https://climate.nasa.gov/effects/ (accessed 2018-02-02)

National Aeronautics and Space Administration (NASA) (n.d.) *Global Temperatures.* Available at: https://earthobservatory.nasa.gov/Features/WorldOfChange/decadaltemp.php (accessed 2017-12-20)

National Aeronautics and Space Administration /GISS (NASA/GISS) (2017). Available at: https://climate.nasa.gov/system/internal_resources/details/original/647_Global_Temperature_Data_File.txt (accessed 2018-01-29)

Nationalencyklopedin (n.d.a*)* industriella revolutionen. Available at: http://www.ne.se/uppslagsverk/encyklopedi/lång/industriella-revolutionen (accessed 2018-02-16)

Nationalencyklopedin (n.d.b) demografi. Available at: https://www.ne.se/uppslagsverk/encyklopedi/lång/demografi (accessed 2018-02-26)

National Climate Assessment (NCA), (2014). *Climate change impacts in the United States, Highlights,* National Climate Assessment. Available at: https://nca2014.globalchange.gov/highlights (accessed 2018-02-02)

Newsom (2013) *One-way ANOVA.* Available at: http://web.pdx.edu/~newsomj/da1/ho_ANOVA.pdf (accessed 2018-01-05)

National Oceanic and Atmospheric Administration (NOAA) (n.d) *Oceanic Acidification: The Other Carbon Dioxide Problem.* Available at: https://www.pmel.noaa.gov/co2/story/Ocean+Acidification (accessed 2018-02-28)

Naturvårdsverket and Statistiska centralbyrån (2017). *Konsumtionsbaserade utsläpp av växthusgaser*. Available at:. http://www.naturvardsverket.se/konsumtionsutslapp (accessed 2018-01-30)

Sayad, S. (n.d.) *k-means clustering.* Available at: http://www.saedsayad.com/clustering_kmeans.htm (accessed 2017-12-19)

School of Geography, University of Leeds (n.d.). *Stepwise linear regression.* Available at: http://www.geog.leeds.ac.uk/courses/other/statistics/spss/stepwise/ (accessed 2017-11-21)

Shindell, D. (2013) *Radiative Forcing in the AR5*, CLA, chapter 8. Available at: http://climate.envsci.rutgers.edu/climdyn2013/IPCC/IPCC_WGI12-RadiativeForcing.pdf (accessed 2018-02-03)

Swedens metrological and hydrological institute (SMHI) (2009) *Klimatförändringar orsakad av människan.* Available at: https://www.smhi.se/kunskapsbanken/klimatforandringar-orsakade-av-manniskan-1.3833 2018-02-16 (accessed 2018-02-10)

Statistics How To (2016). *Tukey Test/ Tukey Procedure/ Honest Significant Difference.* Available at: http://www.statisticshowto.com/tukey-test-honest-significant-difference/ (accessed 2018-01-05)

Statistics How To, (2017). *F statistics/ F Value: Simple definition and Interpretation.* Available at: http://www.statisticshowto.com/probability-and-statistics/f-statistic-value-test/ (accessed 2018-01-07)

Socscistatistics (n.d.) *Effect size calculator for T-test.* Available at: http://www.socscistatistics.com/effectsize/Default3.aspx (accessed 2018-01-11)

Statistiska Central Byrån (SCB) (2004) *Utbildning påverkar dina köpvanor*. Available at: https://www.scb.se/sv_/Hitta-statistik/Statistik-efter-amne/Hushallens-ekonomi/Hushallens-utgifter/Hushallens-utgifter-HUT/22941/22948/Behallare-for-Press/109252/ (accessed 2017-12-09)

Statistiska Central Byrån (SCB) (2009) *Vi lägger mer pengar på fritid än mat.* Available at: https://www.scb.se/sv_/Hitta-statistik/Statistik-efter-amne/Levnadsforhallanden/Amnesovergripande-statistik/Valfard---SCBs-tidskrift-om-arbetsliv-demografi-och-valfard/65291/65298/Behallare-for-Press/Hushallens-fordelning-av-olika-utgifter/ (accessed 2017-12-20)

Statistaska Central Byrån (SCB) (n.d) *Konsumentprisindex mäter prisnivån I Sverige.* Available at: https://www.scb.se/hitta-statistik/sverige-i-siffror/samhallets-ekonomi/prisernas-utveckling/ (accessed 2018-02-14)

StatTrek (n.d)Chi-Square Goodness of fit test. Available at: http://stattrek.com/chi-square-test/goodness-of-fit.aspx?Tutorial=AP (accessed 2018-03-16)

Stockholmresillience (2015). *Planetary boundaries.* Available at: http://www.stockholmresilience.org/images/18.3110ee8c1495db74432636f/1459560224012/PB_FIG33_media_11jan2015_web2.jpg (accessed 2018-01-11)

Sundell, A. (2009) *Gude: Regressionsanalys*. SPSS-akuten. Available at: https://spssakuten.com/2009/12/21/regressionsanalys-1/ (accessed 2018-03-06)

Sundell, A. (2010a) *Guide: Regressionsanalys med dummyvariabler,* SPSS_akuten. Available at: https://spssakuten.com/2010/05/11/guide-regressionsanalys-med-dummyvariabler/ (accessed 2017-11-23)

Sundell, A. (2010b). *Guide: Regressionsdiagnostik- multikollinearitet, SPSS-akuten.* Available at: https://spssakuten.com/2010/10/16/guide-regressionsdiagnostik-–-multikollinearitet/. (accessed 2017-11-17)

Svensktkott (n.d.) *Prisutveckling.* Available at: http://www.svensktkott.se/om-kott/statistik/hur-mycket-kott-ater-vi/prisutveckling/ (accessed 2018.02-01)

Svensktvatten (2016) *Vattnets kretslopp.* Available at: http://www.svensktvatten.se/fakta-om-vatten/vattnets-kretslopp/ (accessed 2018-02-28)

The Pennsylvania State University (n.d.) *10.2- Hypothesis Testing.* Available at: https://newonlinecourses.science.psu.edu/stat200/lesson/10/10.2 (accessed 2018-01-05)

TT (2017, 8 April) Ökat sug efter vegetariskt, *SvD*. Available at: https://www.svd.se/okat-sug-efter-vegetariskt (accessed 2018-02-07)

UCDavis (n.d) *Methods Manual: t-test- hand calculation – for independent samples.* Available at: http://psc.dss.ucdavis.edu/sommerb/sommerdemo/stat_inf/tutorials/ttesthand.htm (accessed 2018-01-20)

United Nation (UN) (2016). *Paris Agreement Signing Cermony-22 April 2016.* Available at: http://www.un.org/sustainabledevelopment/parisagreement22april/ (accessed 2018-01-09)

United Nation (UN) (n.d.a). 17 *Goals to Transform Our World*. Available at: http://www.un.org/sustainabledevelopment/ (accessed 2018-02-03)

United Nation (UN) (n.d.b) Goal 13: *Take urgent action to combat climate change and its impact.* Available at: http://www.un.org/sustainabledevelopment/climate-change-2/ (accessed 2018-01-09)

World Meterological Organisation (WMO) (2016) *WMO provisional Statement on the Status of the Global Climate in 2016*. Available at: https://ane4bf-datap1.s3-eu-west-1.amazonaws.com/wmocms/s3fs-public/2016_WMO_Statement_on_the_Status_of_the_Global_Climate-14-11-16-ver2.pdf?ZmIaubFZknHEGDBpyxTBpTcrNotiDpDo (accessed 2018-01-09)

World Meterological Organisation (WMO) Greenhouse bulletin (2017). Available at: https://ane4bf-datap1.s3-eu-west-1.amazonaws.com/wmocms/s3fs-public/ckeditor/files/GHG_Bulletin_13_EN_final_1_1.pdf?LGJNmHpwKkEG2Qw4mEQjdm6bWxgWAJHa (accessed 2017-12-20)

LCA studies

1. Bryngelsson D, Wirsenius S, Hedenus F, Sonesson U. (2016) How can the EU climate targets be met? A combined analysis of technological and demand-side changes in food and agriculture. *Food Policy. 2016*(59),152-64

2. Cejie J. (2008) *Klimatpåverkan från vilt kött (Climate impact from game meat).* Avalilable at: http://ekotank.blogspot.se/2008/11/klimatpverkan-frn-vilt-ktt.html (Webcite: http://www.webcitation.org/6RfsoOEpH) (accessed December 2016)

3. Flysjö A. (2012) *Greenhouse gas emissions in milk and dairy product chains: Improving the carbon footprint of dairy products*. PhD thesis. Aarhus University, Denmark. 2012

4. Doddridge E. (2015) *The BBQ challeng.* Available at*:* http://doddridge.me/2015/01/02/the-bbq-challenge/ (Webcite: http://www.webcitation.org/6fAYfgpmr) (accessed December 2016)

5. Doddridge E. (n.d.) *Detailed working for the different BBQ scenarios.* Available at: http://nbviewer.jupyter.org/urls/dl.dropbox.com/s/2qp10d2blkkn5m9/BBQs.ipynb?dl=0 (Webcite: http://www.webcitation.org/6fAYjgkl1) (accessed December 2016)

6. Nilsson K, Sund V, Florén B. (2011) *The environmental impact of the consumption of sweets, crisps and soft drinks*. Copenhagen: Nordic Council of Ministers; 2011. TemaNord 2011:509

7. Röös E. (2014) *Mat-klimat-listan Version 1.*1 (Food-climate-list Version 1.1). In Swedish. Swedish University of Agricultural Sciences (SLU). Report 077. ISSN 1654-9406. Available at: http://pub.epsilon.slu.se/11671/7/roos_e_141125.pdf (Webcite: http://www.webcitation.org/6mtTRm6Dn (accessed December 2016)

8. Lidell A-K. (2016) *Food for progress*. Personal communication

9. Davis J, Wallman M, Sund V, Emanuelsson A, Cederberg C, Sonesson U. *Emissions of greenhouse gases from production of horticultural products. Analysis of 17 products cultivated in Sweden.* SIK - The Swedish Institute for Food and Biotechnology, Gothenburg, Sweden. 2011. SR 828. ISBN 978-91-7290-301-2

10. Florén B, von Bah B, Davis J, Flysjö A, Högberg J, Lorentzon K, et al. *Global warming potential for 100 ICA private label food products.* Final report.; 2007. Report No.: UP-07-14423

11. Nilsson K. *Klimatpåverkan från bryggkaffe och snabbkaffe (Climate impact of filter coffee and instant coffee).* In Swedish. Report UPX00221. SIK - The Swedish Institute for Food and Biotechnology, Gothenburg, Sweden. 2010

12. Scarborough P, Appleby P, Mizdrak A, Briggs A, Travis R, Bradbury K, et al. Dietary greenhouse gas emissions of meat- eaters, fish- eaters, vegetarians and vegans in the UK. *Climatic Change. 2014;125*(2),179-92

13. Saxe. *LCA-based comparison of the climate footprint of beer wine spirits*. 2010

14. Amienyo. *Life cycle sustainability assessment in the UK beverage sector*. 2012

15. Ziegler F, Winther U, Hognes E, Emanuelsson A, V VS, H HE. The carbon footprint of Norwegian seafood products on the global seafood market. *Journal of Industrial Ecology. 2013,17(*1),103-16

16. Winther U, F FZ, Hognes E, Emanuelsson A, Sund V, Ellingsen H. *Carbon footprint and energy use of Norwegian seafood products.* SINTEF Fisheries and Aquaculture, Report SFH80 A. 2009;96068

17. Parker RWR, Tyedmers PH. Fuel consumption of global fishing fleets: current understanding and knowledge gaps. *Fish and Fisheries. 2015,16*(4),684-96

18. Konsumentföreningen Stockholm (Stockholm Consumers Association)/SIK – Institutet för Livsmedel och Bioteknik. *Vatten på flaska onödig lyx. Totala koldioxidutsläpp från konsumtionen av buteljerat vatten i Sverige.* In Swedish. (Bottled water unnecessary luxury. Total carbon dioxide emissions from the consumption of bottled water in Sweden.) 2007 [Available at: http://www.konsumentforeningenstockholm.se/Global/Konsument%20och%20Milj%c3%b6/Rapporter/Vatten%20pa%20flaska_onodig%20lyx%20(2)_2007.pdf (accessed December 2016)

Figures

1. Planetary boundaries Stockholmresillience (2015)
2. Change in global surface temperature in relation to 1951-1980 average (NASA, 2018b)
3. Change in GHGE concentration in atmosphere 1850-2012 IPCC (2014c.)
4. Annual GHGE for four scenarios; RCP 2.6, RCP 4.5, RCP6.0 and RCP8.5 and total atmospheric concentrations for respective scenario up until 2100 (IPCC, 2014b).
5. Change in global average surface temperature up until 2100 in °C (Carbonbrief, 2013)
6. Distribution of consumption areas in the household consumption-based emissions (Naturvårdsverket and Statistiska centralbyrån, 2017)
7. Comparison of GHGE in kg CO_2 eq/ kg food product from eight food products based on GHGE data used in Sjörs et al. (2017).
8. Three proposed diets and their corresponding GHGE (Bryngelsson, D., 2015)
9. Price development in sek/kg of ruminant meat in Sweden between 1958-2011 (Jordbruksverket, 2013)
10. Willett residual method for energy adjustment (Willett, W. C., 1997)

13-16. Principal component analysis geometrics (Dunn, 2016)

Appendix

A1 Original LCA-data

Table A1: LCA data of greenhouse gas emissions expressed in kg CO_2 eq. per kg food product or group used for calculations of individual daily greenhouse gas emissions. in Sjörs et al. (2017). For references see *LCA studies* in reference list.

Food products or groups	Kg CO_2e / kg food product	Estimated from
Beef, raw	35,9	Bryngelsson et al (1)
Beef, prepared (fried/boiled/smoked)	45,9	Bryngelsson et al (1)
Minced meat (mix of beef and pork 70/30), raw	27,0	Bryngelsson et al (1)
Minced meat (mix of beef and pork 70/30), prepared (fried)	35,9	Bryngelsson et al (1)
Minced meat (mix of beef and pork 50/50), raw	21,1	Bryngelsson et al (1)
Minced meat (mix of beef and pork 50/50), prepared (fried)	28,0	Bryngelsson et al (1)
Sheep meat (mutton), raw	39,0	Bryngelsson et al (1)
Sheep meat (mutton), prepared (fried)	51,7	Bryngelsson et al (1)
Game, raw	5,2	Cejie (2)
Game, prepared (fried/boiled)	6,7	Cejie (2)
Pork, raw	6,4	Bryngelsson et al (1)
Pork, prepared (fried/boiled, ham)	9,6	Bryngelsson et al (1)
Poultry, raw	2,5	Bryngelsson et al (1)
Poultry, prepared (fried/boiled)	3,6	Bryngelsson et al (1)
Whole milk	1,5	Bryngelsson et al (1) and Flysjö (3)
Semi-skimmed milk	1,3	Bryngelsson et al (1) and Flysjö (3)
Skimmed milk	1,2	Bryngelsson et al (1) and Flysjö (3)
Yoghurt	1,6	Bryngelsson et al (1) and Flysjö (3)
Yoghurt low fat	1,4	Bryngelsson et al (1) and Flysjö (3)
Crème fraiche	6,3	Bryngelsson et al (1) and Flysjö (3)
Crème fraiche low fat	3,8	Bryngelsson et al (1) and Flysjö (3)
Yellow cheese	11,8	Bryngelsson et al (1) and Flysjö (3)
Yellow cheese low fat	10,8	Bryngelsson et al (1) and Flysjö (3)
White cheese	9,7	Bryngelsson et al (1) and Flysjö (3)
Mould cheese	10,0	Bryngelsson et al (1) and Flysjö (3)
Cream cheese	8,1	Bryngelsson et al (1) and Flysjö (3)
Cream cheese low fat	5,1	Bryngelsson et al (1) and Flysjö (3)
Halloumi and feta cheese	16,6	Bryngelsson et al (1) and (4) and (5)
Mozzarella	12,2	Bryngelsson et al (1) and (4) and (5)
Cottage cheese	4,2	Bryngelsson et al (1) and Flysjö (3)
Cream	6,5	Bryngelsson et al (1) and Flysjö (3)
Cream low fat	3,1	Bryngelsson et al (1) and Flysjö (3)
Ice-cream	2,6	Nilsson et al (6)

Butter	13,0	Bryngelsson et al [1] and Flysjö [3]
Butter blends (Bregott)	8,3	Bryngelsson et al [1] and Flysjö [3]
Vegetable oils and margarine	2,3	Bryngelsson et al [1]
Soy drink, oat drink (coconut milk)	0,3	Bryngelsson et al [1]
Eggs	1,3	Bryngelsson et al [1]
Legumes (beans, peas and lentils), dried	0,8	Bryngelsson et al [1]
Legumes (beans, peas and lentils), soaked and boiled	0,3	Bryngelsson et al [1]
Meat substitutes (other than quorn)	2,5	Röös [7] and Lidell [8]
Quorn	4,5	Röös [7]
Fresh fruits and berries, domestic	0,3	Bryngelsson et al [1]
Fresh fruits and berries, imported (other than banana and citrus fruit)	1,3	Bryngelsson et al [1]
Fresh fruit, berries and vegetables, aviation	18,5	Röös [7]
Dried fruit	11,7	Bryngelsson et al [1]
Banana	2,6	Bryngelsson et al [1]
Citrus fruit	1,0	Bryngelsson et al [1]
Tomato	1,0	Bryngelsson et al [1]
Iceberg lettuce	0,4	Bryngelsson et al [1]
Cucumber	1,4	Davis [9]
Root vegetables, onion	0,2	Bryngelsson et al [1]
Broccoli and vegetables not included in "root vegetables, onion"	0,9	Bryngelsson et al [1]
Potatoes, raw	0,2	Bryngelsson et al [1]
Potatoes, prepared (boiled/baked/fried/pommes frites)	0,3	Bryngelsson et al [1]
Pasta, couscous, bulgur, quinoa, dried	0,7	Bryngelsson et al [1]
Pasta, couscous, bulgur, quinoa, prepared (boiled)	0,3	Bryngelsson et al [1]
Flour, grain	0,6	Bryngelsson et al [1]
Breakfast cereals	0,7	Bryngelsson et al [1]
Rice, raw	2,1	Bryngelsson et al [1]
Rice, prepared (boiled)	0,7	Bryngelsson et al [1]
Olives	3,4	Florén et al [10]
Bread and crisp bread	0,7	Bryngelsson et al [1]
Cookies and biscuits	1,2	Bryngelsson et al [1] and Nilsson et al [6]
Soft drinks, fruit syrup	0,2	Nilsson et al [6]
Juice	0,9	Röös [7]
Instant coffee powder	8,3	Nilsson et al [11]
Coffee (prepared)	0,2	Nilsson et al [11]
Tea	0,04	Nilsson et al [11] and Scarborough et al [12]
Snacks (crisps etc)	2,3	Nilsson et al [6]
Foam sweets	4,1	Nilsson et al [6]
Jelly sweets	2,6	Nilsson et al [6]
Milk chocolate	2,9	Nilsson et al [6]
Dark chocolate	1,0	Nilsson et al [6]

Sugar and syrup	4,8	Bryngelsson et al [1]
Jam	3,8	Röös [7]
Nuts and Seeds	1,3	Bryngelsson et al [1]
Wine / liqueur	2,3	Bryngelsson et al [1]
Beer	1,1	Bryngelsson et al [1]
Cider, strong	1,1	Bryngelsson et al [1]
Distilled beverage/ vodka/ rum	3,6	Saxe et al [13] and Amienyo [14]
Salmon, raw. Norwegian aquaculture.	3,8	Ziegler et al [15] and Winther et al [16]
Salmon, prepared (fried/boiled/smoked). Norwegian aquaculture.	4,9	Ziegler et al [15] and Winther et al [16]
Cod, raw. Caught in Norwegian fisheries by various fishing gears.	3,7	Ziegler et al [15] and Winther et al [16]
Cod, prepared (fried/boiled). Caught in Norwegian fisheries by various fishing gears.	4,2	Ziegler et al [15] and Winther et al [16]
Saithe, raw. Caught in Norwegian fisheries by various fishing gears.	3,0	Ziegler et al [15] and Winther et al [16]
Saithe, prepared (fried/boiled). Caught in Norwegian fisheries by various fishing gears.	3,3	Ziegler et al [15] and Winther et al [16]
Haddock, raw. Caught in Norwegian fisheries by various fishing gears.	4,3	Ziegler et al [15] and Winther et al [16]
Haddock, prepared (fried/boiled). Caught in Norwegian fisheries by various fishing gears.	4,8	Ziegler et al [15] and Winther et al [16]
Herring, raw. Caught by the Norwegian pelagic fleet.	1,1	Ziegler et al [15] and Winther et al [16]
Herring, prepared (fried/boiled/pickled). Caught by the Norwegian pelagic fleet.	1,3	Ziegler et al [15] and Winther et al [16]
Mackerel, raw. Caught by the Norwegian pelagic fleet.	1,2	Ziegler et al [15] and Winther et al [16]
Mackerel, prepared (fried). Caught by the Norwegian pelagic fleet.	1,3	Ziegler et al [15] and Winther et al [16]
Crustaceans (shrimps, lobster) without shell, boiled. Europe, bottom trawl.	38,7	Parker et al [17]
Crustaceans (shrimps, lobster), with shell, boiled. Europe, bottom trawl.	14,7	Parker et al [17]
Mussel, without shell. Norwegian aquaculture.	1,8	Ziegler et al [15] and Winther et al [16]
Tuna raw	3,5	Ziegler et al [15], Winther et al [16] and Parker et al [17]
Tuna, prepared	4,0	Ziegler et al [15], Winther et al [16] and Parker et al [17]
Other seafood, passive capture fisheries, raw	0,7	Ziegler et al [15] and Winther et al [16]
Other seafood, passive capture fisheries, prepared	0,8	Ziegler et al [15] and Winther et al [16]
Other seafood, active capture fisheries, raw	2,8	Ziegler et al [15] and Winther et al [16]
Other seafood, active capture fisheries, prepared	3,1	Ziegler et al [15] and Winther et al [16]
Bottled water	0,1	Konsumentföreningen Stockholm/SIK [18]
Unknown (including tap water)	0	

A2 Food categorisation

Table A2: 23 Food categories based on food groups

	Food categories	Food groups
1	Ruminant meat	Ruminant meat, calf, fresh, frozen, cooked
2	Pork	Pork, fresh, frozen, cooked
3	Poultry	Poultry products, dishes Turkey Chicken Other poultry
4	Fish/seafood	Fish, fresh, frozen, boiled Fish and seafood products Fish smoked Fish fried Rom Seafood, squid, fresh, frozen, boiled Seafood, squid, canned
5	Mixed meat	Meat products dishes Horse meat Blood products Giblets Liver pâté "Sylta" Lamb Moose Deer Roe deer
6	Sausage	Sausage, ruminant, pork, lamb, horse Sausage dishes
7	Pizza/hamburger	Pizza, pie, pasty, ready sandwich Hamburger with bread (all meat)
8	Egg	Egg, boiled, raw Egg fried Egg products, dishes
9	Dairy (excluding cheese and yoghurt)	Milk Flavoured milk Milk drinks, chocolate, milkshake, smoothie Milk powder Butter
10	Cheese	Hard cheese Dessert cheese Fresh cheese Cheese dishes Melted cheese
11	Yoghurt/sour milk	Natural yoghurt, sour milk Flavoured yoghurt, sour milk
12	Vegetables	Vegetables, frozen, boiled canned Vegetables, preserved Vegetables, raw Vegetables, fried, wooked (raw, frozen) Vegetable, root veg., legumes prod.

		Vegetable mix Vegetable juice Salad mix Mushroom, raw Mushroom, frozen, boiled, canned
13	Potatoes/root vegetables	Potatoes, raw, boiled, baked Potato products and dishes Root vegetables, frozen, boiled, canned Root vegetables preserved Root vegetables, raw
14	Fruit and berries	Berries, fresh, frozen Fruit, fresh, frozen Fruit, berries, canned Fruit, berries, dried Nectar Juice
15	Legumes	Legumes, fresh (frozen, raw, cooked, canned) Legumes, sprouted Legumes, dried (boiled, canned)
16	Pasta/rice/grain	Pasta dishes Rice, rice noodles Rice dishes Bulgur, couscous Other food grain
17	Grain Products	Hard bread Soft bread Crust Oat, wheat, rye flour Porridge Formula Cereals
18	Nuts/seeds	Nuts, seeds
19	Vegetarian milk prod.	Veg. milk products Veg cheese
20	Vegetarian substitute	Quorn Soy, seitan products
21	Desserts/sweets	Sweet bread, doughnuts Sweet crackers, cookies Desserts Soft cookies Danish Pancakes Other, ice-cream
22	Snacks	Chips/ popcorn
23	Fat	Margarine, fat <70% fat Margarine fat >70% fat Oil Other fat

A3 Result from Principal component analysis

Table A3: Eigen-value table from Principal component analysis showing 23 components of which eleven has eigen values above 1,000.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1,739	7,560	7,560	1,739	7,560	7,560	1,380	5,999	5,999
2	1,440	6,263	13,823	1,440	6,263	13,823	1,337	5,813	11,812
3	1,265	5,499	19,321	1,265	5,499	19,321	1,247	5,420	17,233
4	1,208	5,250	24,572	1,208	5,250	24,572	1,246	5,419	22,652
5	1,191	5,180	29,751	1,191	5,180	29,751	1,231	5,353	28,005
6	1,107	4,813	34,565	1,107	4,813	34,565	1,172	5,096	33,101
7	1,067	4,638	39,203	1,067	4,638	39,203	1,140	4,957	38,058
8	1,049	4,561	43,763	1,049	4,561	43,763	1,130	4,914	42,972
9	1,011	4,397	48,160	1,011	4,397	48,160	1,086	4,722	47,694
10	1,006	4,376	52,536	1,006	4,376	52,536	1,083	4,709	52,403
11	1,001	4,350	56,886	1,001	4,350	56,886	1,031	4,483	56,886
12	,957	4,160	61,046						
13	,948	4,122	65,168						
14	,908	3,948	69,116						
15	,901	3,918	73,035						
16	,865	3,759	76,794						
17	,822	3,574	80,368						
18	,813	3,533	83,901						
19	,796	3,460	87,360						
20	,778	3,382	90,742						
21	,753	3,273	94,016						
22	,726	3,156	97,172						
23	,651	2,828	100,000						

Table A4: Rotated component matrix showing the eleven components with eigen values above 1,000 and respective food categories loading on each component.

Rotated Component Matrix[a]

	Component										
	1	2	3	4	5	6	7	8	9	10	11
Pork	,722										
Egg	,635										
Sausage	,415		,357								
Veg. milk prod.		,683									
Nuts and seeds		,614									
Vegetables, mushrooms		,404							,316		
Desserts			,715								
Fat				,701							
Cheese				,659					,335		
Grain prod.			,357	,435					-,302		
Yoghurt, sour milk					,747						
Fruit, berries					,587						
Pasta, rice, grain						,602					
Mixed meat						,547					
Dairy prod.						,431				,343	
Poultry			-,373			-,416					
Snacks							,756				
Pizza, hamburger							,538				
Fish, seafood								,820			
Potato, root veg.			,384					,481			
Legumes									,678		
Ruminant meat										,835	
Veg. substitute											,906

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.[a]
a. Rotation converged in 12 iterations.

A4 Result from k-mean clustering

Table A5: Final cluster centers from k-mean clustering for eleven clusters

Final Cluster Centers

	Cluster 1	2	3	4	5	6	7
1	-,92119	1,29578	-,10078	-,16036	-,13451	,16523	-,34132
2	9,25753	-1,77327	-,24306	-,11531	-,21248	2,82123	,40638
3	,82336	2,04123	,07657	-,09427	-,16364	,04277	-,88834
4	-,78838	4,82556	-,14111	-,04616	1,64469	1,81298	,31735
5	-2,05720	-3,51437	,12639	-,30114	-,12440	,98778	-,01214
6	-,33171	5,59740	,04275	-,26728	-,47176	-1,88749	2,84458
7	-1,22255	1,47123	,04045	2,46162	-,34808	-,52589	-,27816
8	-,30106	,51995	-,19806	-,03417	,27074	-,07343	-,30147
9	-1,15047	3,77196	-,31228	-,24210	-,40462	-1,93455	-1,00702
10	-,16077	4,41338	3,09769	-,14419	-,12498	-,18205	,05259
11	1,35412	2,64368	-,03760	-,06971	-,13716	28,84905	-,41161

Final Cluster Centers

	Cluster 8	9	10	11
1	-,34525	1,92246	-,26159	,10234
2	-,24899	-,07000	-,33569	1,08534
3	,12205	,31441	-,89656	-,33094
4	-,38832	-,21449	1,12232	-,03122
5	-,06391	,17303	2,36718	,45124
6	-,00770	,03711	-,84646	-,32738
7	-,17601	-,12294	,64342	-,13941
8	-,08310	,31608	-,92151	-,02294
9	,01628	-,22966	-,71001	1,28717
10	-,21913	-,08776	-,52460	,12242
11	,03333	,03124	10,58714	-,24962

Table A6: Summary of k-mean clustering with characteristics for each cluster

Cluster	Components	
1	2	Veg. milk, nuts, seeds, vegetables, mushrooms,
2	Mix	Mix
3	10	Ruminant meat, dairy prod,
4	7	Snacks, pizza, hamburger
5	4	Fat, cheese, grain prod.
6	11	Veg substitute
7	6	Pasta, mixed meat, poultry, dairy
8	Mix	Mix
9	1	Pork, egg, sausage
10	11	Veg, substitute
11	2, 9, (5)	Veg. milk, nuts/seeds, vegetables, mushrooms, cheese, grain prod., (youghurt, sour milk, fruit, berries)

Table A7: Formation of diet groups from clustering, summarizing merged cluster due to similarities in characteristic food

Diet groups	Cluster	Characteristics	Diet group name
1	1+6+10	Veg. milk, nuts, seeds, vegetables, mushrooms,	Vegan
2	11	Veg. milk, nuts/seeds, vegetables, mushrooms, cheese, grain prod., (youghurt, sour milk, fruit, berries)	Pescotarian
3	5	Fat, cheese, grain prod.	Cheese sandwich
4	7	Pasta, mixed meat, poultry, dairy	Spaghetti Carbonara
5	4	Snacks, pizza, hamburger	Junk food
6	9	Pork, egg, sausage	Full English
7	3	Ruminant meat, dairy prod,	Steak Eaters
8	2+8	mix	Swedish Svensson

A5 Consumption of food categories in diet groups

Consumption of food categories in diet groups (kcal)

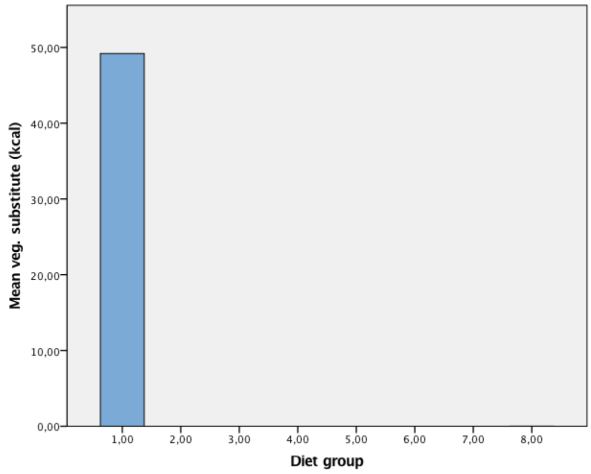


Figure A1: Mean of caloric intake of veg. substitute

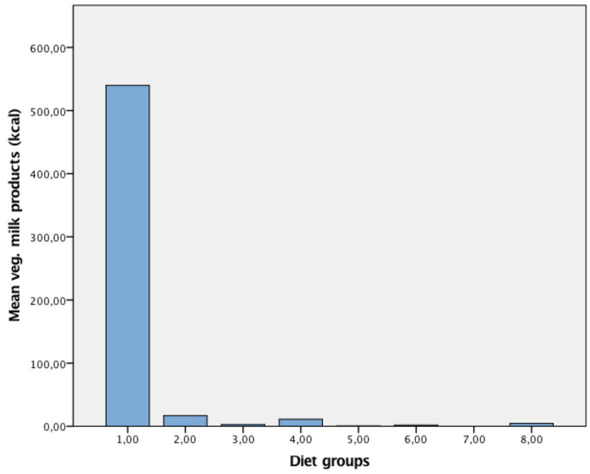


Figure A2: Mean of caloric intace of veg. Milk products

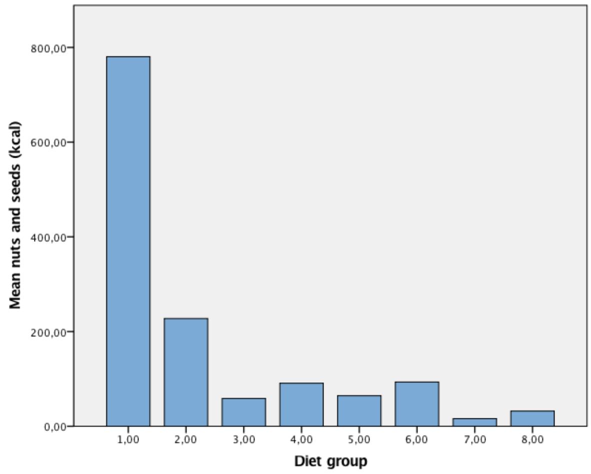


Figure A3: Mean of caloric intake of nuts and seeds

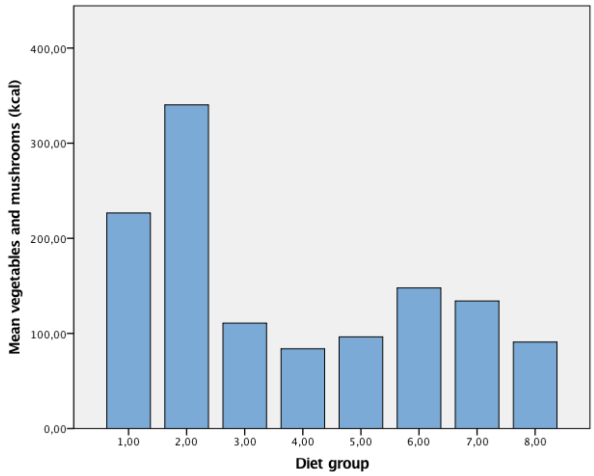


Figure A4: Mean caloric intake of vegetables and mushrooms

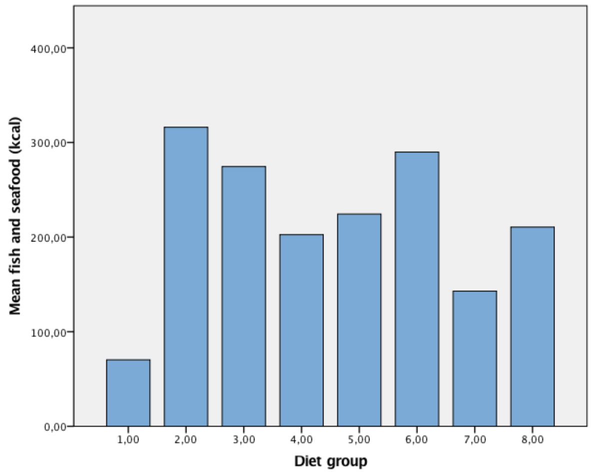


Figure A5: Mean caloric intake of fish and seafood

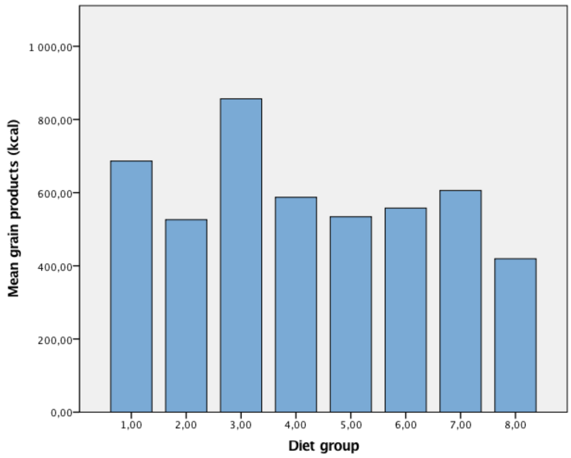


Figure A6: Mean caloric intake of grain products

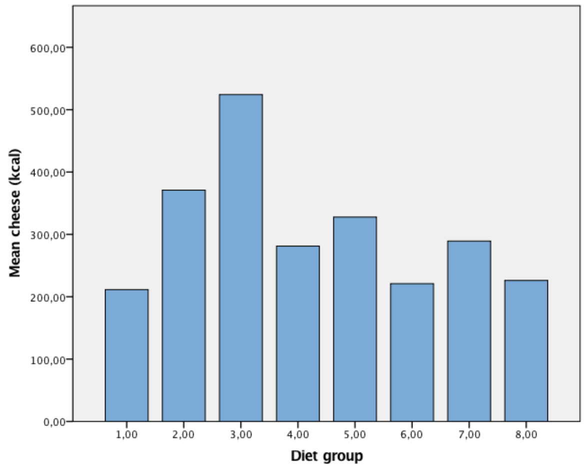


Figure A7: Mean caloric intake of cheese

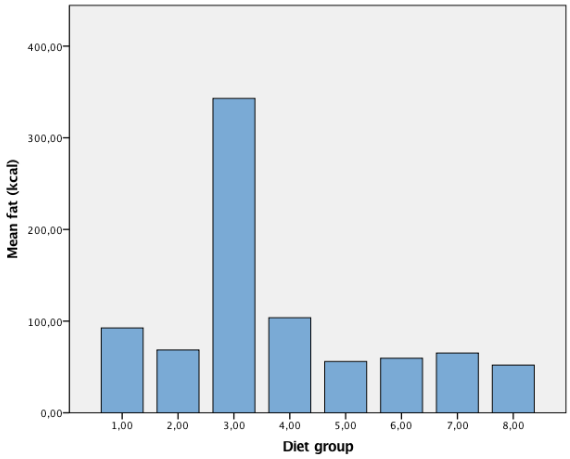


Figure A8: Mean caloric intake of fat

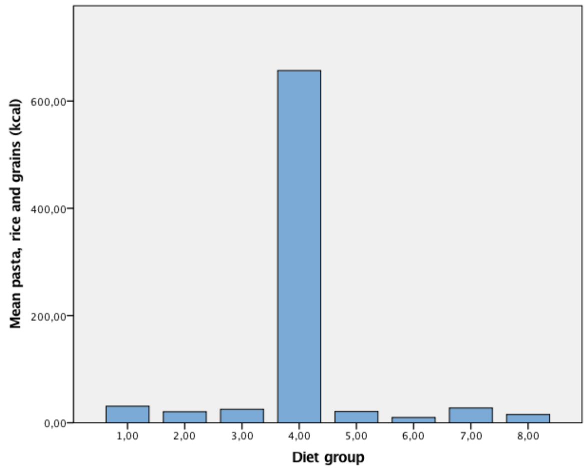


Figure A9: Mean caloric intake of pasta, rice and grains

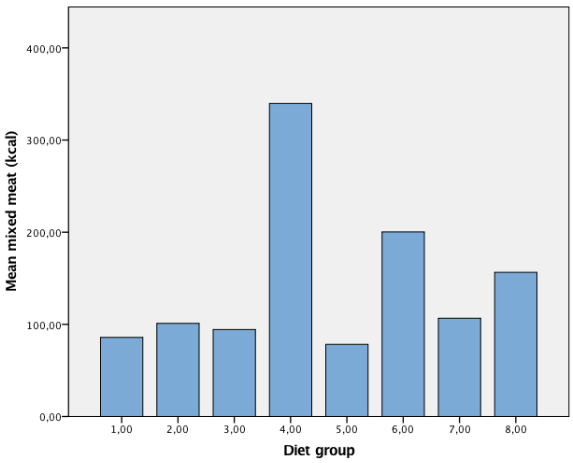


Figure A10: Mean caloric intake of mixed meat

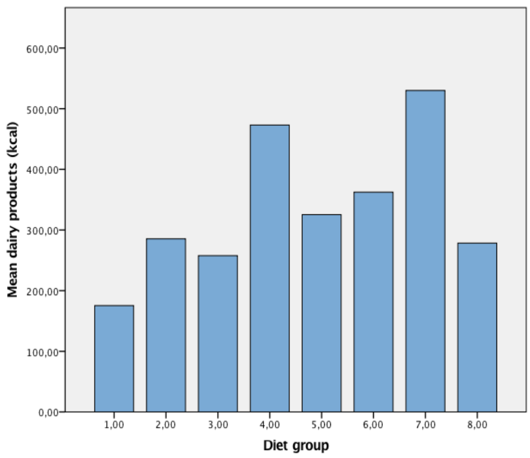


Figure A11: Mean intake of dairy products

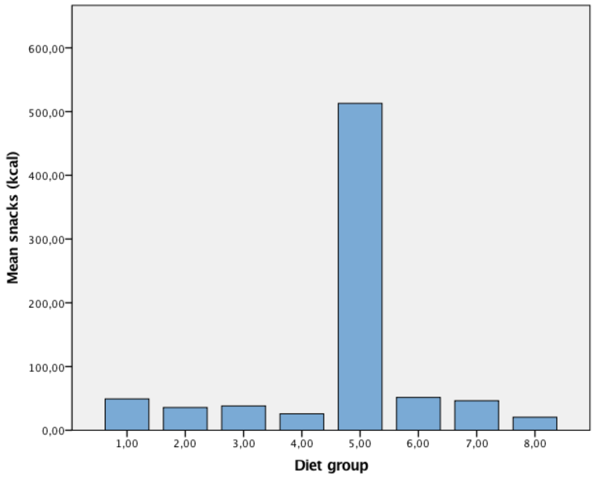


Figure A12: Mean caloric intake of snacks

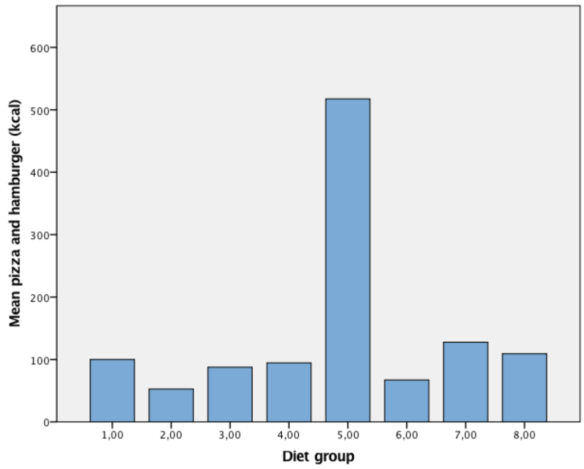


Figure A13: Mean caloric intake of pizza and hamburger

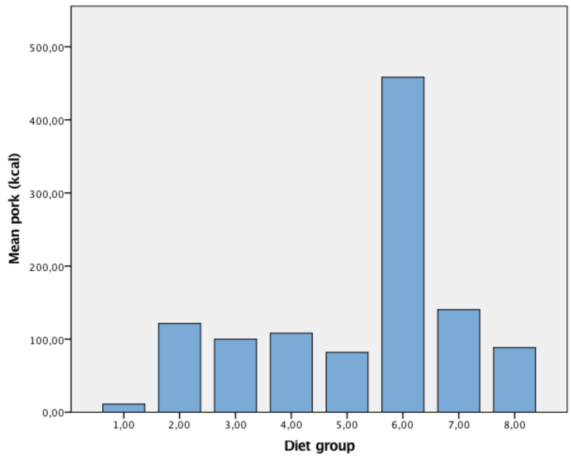


Figure A14: Mean caloric intake of pork

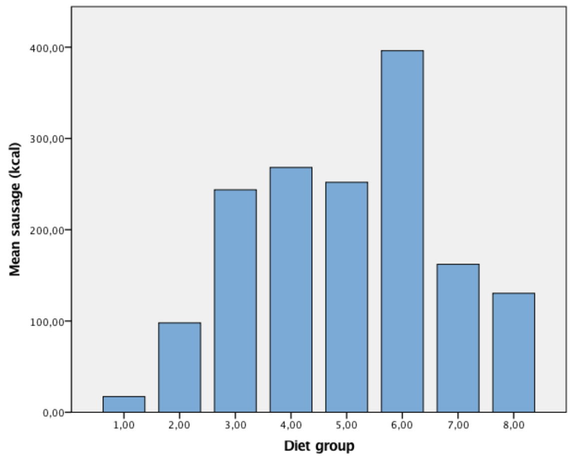


Figure A15: Mean caloric intake of sausage

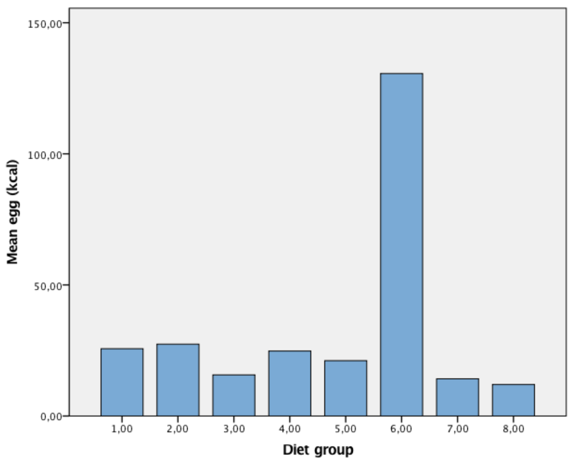


Figure A16: Mean caloric intake of egg

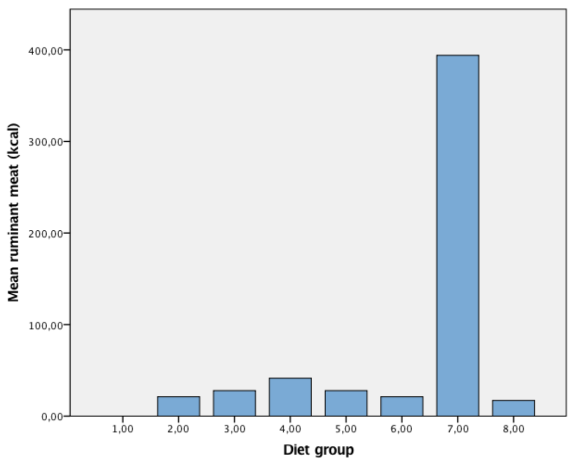


Figure A17: Mean caloric intake of ruminant meat

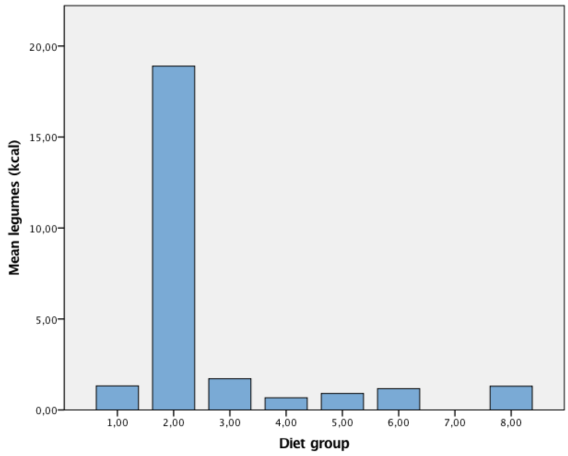


Figure A18: Mean caloric intake of legumes

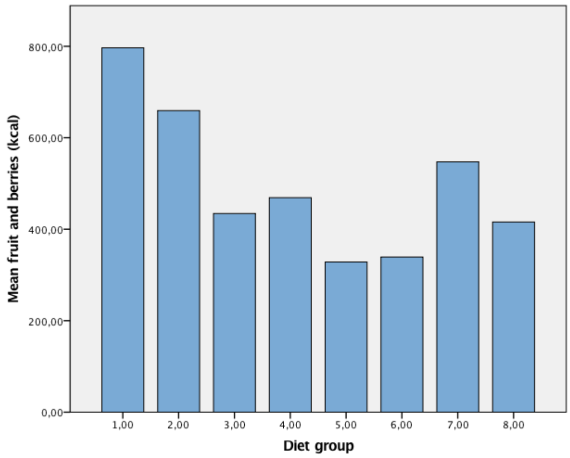


Figure A19: Mean caloric intake of fruit and berries

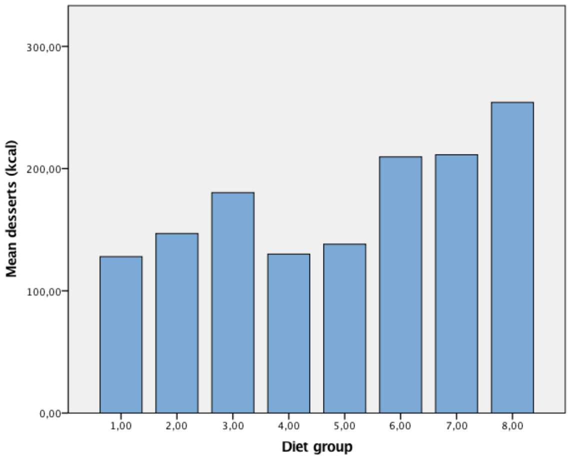


Figure A20: Mean caloric intake of desserts

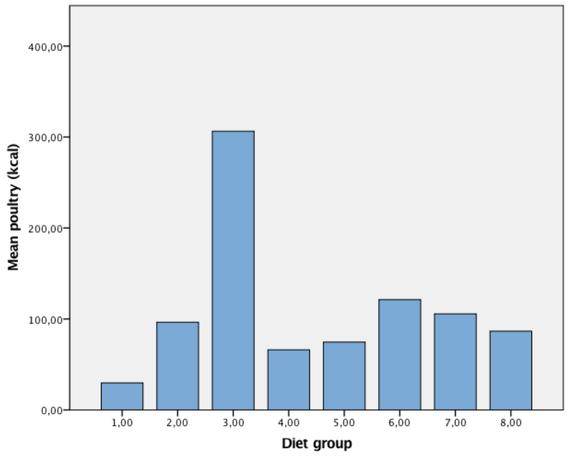


Figure A21: Mean caloric intake of poultry

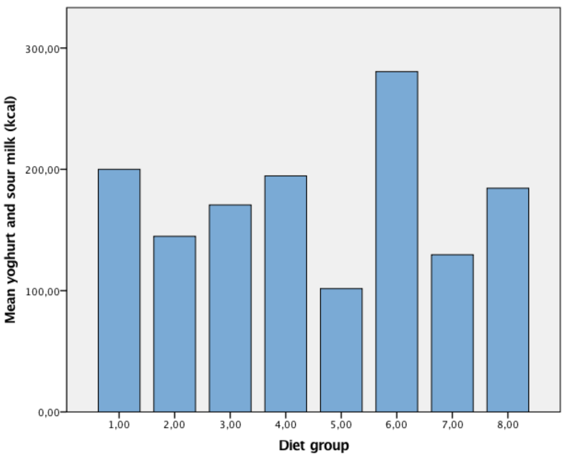


Figure A22: Mean caloric intake of ypghurt and sour milk

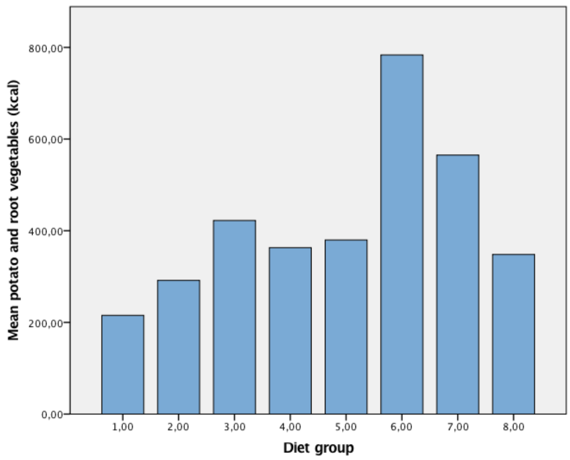


Figure A23 Mean caloric intake of potato and root vegetables

A6 Result from evaluation of socio-demographic variables in diet groups

Table A8: Means of household income and age in diet groups

Diet group	Household income sek	Age years
Vegan	277 852	33
Pescotarian	402 676	51
Cheese sandwich	403 051	49
Spaghetti Carbonara	412 707	44
Junk Food	360 579	38
Full English	435 173	49
Steak eaters	463 635	48
Swedish Svensson	433 191	50

Table A9: Distribution of gender, education, region, single household, children and student in diet groups

Diet group	Gender %		Education %			Region %		Single household %	Children %	Student %
	M	F	Low	Med	High	U	R			
Vegan	25	75	0	58	42	92	8	17	50	8
Pescotarian	27	73	5	35	59	75	25	27	38	7
Cheese sandwich	56	44	10	40	51	75	25	14	44	7
Spaghetti Carbonara	43	57	12	41	47	76	24	17	53	9
Junk Food	60	40	11	42	47	82	18	19	48	9
Full English	73	27	16	43	41	62	38	14	48	6
Steak eaters	64	36	10	42	47	73	27	13	51	8
Swedish Svensson	30	70	13	42	45	73	27	19	38	9