

Consumer preferences for animal welfare:

What can we learn from restaurant, store and web experiments?

Forbrukerpreferanser for dyrevelferd:

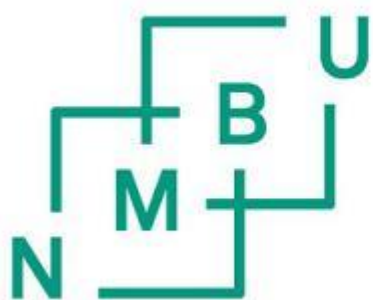
Hva kan vi lære fra eksperimenter i restauranter, dagligvarebutikker og på web?

Philosophiae Doctor (PhD) Thesis

Alexander Schjøll

School of Economics and Business
Norwegian University of Life Sciences

Ås 2014



**Norwegian University
of Life Sciences**

Thesis number 2014: 60
ISSN 1894-6402
ISBN 978-82-575-1223-1

Contents

Contents	i
Preface	iii
Acknowledgements	v
Abstract	vii
Norsk sammendrag	xi
List of papers	xv
Part I: Introduction	17
1. Introduction	1
1.1 Credence attributes related to the production of food products.....	1
1.2 Experiments conducted in restaurants, grocery stores and on the Web.....	3
1.3 Overview of the thesis	5
1.4 Selection of methods	6
2. Research questions	9
3. The cases in the thesis	9
3.1 Overview of cases	9
3.2 Agricultural production in Norway	13
4. Methods used.....	14
5. Relevant consumer studies on animal welfare	17
5.1 Beef studies	17
5.2 Egg studies	22
5.3 Summary of the literature.....	28
6. Summary of papers.....	29
6.1 Paper 1: Eliciting consumer preferences for credence attributes in an haute cuisine restaurant	29
6.2 Paper 2: Country-of-origin preferences for organic food.....	32
6.3 Paper 3: Animal welfare in stores: natural field experiments with positive and negative labeling	34
6.4 Paper 4: Attitude versus action for farm animal welfare: what can we learn from natural field experiments?.....	36
7. Limitations of the thesis	38
8. Generalization	39

9. Contribution.....	40
References	42
Part II: Papers	47
Paper 1	51
Paper 2.....	74
Paper 3.....	100
Paper 4.....	125
Figure 1: The cases presented in the thesis.	10
Figure 2: Percentage of people in various European countries answering ‘very important’ in response to questions regarding the importance of various aspects of animal welfare.	18
Figure 3: Additional price consumers in selected European countries are willing to pay for eggs produced in high-animal-welfare production systems.	23
Figure 4: Norwegian consumers’ rating of attributes when buying eggs.	27

Preface

This thesis is part of the project “Developing robust and economically viable models for cow-calf suckling in organic dairy systems, fulfilling high standards for animal health, welfare and ethics”. The overall purpose of the project is to identify the “ideal” organic rearing system for calves 0-12 weeks of age. By ideal is meant a system that is animal friendly, but also economically sustainable and produce a product that consumers might find attractive.

The Foundation for Research Levy on Agricultural Products (FFL) and the Agricultural Agreement Research Fund (JA) through Norwegian Research Council in Norway (project no. 190424), has founded the project.

I have conducted the work during my stay as a Research Fellow at National Institute for Consumer Research (SIFO), where Research Professor Svein Ole Borgen has been project leader for the consumer part of the project and co-supervisor. Main supervisor has been Professor Frode Alfnes at School of Economics and Business, Norwegian University of Life Sciences (NMBU). Alfnes is also Adjunct Research Professor at SIFO.

Acknowledgements

This thesis has mine name on it, but it is the result of contribution from many other people.

Primarily I would thank my supervisors Frode Alfnes at NMBU/SIFO and Svein Ole Borgen at SIFO. Both have done a splendid job in supervising, because of their huge knowledge in economics, consumer studies and food marketing.

Secondly, I would like to show my gratitude to the partners in the project: Researcher Cecilie Mejdell, Research Fellow Kristian Ellingsen and Research Fellow Julie Føske Johnsen at the Norwegian Veterinary Institute. From NMBU, I would like to thank Head of Department of Companion Animal Clinical Sciences, Ann Margaret Grøndahl and Associate Professor Tore Sivertsen. Grøndahl and Mejdell have been head of the project in different years. They have done a splendid job in inspiring me and let me be part of something bigger than just my studies.

The other partners in the project, who also deserve my gratitude, are: Professor Dan Weary at The University of British Columbia in Canada, Researcher Anne Marie de Passille at Agri-Food Research Centre in Canada, Researcher Cynthia Verwer and Managing Director Jan-Paul Wagenaar at the Louis Bolk Institute in the Netherlands, Professor Lena Lidfors and Assistant Professor Helena Röcklinsberg at Swedish University of Agricultural Sciences, Researcher Leif Jarle Asheim at Norwegian Agriculture Economics Research Institute, Professor Knut Bøe at Department of Animal and Aquacultural Sciences, NMBU, Researcher Emma Brunberg at The Norwegian Institute for Agricultural and Environmental Research, and PhD Fellow Britt Henriksen at Aarhus University.

Members of the steering board in the project also have been good at directing my research so it is relevant for policy makers and other users. I would like to thank Jon Magne

Holten, Jan-Widar Finden, Arne Magnus Aasen and Geir Grosberg. Both project partners and the steering board have contributed to better and more pinpointed research.

I have also had much good help from non-academic contributors in the project. Hans Arild Grøndahl at Grøndalen farm for being so enthusiastic about the project and introducing me to the restaurants purchasing his veal. Further, former chef at Radisson Blu Plaza Hotel, Stephan Hårdi deserves my gratitude for letting me conduct an experiment at his restaurant. I would also express my gratitude to category managers Bent Karlsrud and Ole Førre Skogstø at REMA 1000 for letting me conduct experiments in their stores and providing me with unique sales data.

Further, I would like to thank my present and former colleges at SIFO. You are the best colleagues or friends actually, that I could have. You have supported me in a professional and a social way.

The other PhD students at NMBU School of Economics and Business are also people that have supported me. They have let me be a member of a great team, both professional and social. It is good to know that there are other people that experience the same struggle as I have.

As special thank goes to my former colleague Marthe Hårvik Austgulen for reading part I of this thesis and given me many useful comments. The remaining errors and obscureness are my to blame alone.

Finally, I would like to thank my family for supporting me and showing me the benefits of a long education. I would also appreciate my gratitude to Diana for some nice months while completing this thesis.

Oslo, June 2 2014

Abstract

This thesis uses experimental methods to test Norwegian consumers' preferences for animal welfare. First, it investigates the benefits of using such methods in consumer studies of credence goods in general, and in the case of animal welfare in particular. Consumer researchers have not used such methods extensively, so using them to examine key issues in consumer research is important. It is also important to test the potential of these methods and to see whether they provide the same results as previous studies that use other methods. Second, it examines what consumers think about animal welfare in different sales channels and for different food products.

This thesis consists of four papers. Paper 1 investigates restaurant guests' interest in animal welfare in an haute cuisine restaurant in Oslo. More specifically, the paper has two objectives. First, the paper shows how to conduct experiments in haute cuisine restaurants without disturbing the guests. Several pros and cons of such experiments are discussed. The second objective of the paper is to investigate whether describing an organic veal course as animal-welfare friendly or organic in the menu influenced sales. In addition to manipulation of how the course was described, the price of the course was changed. The result was a small but significant sales increase when describing the course as animal-welfare friendly. There was no effect from describing the course as organic. However, changing the course's price had clear effects. Sales were low when the price was high, which is normal, but also when the price was low. The latter finding is possibly because guests think price is connected to quality, or because they think choosing one of the cheaper courses signals that they are stingy.

Displaying such behavior to a business partner or date is probably something that most people would like to avoid.

Paper 2 also examines veal, but uses another experimental methodology. This paper uses an online choice experiment to measure consumers' propensity to buy foreign food if this food is labeled as organic. From many previous studies, it is well known that consumers prefer food from their home country. Internationally, for example in the European Union, there is a goal to increase international trade of food within the Union. A measure designed to reach this goal is quality labels, such as the mutual EU organic label. These labels are intended to overcome consumers' skepticism toward foreign food. If consumers see a trustworthy label on foreign food, they are more likely to buy it (given they know the label and trust it). The results from the choice experiment in paper 2 cast doubt on this view, suggesting that most Norwegians are not willing to give up Norwegian domestic veal in favor of foreign veal. Furthermore, whether the foreign veal is labeled as organic or comes from a country with a similar culture, such as Denmark, has no influence. The only prospect for increased international trade is among regular organic buyers. The experiment shows that this segment is willing to pay more for Danish organic veal than for Norwegian conventional veal.

Paper 3 tests whether positive or negative animal welfare labeling has the greatest effect on sales of eggs in a Norwegian grocery chain. The egg cartons associated with poor animal welfare (eggs from battery production) were clearly labeled using negative labeling, while the eggs associated with good animal welfare (organic eggs) were clearly labeled using positive labeling. The results show a strong effect of negative labeling, but no effect of positive labeling. This result is in line with the majority of consumer labeling studies and psychological experiments more generally.

Paper 4 is a book chapter discussing the use of natural field experiments in consumer studies. Natural field experiments are field experiments (i.e. experiments conducted outside

the laboratory) designed so that the participants do not know they are part of the experiment. The paper argues that this feature is especially important in empirical consumer studies of consumer goods with an ethical dimension. Examples include environmental issues, child labor and animal welfare. The paper concludes that a method in which participants do not know they are being monitored will provide new insights, but this is only a supplement to other methods traditionally used in consumer research.

The overall contribution is to show how consumer preferences related to food quality can be studied using different methods, and that combining methods can provide new and important insights, because different methods reveal different elements of this field of study. Natural field experiments have some desirable features, as they force participants to reveal their true preferences (i.e. incentive compatible) and are conducted in the setting we are interested in studying. Conducting natural field experiments in stores and restaurants provides a new perspective and valuable new insights into how consumers actually make food choices.

Key words: Animal welfare, consumer studies, natural field experiment, choice experiment, credence good

Norsk sammendrag

Denne avhandlingen benytter eksperimentelle metoder for å teste norske forbrukeres preferanser for dyrevelferd. For det første undersøker den fordelene ved å benytte slike metoder i forbrukerstudier av tillitsgoder generelt og i tilfellet ved dyrevelferd spesielt. Forbruksforskere har ikke benyttet slike metoder veldig mye, slik at å benytte dem til å undersøke nøkkelspørsmål i forbruksforskning er viktig. Det er også viktig å teste potensialet for slike metoder og se om de gir de samme resultatene som tidligere studier som benytter andre metoder. For det andre undersøker avhandlingen hva forbruker mener om dyrevelferd i ulike salgskanaler og for ulike matvarer.

Denne avhandlingen består av fire artikler. Artikkelen 1 undersøker restaurantgjesters interesse for dyrevelferd i en finere restaurant i Oslo. Mer spesifikk, har artikkelen to mål. For det første viser artikkelen hvordan man kan gjennomføre eksperimenter i finere restauranter uten at gjestene forstyrres. Ulike fordeler og ulemper ved slike eksperimenter blir diskutert. Den andre hensikten med artikkelen er å undersøke hvorvidt en beskrivelse av en hovedrett med kalv som dyrevennlig eller økologisk i menyen påvirker salget av retten. I tillegg til å manipulere hvordan retten var beskrevet ble prisen på retten endret. Resultatet var en liten, men signifikant salgsøkning når retten var beskrevet som dyrevennlig. Det var ingen effekt av å beskrive retten som økologisk. Imidlertid var det klare effekter av å endre prisen på retten. Salget var lavt når prisen var høy. Dette er normalt. Det var også lavt salg når prisen var lav. Dette siste funnet skyldes trolig at gjestene mener at pris henger sammen med kvalitet eller at valg av en billig rett signaliserer at man er gjerrig. Det å avsløre en slik holdning til en forretningspartner eller en date er trolig noe de fleste mennesker vil unngå.

Artikkel 2 handler også om kalvekjøtt, men benytter en annen eksperimentell metodologi. Denne artikkelen benytter et online valgekspirement for å måle forbrukernes tilbøyelighet for å kjøpe utenlandsk mat når denne maten er økologisk merket. Fra mange tidligere studier er det kjent at forbrukere foretrekker mat fra hjemlandet. Internasjonalt, for eksempel i den Europeiske Union, er det et mål å øke handelen med mat innenfor unionen. A tiltak for å nå dette målet er kvalitetsmerker, slik som det felles merke EU har for økologisk mat. Disse merkenes intensjon er å motarbeide forbrukernes skepsis til utenlandsk mat. Hvis forbrukerne ser et troverdig merke på utenlandsk mat, vil de være mer tilbøyelige til å kjøpe maten (gitt at de kjenner til merket og stoler på det). Resultatene fra valgekspirementet i artikkel 2 betviler dette synet fordi det viser seg at de fleste nordmenn ikke er villig til å bytte bort norsk kalvekjøtt til fordel for utenlandsk kalvekjøtt. Det viser seg til og med at hvorvidt kjøttet er merket som økologisk eller kommer fra et land med nokså lik kultur, som Danmark, ikke har betydning. Den eneste muligheten for økt internasjonal handel synes å finnes hos forbrukere som kjøper økologisk mat jevnlig. Eksperimentet viser at dette forbrukersegmentet er villig til å betale mer for dansk økologisk kalvekjøtt enn for norsk konvensjonelt kalvekjøtt.

Artikkel 3 tester hvorvidt positiv eller negativ dyrevelferdsmerking har størst effekt på salget av egg i en norsk dagligvarekjede. Eggekartongene assosiert med lav dyrevelferd (dvs. egg fra høns i bur) var tydelig negativt merket, mens eggene assosiert med høy dyrevelferd (dvs. økologiske egg) var tydelig positivt merket. Resultatet er en sterk effekt av negativ merking, men ingen effekt av positiv merking. Dette resultatet er i tråd med majoriteten av forbrukerstudier av merker og psykologiske eksperiment mer generelt.

Artikkel 3 er et bokkapitel som diskuterer bruk av naturlige felteksperimenter i forbrukerstudier. Naturlige felteksperimenter er felteksperimenter (dvs. eksperimenter utført utenfor laboratoriet). Eksperimentene er designet slik at deltakerne ikke vet at de er med i et eksperiment. Artikkelen argumenterer for at denne egenskapen er spesielt viktig i empiriske

forbrukerstudier av forbruksvarer med en etisk dimensjon. Eksempler inkluderer miljøspørsmål, barnearbeid og dyrevelferd. Artikkelen konkluderer med at en metode der deltakerne ikke vet at de blir overvåket vil gi ny innsikt, men metoden er likevel bare et supplement til tradisjonelle metoder i forbruksforskning.

Avhandlingens overordnede bidrag er å vise hvordan man kan studere forbrukerpreferanser knyttet til matkvalitet ved hjelp av forskjellige metoder, og at ved å kombinere metoder kan man få ny og viktig innsikt. Naturlige felteksperimenter har noen ønskede egenskaper, som at de tvinger deltakerne til å avsløre sine sanne preferanser (dvs. incentivkompatibilitet) og er gjennomført i den settingen vi er interessert i å studere. Ved å gjennomføre naturlige felteksperimenter i butikker og restauranter får man et nytt perspektiv og ny innsikt på hvordan forbrukere faktisk gjennomfører sine matvalg.

Nøkkelord: Dyrevelferd, forbrukerstudier, naturlige felteksperiment, valgekspertiment, tillitsgode

List of papers

This thesis is based upon the following four papers:

1. Schjøll, A. and F. Alfnes (2014): “Eliciting Consumer Preferences for Credence Attributes in an Haute Cuisine Restaurant.” Submitted to *Scandinavian Journal of Hospitality and Tourism*.
2. Schjøll, A. (2014): “Country-of-origin preferences for organic food.” Submitted to *European Review of Agricultural Economics*.
3. Schjøll, A.; F. Alfnes and S. O. Borgen (2014): “Animal welfare in stores: natural field experiments with positive and negative labeling.” Second submission to *Food Policy*.
4. Schjøll, A.; F. Alfnes and S. O. Borgen (2013): “Attitude versus action for farm animal welfare: what can we learn from natural field experiments?” In L. Hansson, U. Holmberg and H. Brembeck (Eds.) *Making Sense of Consumption*. Gothenburg: University of Gothenburg. Pp. 231-245.

Part I: Introduction

1. Introduction

There is increasing interest in and use of credence attributes in research related to food products, such as organic production, country of origin and animal welfare. This thesis presents results from experiments conducted in a restaurant, in grocery stores and on the Web to investigate and better understand consumer preferences and behavior related to these attributes. Two types of products have been investigated: veal and eggs. The veal study was a requirement of the project of which this thesis forms a part, and eggs were included because credence attributes related to production can be very important in differentiating the alternatives.

1.1 Credence attributes related to the production of food products

Darbi and Karni (1973) introduced the term credence attributes, which are attributes that a consumer can never verify during the lifespan of a consumer good. Unlike search attributes, they are not verifiable before purchase, nor are they verifiable after consumption, unlike experience attributes. In short, credence attributes can never be verified. A consumer then needs to trust labels, or any other signal a producer uses to claim a credence attribute.

For consumers, there are many relevant credence attributes in the food market. Examples include animal welfare (Dentoni et al. 2011; Heerwagen et al. 2013), organic production (Aertsens et al. 2009; Van Loo et al. 2012), country of origin (Cicia and Colantuoni 2010; Lobb and Mazzocchi 2007) and fair trade (Rygh 2007; Teyssier et al. 2012).

As long as consumers allow credence attributes to play a role in their choice decision, it is important for producers and retailers to know which attributes consumers care about in order to offer popular products. Such information is also valuable for policy makers who design the regulations that affect food producers. Animal welfare is heavily regulated, and includes factors such as amount of space, access to pasture, etc. The animals' needs are, of

course, the most important factor in regulating agricultural production. For example, the objective of the Norwegian animal welfare laws is to promote good animal welfare and respect for animals (Lovdata 2009). However, it is important that the regulation is consistent with ordinary people's understanding of animal welfare, because regulations in general should mirror the national will, otherwise people will not obey the regulations.

In addition, special interest groups, such as animal welfare organizations, like to know consumers' attitudes toward credence attributes related to their cause in order to generate support for their activities. Interest groups' campaigns are more likely to be successful if they are consistent with consumers' attitudes.

Credence attributes are not only difficult for consumers; they are also difficult for researchers to study as well. Many credence attributes are related to ethical issues. Norms guide us in dealing with ethical issues, but norms are only consultative, and we are not obliged to follow them in the same way that we are obliged to follow laws. However, not following norms comes at a cost, in that while you might not be sent to prison, you are likely to be punished in a less severe way. Few people are likely to admit that they buy food that is harmful to the environment and is produced under poor animal welfare conditions, but this does not mean there is no market for "unethical" alternatives. A large selection of products that are low-priced compared with ethically produced products may indicate that in general, price is a more important choice attribute than most credence attributes. It seems that saving money is more important than saving the world in an individual choice setting.

Admitting that one chooses the least ethical alternative often comes at a social cost. In the market, this social cost is balanced with the price, but in surveys and focus groups there is no cost associated with stating that you buy the superior product. Therefore, people too often state that they buy products of high ethical quality in surveys or focus groups because their

claims will never be scrutinized. Consequently, methods measuring actual actions seem necessary in order to obtain a broader picture in consumer studies with ethical aspects.

1.2 Experiments conducted in restaurants, grocery stores and on the Web

Economic experiments are part of a new branch of economics known as behavioral economics. Behavioral economics improves the realism of economic models by using psychological underpinnings within the economic fundament (Camerer and Lowenstein 2004: 3). Initially, only laboratory experiments were used, and these were conducted in a closed environment where the experimenter can ideally control “everything”. One of the earliest, and probably most cited, economic experiments is described in Kahneman and Tversky (1979). They used very simple laboratory experiments to test whether expected utility holds. Their rejection of the expected utility theory led to the development of prospect theory, and later a common understanding among most economists that economic theory should be experimentally tested if possible.

Daniel Kahneman received The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel in 2002 for integrating insights from psychological research into economic science (Nobelprize.org 2013a). The other laureate that year was Vernon L. Smith, who received the price for “having established laboratory experiments as a tool in empirical economic analysis, especially in the study of alternative market mechanisms” (Nobelprize.org 2013b). Using his double-prize auction, he shows in Smith (1962) that the market cross, with a decreasing demand curve and increasing supply curve, actually existed in small samples and with little information among market actors. “Market theory actually works, so evidence was not necessary” was the critique Smith met (Smith 1991: 151). Experimental economists have used laboratory experiments to both support and reject various economic theories, as seen from these two examples.

Economic scholars have become interested in the effect of context on how people behave. People behave differently in different contexts, so testing for this is necessary. In addition to context, increased relevance has been an important reason for conducting experiments in the field (Harrison and List 2004). As a result, the number of field experiments has increased. For the complete history of field experiments, see List (2011).

Parallel to the development of laboratory and field experiments, researchers have developed design and estimation methods for choice experiments (CEs). Jordan Louviere and Daniel McFadden are two of the most important contributors to the development of CEs, a methodology based upon random utility theory (RUT). The basic idea is that an individual would choose the option from a choice set if and only if the utility from a specific alternative is larger than the utility from all other alternatives (Louviere et al. 2010). When consumers choose, many attributes are considered. The researcher does not know all of these attributes or how consumers weight them a priori. As a result, consumers' choices seem random to the researcher. Therefore, RUT models, estimated by discrete choice methods where the estimated parameters maximize the probability of an individual choosing the specific choice made, seem appealing.

In a CE, a good is a bundle of independent attributes. A steak, for example, would differ from other steaks in terms of breed, fat content and place of production, just to mention a few attributes. Consumers care about these attributes, and the main objective of a CE is to evaluate these attributes.

In a CE, the respondent makes a discrete choice; he only chooses one object from a range of alternative objects, i.e. one steak from a collection of many steaks. Daniel McFadden (1976, 1968) develops a multinomial logit model, the most basic choice model (Hensher et al. 2005: 308). Unlike the standard logit model, this model allows the dependent variable to have more than two outcomes. McFadden's work laid the foundation for an array of econometric

models describing discrete choice. Today a CE is a powerful tool for analyzing consumer preferences, and new techniques are constantly emerging, mainly because of continually improving computer technology.

Most researchers conduct their CEs online these days, as a part of a Web survey. This makes data collection easy, and one can quickly obtain information about thousands of choices made by the participants. Because CEs are easy to conduct and very flexible, they have been used to study many topics, for example choice of airline (Eckert et al. 2012), choice of cured ham (Gracia et al. 2011) and choice of place for recreational fishing (Hunt et al. 2010).

One way to distinguish between different methods used in consumer research is analyzing whether they use stated or revealed preferences. Focus groups, surveys and most CEs are examples of stated preference methods, where participants state their attitudes. Field experiments and most laboratory experiments use revealed preferences. With revealed preferences, participants' actual actions are measured.

1.3 Overview of the thesis

This thesis is part of the project “Developing robust and economically viable models for cow-calf suckling in organic dairy systems, fulfilling high standards for animal health, welfare and ethics”. Hence, an examination of veal was required by the project. Norway has a low level of consumption of veal compared with most other European countries (European Commission 2012). According to Norwegian veal producers, most of their production is currently sold to restaurants, which are interested in premium products.

As most veal is sold to restaurants, we (my supervisors and I) decided to conduct a restaurant experiment, and contacted a restaurant that we knew had purchased veal from the largest organic veal producer. The results of this experiment are presented in paper 1. To examine the preferences of more ordinary veal products, we conducted a Web survey of

minced veal including factors such as place of origin, organic production and animal welfare labels. The results of this Web survey are presented in paper 2.

Because of certain limitations in the case of veal (i.e. low sales of veal and few known products), we decided to analyze a second case in which we could investigate consumer behavior related to animal welfare in stores. We examined eggs, because they are a product most Norwegians buy on a regular basis. In addition, grocery stores use animal welfare as one of the attributes to differentiate their types of eggs. New regulations related to the production of eggs have also changed the products available in the stores. Fortunately, Norway's second-largest grocery chain decided to go one step further than the new regulations. They first introduced a new carton with negative animal welfare information, and later removed the lowest animal welfare egg alternative. These changes, together with our own manipulations, form the basis for papers 3 and 4.

1.4 Selection of methods

Why people eat as they do is a question that has motivated consumer researchers for decades. A better understanding of food choices will benefit all stakeholders. Producers would like to know more about food preferences in order to supply products more in line with consumers' needs, and hence increase profitability. Meanwhile, consumers will enjoy increased utility if products are more in line with their needs. Policy makers should also care about consumers' food preferences in order to prevent lifestyle diseases and maintain food security.

Researchers have put considerable effort into determining why people eat as they do. However, a problem often arises when dealing with consumers' food preferences: they are difficult to measure, hence measurement errors of various types are likely to occur.

Animal welfare is a good case study when examining credence attributes related to food, because it is an ethical issue whereby political correctness often casts a shadow over

consumers' true attitudes. That is, consumers often give politically correct answers instead of disclosing their real attitudes and typical behaviors. See for example Lusk and Fox (2003).

In order to obtain a broader perspective on consumers' preferences for animal welfare, we use new methods to determine whether the results presented in the literature hold. Field experiments have many advantages (see below), but in general it is good to use different methods to examine a phenomenon. Alexander et al. (2008) outline several reasons for using multiple methods. First, multiple methods increase the validity and reliability of the research conducted. This is relevant for consumer studies of animal welfare because most studies conducted to date are studies where participants know they are being monitored. For ethical issues such as animal welfare, this may imply deviation from true preferences.

Complementarity is also a reason for using different methods in a study. For example, we are not aware of any field experiments in the study of animal welfare. Using different methods will most likely reveal different dimensions of the complex issue of animal welfare. Finally, by using different methods, we seek to develop more accurate research instruments. Most studies of consumer preferences for animal welfare use different stated preference approaches, i.e. respondents can state something without being held accountable for their answers. In some of the studies presented here, participants choose and buy real products in stores and restaurants. Hence, revealed preference methods are used.

Economic experiments are very suitable for consumer studies of animal welfare. This thesis presents several controlled experiments. Holding everything constant except the item under investigation is a characteristic of a controlled experiment (List 2011). Even though there are numerous explanations of why consumers eat as they do, experiments can provide new knowledge because they use randomization as an instrument variable, balancing unobserved variables with the observed and treated variables (op cit.). By doing this, in

combination with proper use of statistical significance testing, we can ensure that the findings are reliable.

This thesis uses two field experiments, because other researchers have not conducted such experiments investigating consumer preferences regarding animal welfare. List (2011) offers three advantages of field experiments. (1) They provide a distinctive and new source of empirical evidence. Researchers have not conducted any field experiments previously in the field of consumers' preferences regarding animal welfare. Therefore, it is appropriate to introduce this method within this field of research. (2) Conducting field experiments provides an opportunity to specify and address economic questions of interest. The alternative would be waiting for the event to occur or hoping that a proper econometric specification makes it possible to test the relationship. In Norway, there is debate over whether animal welfare legislation should be changed; hence, empirical evidence is needed now. (3) Field experiments strengthen the connection between economic theory and empirical evidence from the real world. This is highly relevant for animal welfare, where research on consumer preferences often produces results that are opposite to real market behavior. See Andersen (2011) for an example.

The field experiments used in this thesis are all natural field experiments. Such experiments are conducted in the field without participants knowing they are part of an experiment (List 2011). This has the advantage that strategic behavior from participants is eliminated, and the researcher is able to monitor "true" behavior. These experiments are good at separating cause and effect, and use a completely random sample, because no one actively decides to participate in the experiment (op cit.).

Consumers seem to be interested in the ethical aspects of the food they eat. Of course, individual factors as taste, price, brand loyalty, etc., are the most important factors when consumers select their food (see paper 3), but ethical aspects such as child labor,

environmental issues and animal welfare also play a role. Growth in alternative food sales channels may be due to increased interest by consumers in the social aspects of food production.

2. Research questions

This thesis seeks to answer three research questions:

1. What can we learn from using natural field experiments in consumer studies? What new knowledge can be acquired that was not obtainable using other methods?
2. How does consumers' interest in animal welfare vary between types of food products (veal and eggs) and sales channels (restaurants and grocery stores)?
3. How do consumers grade different food-related credence attributes, for example country of origin versus animal welfare?

3. The cases in the thesis

3.1 Overview of cases

All four papers in this thesis present data from cases illustrating different aspects of consumers' actions related to animal welfare in Norway. The following figure summarizes the cases:

Paper	Objective	Empirical method
1: Veal in a high-class restaurant	Test whether organic is a more attractive attribute than “pure” animal welfare.	Testing willingness to pay (WTP) from restaurant menu descriptions. Main concept: does it matter whether the menu promotes the veal course as organic or animal friendly?
2: Veal in a stated preference choice experiment	Test whether it is animal welfare labeling, price, or country of origin that is most important when consumers buy prepacked meat.	A choice experiment conducted online with a representative sample of Norwegians. Respondents rank three packages of prepacked veal with different attributes.
3: Eggs in grocery stores	Test whether it is positive or negative information about animal welfare that influences sales the most.	Test different types of animal welfare labels on eggs in a grocery store setting. Eggs produced with a low level of animal welfare had a negative label, while eggs produced with high level of animal welfare had a positive label.
4: Natural field experiments in consumer studies	A theoretical discussion of the benefit of using natural field experiments in consumer studies.	Present the same case as in the previous paper, but provide a general discussion of the pros and cons of natural field experiments.

Figure 1: The cases presented in the thesis

We study consumers’ preferences for veal as part of the broader project of which this thesis is a part. However, as argued below, studying veal is also justifiable from a scientific viewpoint. We chose to also examine preferences for eggs independently.

By selecting these cases, the thesis seeks to satisfy several criteria in case selection. First, it maximizes the variation in maturity. Organic veal is a more or less nonexistent product in the Norwegian meat market. As mentioned, the consumption of organic meat in general is very low in Norway. Per capita consumption of veal (conventional and organic) was only 400 g in 2012 (Animalia 2013). It is not easy to study the Norwegian veal market for either conventional or organic veal, simply because the market is not yet well established. However, the organic-egg market is very different. Organic eggs have a decent market share, and are available in most grocery stores in Norway. Therefore, this is a product with which

Norwegians have a relationship, unlike veal. Another argument along the same dimension is the communication of animal welfare information. It is difficult to find meat labeled as animal friendly in Norway, but in the egg market, animal welfare is a well-established choice attribute, despite the fact that no label only dealing with animal welfare exists. This implies that consumers should choose organic food because in general that is the most animal-welfare-friendly alternative. A recent Norwegian review concluded that there are small differences in animal welfare and health when one compares conventional and organic production, but there is a slight tendency toward higher welfare in organic production (Norwegian Scientific Committee for Food Safety 2014).

Using Patton's (1990) terminology, this thesis uses two kinds of case selection techniques. The first is maximum variation, where one seeks variation along one dimension. With the selection of one nonexistent market (organic veal) and one well-established market (organic eggs), all possible cases are, in a sense, covered when it comes to the extent of animal welfare within a segment because the extremes are covered. The veal market can learn something from the egg market in terms of how to become a successful product.

The second kind of case selection technique used in the thesis is sampling of politically important cases. In some European countries, consumption is quite high, so organic veal in Norway may have a decent market potential based upon sales in some other European countries (European Commission 2012). Most likely, the project of which this thesis forms a part received funding because raising calves for slaughter is not common in Norway, and the government may want to increase calf production in order to produce more meat. In Norway, the most common cow breed is the "Norwegian Red", a breed that is suitable for both meat and milk production (Vangen et al. 2007). These days, Norway produces too much organic milk compared with demand. It is not possible to export the surplus milk given Norway's current agricultural policy (see Section 3.2 for an explanation), so it has to be used

domestically. To date, the solution has been to sell it as conventional milk; in fact, only 38 percent of organically produced milk is sold as organic (Norwegian Agricultural Authority 2013). Of course, this means a loss for dairies, because conventional milk has a lower selling price compared with organic milk. Hence, any solution that provides a way to reduce the number of dairy cows without closing down farms (again, see to Section 3.2 for an explanation of why this is not an option) is of interest to politicians. Therefore, the organic veal case is important because it is both a market that is not yet fully established and a market that may offer a solution to practical market problems.

Even though there are only two cases of animal welfare studied in this thesis, it should be possible to generalize from these cases. There are at least two reasons why this might be so. First, by studying veal and eggs, we have one food product (veal) that requires animals to be slaughtered and one product (eggs) that is produced by animals that remain alive. It could be that consumers think more of the slaughtering process when they think about animal welfare, and therefore are more concerned about animal welfare in meat production than in egg production. Second, the cases are very different. Most Norwegians eat eggs. Per capita consumption of eggs in 2012 was almost 13 kg (Animalia 2013). Consumption of veal, however, is limited in Norway. It is likely that organic veal will never be a high-volume product, even though there is a potential market for it. However, Flyvbjerg (2006) argues that such “hopeless” cases have a value because if they are successful, all cases will be successful. Therefore, a case with limited potential, such as organic veal, is useful for identifying the critical factors for success.

To sum up the case selection, organic eggs are suitable because they are an extreme case compared with organic veal. In addition, meat is a product from an animal that lived and was slaughtered. Eggs, in contrast, are a product from an animal that still lives. In this sense, both aspects of animal welfare are investigated.

3.2 Agricultural production in Norway

In a Western European context, Norwegian agricultural policy is sufficiently different to warrant a brief review. Norway is not a member of the European Union (but it is a member of the European Economic Area), so the country has an independent agricultural policy.

However, it follows common EU regulations, for example in relation to animal welfare and organic production.

Organic food is a niche market in Norway, and the general market share is around one percent (Norwegian Agricultural Authority 2013). In the meat segment, the market share is especially low. Therefore, the future potential market for animal welfare products seems limited. However, there are positive aspects of the market worth noting. First, the market share for organic eggs is five percent (op cit.), so in some market segments Norwegians buy organic. Second, many consumers are interested in buying organic. The SIFO survey from 2013 shows that around 50 percent of Norwegians look for organic alternatives when shopping for food.

Norwegian agricultural policy deviates from the EU's policy in two ways. The first is with regard to the goals of the agricultural policy. The main goal for Norway is to stimulate food production, but there are other goals. However, these other goals may be in conflict with efficient production. According to the last white paper on agriculture to the Storting (the Norwegian parliament), agricultural production is an important part of Norwegian regional policy (Ministry of Agriculture and Food 2011).

Another way in which Norwegian agricultural policy is different relates to the role of free trade. In the EU, free trade is the mantra within the Union. Norway does not have a policy based upon free trade of agricultural products. Because Norwegian food production is more expensive than European food production, Norwegian products need to be protected from foreign competition. Import taxes, quotas and embargos are measures used to protect

Norwegian farming. For veal, there is a general embargo, which can be lifted when domestic production is too low to meet domestic demand. Due to problems meeting rising demand for beef in Norway, there have been greater imports in recent years. For eggs, there is a general embargo, and all eggs sold in Norway are Norwegian. Norwegian agricultural policy thus makes food sold in Norway more expensive than food sold in other European countries, based on current prices (Pettersen et al. 2011).

4. Methods used

This thesis uses different types of experiments to investigate consumers' preferences for animal welfare. As mentioned in the previous section, the two cases selected are very different, as are the types of experiments used.

The first paper presents a natural field experiment conducted in an haute cuisine restaurant. In the case of organic veal, there are several reasons why this method is appropriate. First, the small amount of organic veal that Norwegians eat is consumed in restaurants. There is no official consumption data in relation to organic veal, but the leading organic veal producers all sell to restaurants (own investigation). It is not possible to find organic veal in grocery stores; therefore, a restaurant is the proper sales channel to investigate. The reason why we use an haute cuisine restaurant, and not a family restaurant, is twofold. As Norwegians consume so little veal, they probably view it as something exclusive and expensive, food that one would expect to find in a good restaurant. In addition, many experiments have been conducted in restaurants, but very few of them take place in haute cuisine restaurants. This sales channel allows us to investigate whether choices in such restaurants differ from those associated with other sales channels for food. In an haute cuisine restaurant, the researchers must minimize disturbance to guests. The diners are paying a lot of money for their meal, and so the challenge is to obtain the required data without harming their

dining experience. In a natural field experiment, participants are most likely unaware that they are part of an experiment, and so conducting this type of experiment in this way helps to avoid disturbances.

Paper 2 also involves organic veal, but this time a CE method is used. The reasons for conducting this online experiment were to provide a different perspective on consumers' preferences and to make use of the advantages offered by a CE. The CE has several advantages that can contribute to an improved understanding of the product being investigated. First, it is possible to test hypothetical situations in a CE. As mentioned, organic veal is not available in grocery stores, so it is not possible to do field experiments involving this product. A field experiment for veal in grocery stores would require the development of a real organic veal product to sell in grocery stores so that sales could be observed, which is beyond the scope of this thesis. Therefore, this paper develops a "virtual" product instead. An online experiment where respondents make choices based on pictures of products, instead of the actual products themselves, is a way to avoid production development costs. By letting consumers choose between pictures of prepacked veal, one mimics the choice situation at the meat counter in a grocery store. Thus, it is possible to test the product in a sales channel, even though it does not exist in this sales channel.

The CE is a stated preference technique, while the restaurant experiment is a revealed preference technique. The restaurant experiment provides a limited data set because of the amount of meat available, and the fact that each guest makes one choice only. In a CE, each respondent makes several choices, ensuring a rich dataset, which makes it possible to test various econometric models, revealing patterns that are difficult to find in the field.

In the third paper, we examine organic eggs in grocery stores. As mentioned earlier, eggs are consumed regularly by most Norwegians. There have been many surveys and laboratory CEs conducted in relation to eggs, but there have been no field experiments.

EU has issued a ban on traditional cages for laying hens. This is the most extensive change in animal welfare legislation in recent years, and it is therefore important to study consumers' reaction. Grocery stores are probably the most important sales channel for eggs, making it natural to conduct experiments in a store setting. In addition, by including grocery stores in the thesis, we cover both food eaten at home (from grocery stores) and food eaten away from home (restaurants).

As mentioned above, there are no field experiments on consumers' preferences toward animal welfare. However, there are several laboratory experiments (Section 5 provides an overview). List (2007) argues that field experiments can provide a link between laboratory experiments and naturally occurring data. On the one hand, we have data from various CEs, showing that consumers care about animal welfare; see for example the meta-analysis by Cicia and Colantuoni (2010). On the other hand, we have market data (scanner data) indicating that food produced using high standards of animal welfare does not sell very well. In a CE, the setting is strongly controlled, while in natural field experiments the researcher has little control over the environment. On the "control scale," data from these two methods represent the extremes. Field experiments are positioned more in the middle of this distribution, and therefore are also useful for highlighting certain results from CEs or naturally occurring data.

A common critique of field experiments is the lack of sample representativeness, and hence the results are not generalizable. List (2007) argues that sample representativeness is not the most crucial aspect of generalization, and that a representative environment is often more important. The laboratory, or a survey, is a strange environment for the participants, which may potentially change respondents' behavior, leading to non-generalizable results, even though the sample is representative.

This thesis presents the results of field experiments on veal and eggs. In these two experiments, it is most important to investigate the preferences of consumers interested in buying veal and eggs, not consumers that are outside the market, such as vegans for example. Hence, a representative sample of all Norwegians would be of little use, because not all Norwegians are in the market for these food products.

The thesis aims for maximum variation in sales channels, just as it aims for maximum variation in cases. In the grocery store setting, we study animal welfare in an ordinary consumption situation, while in the restaurant setting; we study animal welfare in a unique way because most consumers seldom visit haute cuisine restaurants. Finally, the CE makes it possible to let Norwegian consumers rank different credence attributes simultaneously, not only separately as in most previous Norwegian studies.

5. Relevant consumer studies on animal welfare

Many studies have investigated consumers' preferences for animal welfare. Here, we provide a brief review of this literature. Only studies related to beef (there are no studies of consumers' preferences for veal) and eggs are included in this review.

5.1 Beef studies

The Welfare Quality® project from 2008 reviewed studies on consumers' preferences toward animal welfare in seven European countries. Figure 2 shows that animal welfare is of less importance for Norwegian beef buyers compared with beef buyers in other European countries.

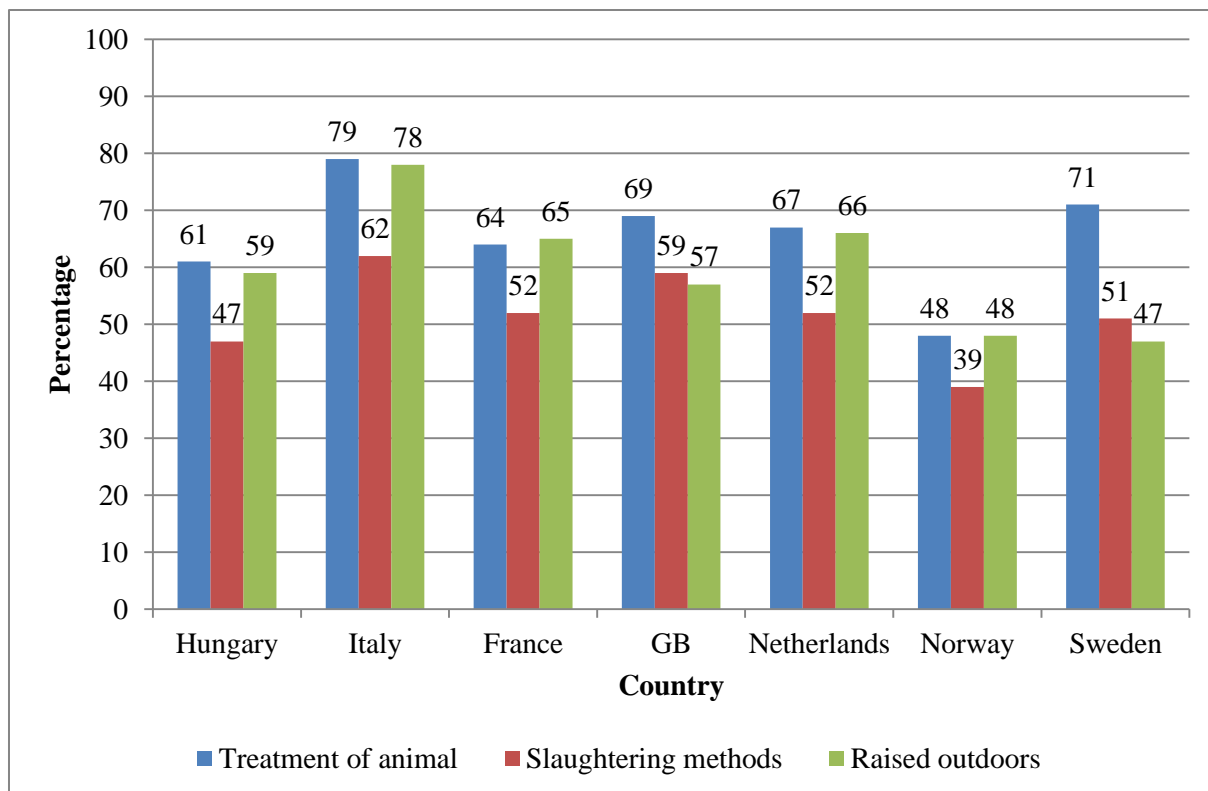


Figure 2: Percentage of people in various European countries answering ‘very important’ in response to questions regarding the importance of various aspects of animal welfare.

Source: Kjærnes and Lavik (2008), Table 3.3.

In Sweden, Carlsson et al. (2005a) find a higher willingness to pay (WTP) among consumers for a ban against the use of genetically modified (GM) cattle than for slaughter of cattle in mobile abattoirs. Furthermore, consumers have the lowest WTP for cattle that graze outside for the entire year. Consumers therefore seem to differentiate what they think is important when it comes to animal welfare.

In another Swedish study, Carlsson et al. (2007a) measure WTP for mobile abattoirs. In their CE, they estimate a WTP of 3.46 SEK/kg, which is higher than the estimated production costs of such abattoirs. Thus, there seems to be a WTP to avoid painful cattle transportation to the abattoir.

The last Swedish study worth mentioning is Carlsson et al. (2005b). In their CE, they measure WTP for several animal welfare attributes for different kinds of animals. For beef-

producing cattle, the WTP is highest for a ban on GM fodder. When animal welfare is linked to private food attributes such as food safety, the WTP for animal welfare is highest.

Carlsson et al. (2005b) also study animal welfare and milk production in Sweden. For milk, the highest WTP is also for a ban on GM fodder. A more interesting finding is a very low WTP for a long relationship between cow and calf after the calf is born. This thesis is part of a project that seeks to study the benefits of a prolonged relationship between cow and calf. This Swedish study finds a very low WTP among consumers for this welfare improvement. In our restaurant experiment, we used the term “happy calves” without explicitly stating what that expression meant. However, a relationship between cow and calf is an implied part of this “happiness.”

There exists one WTP study of calves’ welfare. The study by Schumacher et al. (2012) is not a consumer study, but investigates cattle feeders’ WTP for certified health programs that can improve calves’ health. Three programs were tested, varying in terms of number of days weaned. The most popular program weaned the calves for at least 30 days and a third party, such as a veterinarian, did the certification. The study also finds a positive WTP for age of calves and verification programs. The feeders agree that such programs will increase calves’ welfare and increase meat quality. However, they do not think the programs will lead to higher meat prices; the programs will therefore be profitable for consumers.

Another study conducted among US calf feeders is that of Schulz and Tonsor (2010). Their study investigates whether the feeders would like to pay for a traceability system for calves. The main reason why the feeders would like a national traceability system is the possibility that the system may prevent diseases. Disease prevention is clearly a costly part of animal welfare, so it seems that the feeders see the value of animal welfare.

There are not many Norwegian studies about consumers’ WTP for animal welfare. However, Alfnes (2004) studies the WTP of Norwegians for beef from various countries, and

the use of growth hormones. For consumers, hormone use is probably more related to expected taste and health risks rather than animal welfare. Nevertheless, US hormone-treated beef had the lowest WTP among the alternatives. In addition, respondents prefer domestic meat compared with foreign meat. If meat is foreign, it should come from neighboring Sweden. This result is the same as that found in paper 2 in this thesis.

In the US, Lusk and Fox (2002) estimate the WTP for two mandatory beef labeling schemes. Consumers are willing to pay 17 percent more for a labelling scheme that identifies hormone-treated beef compared with what they will pay for no labeling scheme. They are also prepared to pay 11 percent more for a label indicating the use of GM corn as fodder compared to no label. This result is also in line with the finding in paper 2, whereby consumers are very skeptical about the hormone treatment of food.

Bennett et al. (2002) find that among British students, 58 percent desire an export ban on live animals sent for slaughter within the EU and in countries outside the Union. From an animal welfare perspective, such transportation stresses the animal, and therefore transportation of meat products would be better. However, such a ban would lead to higher meat prices, because it is then more difficult to exploit economies of scope and scale. The WTP for such a ban is quite low, therefore this study shows that consumers do care about animal welfare, but are unwilling to pay for it.

Among beef consumers in Italy, Scotland and Spain, Bernués et al. (2003) conduct a cluster analysis with extrinsic attributes for beef. In their study, animal welfare is of high importance for younger people living in large cities, mostly in Italy, and for older people in Scotland living in medium-sized cities. My thesis does not conduct any segmentation analysis, so it is difficult to say whether the same findings would apply in Norway. Chilton et al. (2006) conduct a cost–benefit analysis of the welfare of dairy cows among consumers in Northern Ireland. They find a large positive net benefit associated with

improving cows' welfare, indicating that consumers are willing to pay to reduce lameness by increasing the opportunities for cows to lie down. The high WTP estimate probably reflects the fact that consumers can easily relate to this issue, just as they can in the case of moving hens from battery systems to free-range systems.

Napolitano et al. (2007) study the effect of sensory evaluation using information about animal welfare for Podolian cattle. These cattle from Southern Italy are traditionally reared under excellent animal welfare conditions. The study finds a higher grade of sensory evaluation when information about animal welfare and nutritional properties is presented compared with when it is not. Thus, there seems to be a correlation between taste and animal welfare. It is difficult to say whether the same correlation applies in Norway, but in paper 2, organic food (the food associated with the highest level of animal welfare) receives a low ranking in terms of taste but a high ranking in terms of animal welfare. For conventional food, the opposite is the case.

In non-hypothetical laboratory auctions, Dickinson and Bailey (2002) asked their US participants, who were connected to a university, to bid for roast beef sandwiches. It was possible to upgrade the sandwiches along several traceability dimensions. Basic traceability received the lowest WTP, while WTP for extra food safety assurances received the highest WTP. Animal welfare was in the middle between these two extremes. Trained professionals had the highest WTP for animal welfare, while professional staff had the lowest. Again, there is a positive WTP for animal welfare, especially when animal welfare is combined with food safety. Consumers seem to think about themselves first, then the animals. Norwegians seem to act in the same way, cf. paper 2.

A rather peculiar study of consumers' interest in animal welfare regarding beef is that of Tonsor and Olynck (2011). Instead of studying consumers directly, the study investigates the effect of animal welfare media coverage on US meat demand during the period 1982–

2008. First, the authors notice an increase in the number of articles covering animal welfare issues in the US media. Second, they find a significant, but weak, negative effect on beef demand in the long run. Finally, the study also finds a positive effect on beef demand of articles mentioning both the industry and consumer groups. They find the same effect for similar articles about the poultry industry, so beef and poultry seem to be substitutes when it comes to animal welfare. There have been no media studies of animal welfare in Norway, but the study of Tonsor and Olynck (2011) indicates that consumers use information about animal welfare received through the media when they make choices in the grocery store.

5.2 Egg studies

The Eurobarometer survey provides a broad overview of European consumers' attitudes toward animal welfare. However, Norway was not part of this survey. Figure 3 illustrates how much more consumers from selected European countries are willing to pay for increased welfare for the hens. The figure should be interpreted as follows. In EU 25 34 percent are not willing to pay any additional price for eggs produced in high-animal welfare production systems. 25 percent are willing to pay additional five percent more and so on. As can be seen in the figure, WTP for higher welfare is highest in the other two Scandinavian countries. One can therefore expect Norwegians to also have a high WTP, because the Scandinavian countries are very similar culturally.

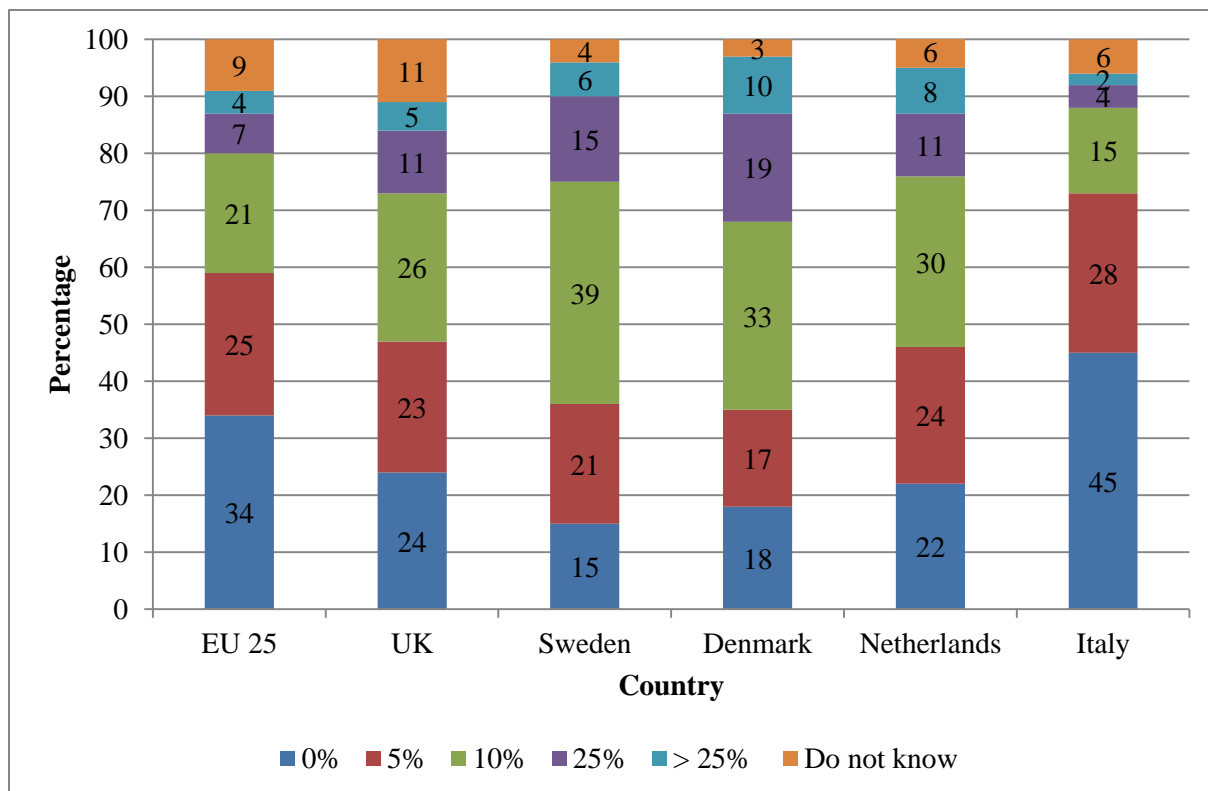


Figure 3: Additional price consumers in selected European countries are willing to pay for eggs produced in high-animal-welfare production systems.

Source: Eurobarometer (2005).

Andersen (2011) estimates the revealed WTP for animal welfare in egg production using a consumer panel in combination with a survey. From the panel, Andersen registers the price of purchased eggs, and from the survey she obtains data on attitudes toward animal welfare. She uses consumer purchase data, but does not conduct any experiments because there is no manipulation of the variables, just observations. She concludes that when consumers claim they purchase organic eggs for animal welfare reasons, they are often not being truthful. By comparing actual purchase behavior when prices fluctuate, she discovers that attributes other than animal welfare, such as brand awareness and store type, influence purchase decisions more. The socio demographic factor with the biggest impact on the purchase of organic eggs is whether or not the household is located in the capital city

(Copenhagen). This study comes closest to a field experiment among the consumer studies of animal welfare, and is therefore an important reference point for this thesis.

Gerhardy and Ness (1995) conduct a laboratory experiment in the UK about eggs. In their study, price is an important factor for only 10 percent of respondents, implying that it has no influence on demand for the majority. Furthermore, for most respondents, the production method is the most important attribute. Finally, one-third of respondents are focused mainly on the place of origin.

The main lesson from the study is that consumers are heterogeneous when it comes to preferences for eggs. A broad range of alternative products is therefore necessary to provide choice in the market. In paper 3, we test the market potential for increased sales of organic eggs, and it appears that the Norwegian market is more homogeneous than that in the UK, because there is less room for alternatives.

In the study by Carlsson et al. (2007b), it is found that Swedish consumers would be willing to pay SEK 10.84 for six eggs produced in a free-range system if the government bans eggs from battery systems. The authors call this the legislative solution. A market solution, where both free-range and battery systems exist but all eggs are labeled, results in a WTP of only SEK 8.40 for the free-range eggs. These two WTP estimates are not significantly different from each other; hence, one cannot justify a ban on battery production systems based solely on these results. Paper 3 sheds light on a “voluntary” ban on battery eggs, even the new enriched cages, because one grocery chain removed these kinds of eggs. Our paper shows that consumers are willing to switch to free-range eggs, but not to organic eggs.

Burrell and Vrieze (2003) analyze a representative sample of Dutch egg consumers in order to discover to what extent ethical motives are important when buying eggs. More precisely, they want to explore the paradox of why so few in their sample claim to purchase

eggs from high-animal-welfare systems while so many of their respondents state that hen welfare is important to them.

The individuals with the greatest concern about the welfare of the hens tend to be older females who are well educated, have high incomes, have a vegetarian in the household and are members of an animal welfare organization.

However, the household's per capita egg consumption, regardless of whether they pay attention to price, whether they live in urban or rural areas and whether they are aware of the factors influencing hens' welfare does not seem to significantly influence how they view the importance of the hens' welfare.

From their binary logistic model, where the choice is to buy battery eggs or not, they have the following main findings:

- Ethical intention (intention to buy measured in WTP) cannot predict whether or not a consumer will buy battery eggs.
- Adding ethical concern (how much they care) as an independent variable in the regression removes the influence of ethical intention. Thus, in the model, intention is a better measure of action than concern.
- Ability to recognize the rearing system's label and consumer attitudes are important factors modifying the link between concern and behavior. In particular, concerns about the price are more important than income in the prediction of the purchase of battery eggs.
- Consumers living in rural areas are more in favor of buying battery eggs than consumers living in urban areas (consistent with the Danish study).

Overall, the study concludes that intention alone cannot explain consumers' behavior. However, what can explain behavior is price and rearing system. Gender and age are

important factors in determining ethical concern, but these factors cannot explain how concern translates into ethical behavior. More specifically, membership of an animal welfare organization does not act as a substitute for ethical consumerism. Furthermore, the study reveals a low level of knowledge about hen-friendly production systems, which can explain the weak link between ethical concern and ethical behavior.

In the study, respondents are willing to pay 32 percent more for eggs produced by hens that have experienced a higher level of welfare compared with battery eggs.

The study of Burrell and Vrieze (2003) is an important reference point for the animal welfare studies presented in the thesis. Using field experiments means it is difficult to determine why consumers behave the way they do, therefore we can only expect to see what they do.

Therefore, this thesis cannot analyze consumers' intentions.

In Italy, it is mandatory to label all egg cartons with a one-digit code indicating the production method. In a study by Vecchio and Annunziata (2011), only 11 percent of the respondents were able to give a precise explanation when shown an egg carton and asked to explain the meaning of the code. In their survey, 58 percent of respondents claimed to be very interested in the welfare of egg-laying hens, while more than 50 percent thought the welfare of hens in Italy was inadequate.

A mandatory labeling scheme for eggs in Norway has been debated (Rotevatn 2014). There is no text indicating the production method on most Norwegian battery-egg cartons, however on non-battery-egg cartons, the production method is clearly stated. This was one of the main reasons for conducting one of the experiments in paper 3.

Even though consumers state that animal welfare is important, the study by Vecchio and Annunziata (2011) concludes that freshness, appearance and nutritional value are the most important choice attributes for Italians when buying their food. This leaves animal welfare as one of the less important factors in relation to choice in the Italian egg market.

These findings are in line with the Norwegian survey of Kvakkestad et al. (2011). In this survey, the researchers asked respondents to state which attributes are important when choosing eggs. Figure 4 provides an overview.

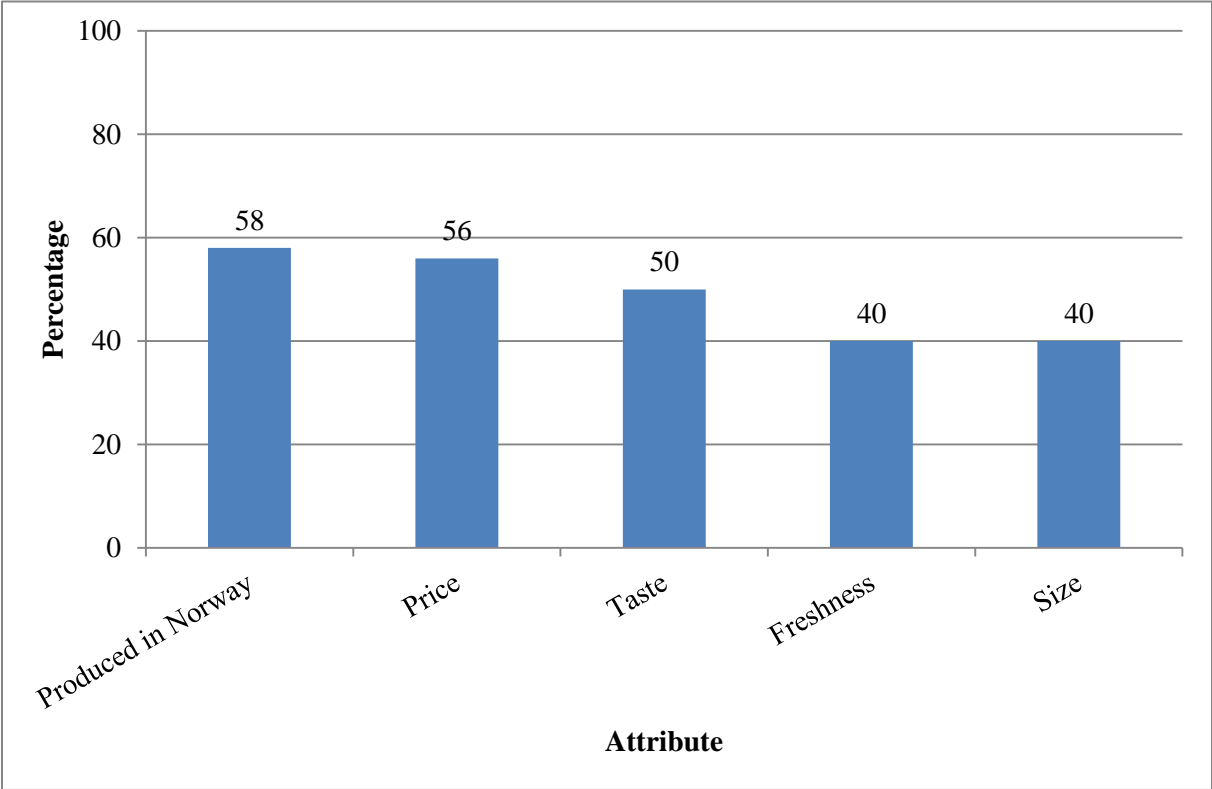


Figure 4: Norwegian consumers' rating of attributes when buying eggs.

Source: Kvakkestad et al. (2011), Table 6.

In addition to the attributes mentioned in the figure, the respondents also rank attributes such as free range and outdoor access, free range only, animal welfare, organic fodder, produced by an agricultural cooperative, short travelled, locally distinctive, long shelf life, color of the eggshell and extra-yellow yolk. All of these attributes had low rankings. In paper 2, veal is the product studied, and “Norwegian produce” is found to be the most important attribute.

Norwood and Lusk (2011) conducted an experimental auction with 100 consumers in three US cities. The mean bid for 12 battery eggs was \$0.92, while respondents were willing to pay \$1.53 for the same number of indoor free-range eggs. For eggs coming from indoor free-range aviary systems, the mean WTP was \$1.87. Consumers were willing to pay the most for organic eggs, with a mean WTP of \$2.23. These WTP numbers confirm that US consumers rank animal welfare according to the rearing systems used.

A meta-analysis of nine WTP studies of egg-laying hens conducted by Lagerkvist and Hess (2011) produced three main findings: increased income means increased WTP; older people have a lower WTP than younger people; and, on average, consumers in France and Germany have a higher WTP than consumers in the other tested countries.

In their preferred model, Lagerkvist and Hess (2011) find the following effects of changed welfare on WTP for egg-laying hens:

- More hours of indoor light reduces WTP estimates.
- Removal of isolation increases WTP estimates.
- Improvement of relevant amenities increases WTP estimates.
- Changes in regulation or labeling of animal welfare reduce WTP estimates.

The meta-study says that legislation and labeling can result in a loss to consumers, because sellers need to segregate products, which is a cost partly borne by consumers. They also note that simply knowing that other consumers purchase food produced in low-animal-welfare environments can create a loss for some consumers.

5.3 Summary of the literature

The idea of this overview is to provide a glimpse of the large body of literature on consumers' interest in animal welfare. Much is already known, but there are knowledge gaps when it

comes to products tested, methodologies used and sales channels. This thesis fills some of these gaps, at least partly.

6. Summary of papers

6.1 Paper 1: Eliciting consumer preferences for credence attributes in an haute cuisine restaurant

In this paper (coauthored with Frode Alfnes), we conduct a natural field experiment in an haute cuisine restaurant in Oslo. The purpose of this study is twofold: the main objective is to show how it is possible to conduct a natural field experiment in an haute cuisine restaurant without interrupting the daily running of the restaurant. Hereafter, we present the pros and cons of using such a methodology. Our second objective is to explore how labeling and pricing affects customers' choices in the restaurant.

By manipulating how the restaurant menu presents an organic veal course, we seek to understand whether the text describing the course and the price of the course affect sales of the course being studied.

We present five arguments for conducting experiments in a restaurant. First, we test organic veal as a food product. Organic veal has very limited sales in Norway, and what little sales there are occur in restaurants. Hence, a restaurant is the proper sales channel for testing organic veal's sales potential. Second, we test consumers' preferences for organic food and animal welfare while avoiding hypothetical choice situations, i.e. choices that a consumer would never actually face.

Animal welfare is an ethical issue. For such consumer issues, there is a "gold standard" of opinions, and there is a risk that consumers will pretend to meet this standard. In a restaurant, consumers would not expect researchers to monitor them as in a laboratory. When monitoring occurs without consumers' knowledge, we minimize the risk of deviation

from true preferences, i.e. social desirability bias. This is the third reason for conducting the experiment in a restaurant.

The fourth reason relates to context. Other studies show that people make different choices in different settings. Context seems to matter, so it is important to test its effect. In this paper, our context is the restaurant.

Finally, the restaurant sector, and food consumed away from home in general, is an important sales channel for food, especially given that 24 percent of Norwegians' spending on food and drinks is related to restaurant services (Statistics Norway 2013). Clearly, this sales channel has not received as much attention in the literature as it should have, based on its sales value.

There are four main disadvantages of conducting experiments in restaurants. First, a restaurant is not a place where people expect to be tested. Instead, they expect to have a nice dining experience, meaning that researchers must avoid disturbing them, but it is difficult to obtain useful background information from the diners without doing so.

Second, the field is not a "sterile" environment like a laboratory. In the field, the researcher cannot control everything, so there is a huge risk of a lot of "noise" in the data, making interpretation difficult.

Third, in a natural field experiment there is no opportunity to screen, match or randomize participants into the different treatments without revealing that an experiment is in progress. Thus, there is self-selection, meaning that the sample is not representative of the population as a whole.

Finally, restaurant guests make only one choice. They select one main course, in contrast to regular choice experiments (CEs) where respondents make a sequence of choices, ensuring a much richer dataset than can be obtained in a restaurant. Each participant provides the researcher with much more data in a standard CE than in a restaurant experiment. This

fact makes it more difficult to obtain significant results in a restaurant experiment compared with regular CEs.

In the experiment, we described the course as organic, or as animal-welfare friendly, or as both. In addition, we had three price levels for the course: one below the average price for the main courses, one at the average price, and one above the average price. We only manipulated the veal course, and kept everything else constant during the period of the experiment.

The results show that in general, the text manipulations had a small effect on sales of the veal course. The sales data show that describing the veal course as animal friendly significantly increased sales, but describing it as organic had no effect. However, changing the price had a more distinct effect. Interestingly, both a high price and a low price, compared with the average price, reduced sales. Low sales when the price is high are in line with standard economic theory. What is strange is that sales do not increase when the price is low. Most likely, there is a signaling effect here. Dining in a restaurant occurs mainly with companions. In such situations, choosing the cheapest course on the menu might send bad signals to those companions. Therefore, the demand curve for the veal course seems to have a peak, i.e. there is a price that maximizes sales. However, lower sales when price is low are not as strong as the regular effect of higher prices on sales.

We conclude by stating that restaurant experiments are a useful tool when studying consumer preferences toward food, but they have advantages and disadvantages. We believe that there are two reasons why our manipulations did not have significant effects. First, purchases of organic food by Norwegians are low in general, and so it is to be expected that this should not be substantially different in a restaurant compared with a grocery store. Second, animal welfare, like any ethical issue, is not something consumers like to think of

while enjoying a nice dining experience. A restaurant is not the place to start trying to make the world a better place, it would seem.

6.2 Paper 2: Country-of-origin preferences for organic food

This paper uses an online CE to test for credence preferences when buying veal. The purpose of this paper is to study whether the European organic labeling scheme can affect the strong preferences for domestic food found in most previous studies.

The background for the study is the European Union's goal to increase trade in agricultural products between the member countries. From many studies such as those of Ehmke et al. (2008) and Tonsor et al. (2013), it is well known that consumers prefer food from their home country. In order to limit the country-of-origin (COO) effect, a common quality label may be beneficial. If consumers believe the quality is the same, regardless of the place of origin, they have no reason to prefer domestic food other than a desire to support local producers. This is one reason for having a common organic label within the EU. Using the US organic label, Onozaka and Thilmany McFadden (2011) test the effect of an organic label on imported food among US consumers. However, in the US, there is no common international organic label. The only region with a common organic label is Europe.

The sample consists of Norwegians eating meat and being the main purchasers of groceries in their household. In total, 953 respondents participated in both a survey about food selection in general and a CE involving veal. In the CE, respondents ranked three pictures of veal packages. The packages were similar, and only differed in terms of price, animal welfare labeling and COO. Prices ranged from NOK 45 to NOK 85 for a 400-gram package. Animal welfare had three levels, conventional (no label), Friland and organic. The Friland label is quite unfamiliar to Norwegians, but was used because it mimics the organic label in all respects other than the use of organic fodder. That is, Friland is a pure animal welfare label. Consumers have many reasons for buying organic food, including those relating to health and

environmental issues. Therefore, the study had to use a label that was solely concerned with animal welfare in order to study consumers' interest in animal welfare. In relation to COO, The labels had text stating that the veal was a product of either Norway, or Denmark, or Poland.

From the results of our survey, taste, freshness and smell were the most important choice attributes when buying food in general. COO and animal welfare were not important. Those respondents who stated that they purchase organic food regularly were asked to rank organic food compared with conventional food. For them, low levels of pesticides/medicine rests and animal welfare were the attributes ranked highest.

The CE had a clear finding of COO being the most important attribute. For the full sample, all foreign veal packages received negative WTP estimates. That is, Norwegians seem to prefer Norwegian meat even when the foreign alternatives are organic. The WTP for Polish meat was particularly low.

For the respondents who buy organic food regularly, the picture was a little more balanced. They had the highest WTP for Norwegian organic and Friland meat, but unlike the full sample, they were willing to pay for Danish meat (both organic and Friland), but not for Polish meat of any kind.

The paper finds no reason to believe that a common international quality label can overcome the preference for domestic food found in most previous studies. COO is the most important credence attribute, so a quality label cannot overcome this effect. Quality labels are not enough to increase international trade in food. What consumers seek is trust in the country itself because they link evaluations of a country with food from that country. In addition, international trade in food has huge potential, but most likely between countries that are close to each other both culturally and geographically. Only then can the necessary trust be developed.

6.3 Paper 3: Animal welfare in stores: natural field experiments with positive and negative labeling

This paper (coauthored with Frode Alfnes and Svein Ole Borgen) examines whether it matters how producers transmit information about animal welfare to consumers. Specifically, we test whether positive labeling of high levels of animal welfare or negative labeling of low levels of animal welfare has the greatest influence on sales.

We use eggs in store for our study, because animal welfare is one of the choice attributes in this food segment. In Norway, it is possible to choose between three types of egg production systems: battery systems (hens living in cages), barn systems (hens moving about freely indoors) and organic systems (hens moving about freely both indoors and outdoors). The paper presents three natural field experiments. The first and second of these experiments were conducted within all stores of Norway's second-largest grocery chain. The third experiment was conducted in only two stores within the chain.

The first experiment involved the launch of an “unpleasant” carton for eggs from battery systems. This carton was black and white, and there was no doubt that the eggs came from a battery system. The carton replaced a neutral carton with no information about the production method. We believe that this new carton was a form of negative labeling. Here, as in the following experiments, we looked at how sales changed after the introduction of the new carton.

The second experiment looked at the effect on egg sales when battery eggs were removed. Consumers who were not concerned about animal welfare had to buy eggs from higher-animal-welfare production systems. The results show that consumers of battery eggs switched to barn eggs, while the organic market share remained stable.

The third experiment used positive information about animal welfare by promoting organic eggs. This was done by either placing a sticker on the organic-egg cartons, or putting

up a poster on the egg shelf, or a combination of the two. The sticker looked like an authentic label for animal welfare (no such label exists in Norway, so we had to design our own), and poster also looked authentic, urging shoppers to buy organic eggs to promote the welfare of hens. An organization for animal protection was the apparent sender of this message. The label and the poster are examples of positive labeling.

As with other studies of positive versus negative labeling, our results showed that negative labels influence consumers much more than positive labels. Introducing the unpleasant battery-egg carton (experiment 1) immediately reduced sales of battery eggs by 14 percentage points. In the third experiment, the positive information we provided for the organic eggs had no effect. According to our econometric model, price is the only variable that significantly influences sales. During the two years for which we have sales data, organic eggs maintained a more or less constant market share.

We believe that animal welfare has a huge potential impact on the egg market, but how producers transmit information about the issue to consumers is important. Traditionally, positive information has been used, but such information does not stand out among the myriad in-store marketing, which is probably why we did not see any effect from our positive labeling. Conversely, consumers would have found the negative labeling unusual, and so when it was introduced, they would have noticed it.

In addition, our findings can explain the mismatch found in other studies between action and attitude toward animal welfare. Many surveys find that consumers are very concerned about animal welfare, and yet welfare-friendly food has a very low market share.

Some researchers claim that consumers deliberately lie about their opinions and shopping habits in surveys because they like to appear to be “politically correct.” Our findings question this belief. We believe that the lack of overlap between action and attitude is simply due to the way in which consumers gain information. Instead of promoting the food produced

using high-animal-welfare methods, it is better to highlight the food produced using low-animal-welfare methods. When this is done, consumers' food selections in the store are more in line with what they say they do in surveys.

6.4 Paper 4: Attitude versus action for farm animal welfare: what can we learn from natural field experiments?

The purpose of this paper (coauthored with Frode Alfnes and Svein Ole Borgen) is to discuss how natural field experiments can contribute to more valid and reliable studies of consumer behavior. We use organic eggs as a case study.

The market share for organic eggs, those produced using the highest level of animal welfare, is very low in Norway. However, a previous survey by Kjærnes and Lavik (2008) finds that the hens' welfare is among the three most important factors when Norwegians choose their eggs. Thus, there seems to be a mismatch between action (actual sales) and attitude (responses in surveys).

This mismatch is common in relation to many ethical consumer issues, such as child labor and environmental concerns. A possible reason for this is the different roles people have. In stores, people act as consumers, thinking about what is best for them at point of purchase. Low price and taste are important factors in a purchase setting. In a survey, on the other hand, people act as citizens. Therefore, consumers will not necessarily live up to their stated values or preferences. One reason for this could be that they expect someone else to deal with the problem, not them as consumers. This is especially true in Norway, where the inhabitants are used to the government taking care of them in many ways, possibly including securing high levels of animal welfare.

We conducted an in-store natural field experiment using eggs in order to test whether the motivation to buy food labeled as animal-welfare friendly is present in consumers' minds. We wanted to see to what extent the issue of animal welfare would cause a spontaneous

action, just as surveys test for spontaneous attitudes. In order to look for a spontaneous action, we could not use in-store marketing involving a huge campaign, as is usual for marketers, because that would have enabled consumers to reflect on animal welfare issues prior to a visit to the store. In marketing terminology, we used incidental exposure, i.e. marketing that does not force consumers to make up their mind about the topic. In our experiment, we used “just another label” as our marketing technique. In both surveys and marketing in general, forced exposure is often used. This forces consumers to make an active decision, because they are bombarded with impressions. Such a technique could possibly produce a larger sales increase but would not trigger immediate reflection at the point of purchase, because at least some consumers would most likely have the message from the campaign in back of their mind, which we intend to measure. A campaign message would presumably influence consumers’ decisions; therefore, we could not launch it. For those consumers concerned with animal welfare, this treatment might act as a reminder of the benefits of buying animal-welfare-friendly products.

We discuss the validity and reliability of natural field experiments, using our egg experiment as a case study. We conclude that traditional methods in consumer research, such as surveys and focus groups, fail to mimic a real-life setting. In our experimental setting, consumers do not know that they are being monitored, so they act “naturally.” Therefore, our experiment is much more in line with actual market observations than a survey about attitudes to animal welfare. These experiments measure actions, not attitudes as measured by survey and focus groups. Hence, while natural field experiments cannot replace these traditional methods, they can complement them.

7. Limitations of the thesis

There is no such thing as a perfect study. All studies have limitations. In this thesis, the major limitation is the number of case studies. There are only two food products studied, so more case studies would strengthen the analysis.

Another limitation is the experimental design. Papers 1, 3 and 4 present data from field experiments. In order to conduct such experiments, the researcher needs support from someone outside academia. In this thesis, the manager of a restaurant and the managers of two grocery stores were the people setting our limits. The research could not be conducted exactly the way we wanted, and therefore the data are not as good as they could be. For example, we would have preferred to conduct more experiments. In evaluating the research, one should weigh this limitation against the benefits of conducting field experiments.

In addition to there being only two case studies, in at least one of the studies there is a problem with a limited number of observations. In the restaurant study, the limited amount of veal meant that only a small number of servings were available. This is probably the main reason why there are few significant findings in that paper. However, we would expect to obtain significant findings if the sample were larger.

All data collection takes place in Norway, which is the reference frame for the analysis. Norway is a special country when it comes to agriculture, with its strict import regulations for food and a policy based upon rural settlement and self-sufficiency. Therefore, our results may not be generalizable to other European countries, where free trade of agricultural products is more prevalent.

Several of the papers have weak or few significant findings. Why is that? Is the research design weak, or is it just that there is no linkage between the variables in the study? In this thesis, we investigate consumers' preferences for animal welfare. There are several such studies, but replicating the existing ones would probably provide little new knowledge.

Therefore, this thesis uses new methods in the field of consumer studies of animal welfare to test how robust the existing findings are. Overall, the interest in animal welfare found by the methods involving stated preferences methods does not translate into sales in our experiments. We find that only when low-animal-welfare food products are clearly stated as such do consumers switch to alternatives involving higher levels of animal welfare.

There are no previous field studies on this topic, thus it is not unexpected that these pioneering studies have several shortcomings. In future studies, we would seek to better engage with our nonacademic partners in order to improve data quality.

8. Generalization

This thesis examines consumers' preferences toward animal welfare, which is an ethical consumer issue. An important question to ask is whether the results also hold for other ethical consumer issues, such as child labor and environmentally friendly production.

This thesis tests consumer preferences for a credence attribute. As explained in Section 1.1, a consumer can never verify a credence attribute as being true. For example, it is impossible to distinguish a conventional egg from an organic one when the wrapping is gone. This is true for all consumer products that claim to be ethically produced. In this sense, the results here are transferable, not only to other kinds of ethically produced food, but also to clothes that label themselves as “no child labor.”

The ethical attribute is an example of a public good. If a cow experiences increased welfare, that is to the benefit of all people, not only those customers buying beef from the now “happy” cow. It is not possible to canalize the increased welfare to only those consumers buying the food. It is standard economic theory that we get low production of public goods in an unregulated market because of free riding. However, if consumers see a positive linkage between animal welfare and food quality, there is a private benefit from eating animal-

welfare-friendly food. It is the consumers who eat the food who experience the higher quality of the food, not the consumers who do not eat it. Then an individual consumer gains some of the collective aspect of animal welfare. At least to some extent, it should be possible to get close to the socially optimal level of animal welfare in society. Such a result is in line with the Coase theorem (Coase 1960), where property rights are expected to solve problems that arise from external effects. This reasoning should also hold for other ethical issues.

Hence, the results in this thesis are generalizable to other ethical consumer issues, not only for food, but also for textiles and apparel, where the same problems are present.

9. Contribution

The thesis makes two main contributions, one within the field of economic methodology and the other in terms of the understanding of consumer preferences in relation to animal welfare. In consumer research, it is important to map consumers' preferences regarding different issues. Studies of preferences in relation to animal welfare have increased in recent years, but the number of experiments is limited. Most likely, the studies presented in this thesis are the first field experiments conducted in this area. Certainly, no natural field experiments have previously been conducted. This is a flaw, because for ethical issues the field context can add realism, solving the mismatch between action and attitude that we often observe. In addition, we can acquire new insights in the field that we cannot obtain in the laboratory. The sales effect of negative information is an example of this.

The thesis has also contributed to economic methodology. These days, a CE is the typical way to measure consumers' preferences for goods not traded in a market. Carlsson et al. (2005b) discuss the possibility of combining CEs with revealed data as one of the main advantages of CEs. Unfortunately, there are not many studies along these lines. While my thesis does not combine different types of data in the same manner, as Carlsson et al. would

like to see, it does use the experimental design developed for CEs in the field. This is at least a small step in the right direction. By combining CEs with natural field experiments, one avoids the most important drawback of CEs, namely the hypothetical setting. In fact, the combination of techniques brings together the best of both worlds. A CE mimics how consumers derive utility from goods, allows exchangeability between attributes, mimics an actual purchase situation and often finds results in line with actual market behavior (op cit.). Natural field experiments avoid the hypothetical bias and other biases that can affect the results. A combination of CEs and natural field experiments is a methodology that is highly valid and reliable. This combination has several benefits, which should make it a popular method for conducting consumer research in the future. It is possible that one could use this combination in environmental and health-care studies, for example, or in any other field where biases appear. This thesis shows how this is possible.

References

- Aertsens, J.; W. Verbeke; K. Mondelaers and G. Van Huylenbroeck (2009): "Personal determinants of organic food consumption: a review." *British Food Journal*, vol. 111, no. 10, pp. 1140-1167.
- Alexander, V. D.; H. Thomas; A. Cronin; J. Fielding and J. Moran-Ellis (2008): "Mixed Methods." In N. Gilbert (Ed.) *Researching social life. Third edition*. London: Sage. (Pp. 125-144).
- Alfnes, F. (2004): "Stated Preferences for Imported and Hormone-Treated Beef: Application of a Mixed Logit Model." *European Review of Agricultural Economics*, vol. 31, no. 1, pp. 488-500.
- Andersen, L. M. (2011): "Animal Welfare and Eggs. - Cheap Talk or Money on the Counter?" *Journal of Agricultural Economics*, vol. 62, iss. 3, pp. 565-584.
- Animalia (2013): *Kjøttets tilstand 2012*. [State of the meat.] URL: http://animalia.no/upload/Kj%C3%B8ttets_tilstand_2012/Kj%C3%B8ttets_tilstand_2012.pdf [Date of reading: 27.07.2013]
- Bennett, R. M.; J. Anderson and R. J. P. Blanley (2002): "Moral intensity and willingness to pay concerning farm animal welfare issues and the implications for agricultural policy." *Journal of Agricultural and Environmental Ethics*, vol. 15, iss. 2, pp. 187-202.
- Bernués, A.; A. Olaizola and K. Corcoran (2003): "Extrinsic attributes of red meat as indicators of quality in Europe: an application for market segmentation." *Food Quality and Preference*, vol. 14, iss. 4, pp. 265-276.
- Burrell, A. and G. Vrieze (2003): "Ethical motivation of Dutch egg consumers." *Tijdschrift voor socialwetenschappelijk onderzoek van de landbouw*, vol. 18, no. 1, pp. 30-42.
- Camerer, C. F. and G. Loewenstein (2004): "Behavioral Economics: Past, Present, Future." In C. F. Camerer, G. Loewenstein and M. Rabin (Eds.) *Advances in Behavioral Economics*. Princeton, NJ: Russell Sage Foundation Princeton University Press. (Pp. 3-52).
- Carlsson, F.; P. Frykblom and C. J. Lagerkvist (2007a): "Consumer willingness to pay for farm animal welfare: mobile abattoirs versus transportation to slaughter." *European Review of Agricultural Economics*, vol. 34, no. 3, pp. 321-344.
- Carlsson, F.; P. Frykblom and C. J. Lagerkvist (2007b): "Farm Animal Welfare-Testing for Market Failure." *Journal of Agricultural and Applied Economics*, vol. 39, no. 1, pp. 61-73.
- Carlsson, F.; P. Frykblom and C. J. Lagerkvist (2005a): "Using cheap talk as a test of validity in choice experiments." *Economics Letters*, vol. 89, iss. 2, pp. 147-152.
- Carlsson, F.; P. Frykblom and C. J. Lagerkvist (2005b): "Consumer Preferences for Food Product Quality Attributes from Swedish Agriculture." *AMBIO*, vol. 34, no. 4, pp. 366-370.
- Chilton, S. M.; D. Burgess and W. G. Hutchinson (2005): "The relative value of farm animal welfare." *Ecological Economics*, vol. 59, iss. 3, pp. 353-363.
- Cicia, G. and F. Colantuoni (2010): "Willingness to Pay for Traceable Meat Attributes: A Meta-Analysis." *International Journal on Food System Dynamics*, vol. 1, no. 3, pp. 252-263.
- Coase, R. (1960): "The Problem of Social Cost." *Journal of Law and Economics*, vol. 3, no. 1, pp. 1-44.
- Darby, M. R. and E. Karni (1973): "Free competition and the optimal amount of fraud." *Journal of Law and Economics*, vol. 16, no. 1, pp. 67-88.

- Dentoni, D.; G. Tonsor; R. Calantone and H. C. Peterson (2011): “‘Animal Welfare’ Practices along the Food Chain: How Does Negative and Positive Information Affect Consumers?” *Journal of Food Products Marketing*, vol. 17, iss. 2-3, pp. 279-302.
- Dickinson, D. L. and D. Bailey (2002): “Meat Traceability: Are U.S. Consumers Willing to Pay for It?” *Journal of Agricultural and Resource Economics*, vol. 27, no. 2, pp. 348-364.
- Eckert, C.; J. J. Louviere, and T. Islam (2012): “Seeing the forest despite the trees: brand effects on choice uncertainty.” *International Journal of Research in Marketing*, vol. 29, iss. 3, pp. 256–264.
- Ehmke, M. D.; J. L. Lusk and W. Tyner (2008): “Measuring the relative importance of preferences for country of origin in China, France, Niger, and the United States.” *Agricultural Economics*, vol. 38, iss. 3, pp. 277–285.
- Eurobarometer (2005): *Attitudes of consumers towards the welfare of farmed animals. Special Eurobarometer 229/Wave 63.2*. Brussels: European Commission. URL: http://ec.europa.eu/public_opinion/archives/ebs/ebs_229_en.pdf [Date of reading: 17.07.2012]
- European Commission (2012): *Table 4.15.1.2 Net production of beef/veal in the EU*. URL: http://ec.europa.eu/agriculture/statistics/agricultural/2012/pdf/d15-0-415_en.pdf Brussels: European Commission. [Date of reading: 03.08.2013]
- Flyvbjerg, B. (2006): “Five Misunderstandings About Case-Study Research.” *Quality Inquiry*, vol. 12, no. 2, pp. 219-245.
- Gerhardy, H. and M. R. Ness (1995): “Consumer preferences for eggs using conjoint analysis.” *World’s Poultry Science Journal*, vol. 51, iss. 02, pp. 203-214.
- Gracia, A.; M. L. Loureiro and R. M. Nayga Jr. (2011): “Are Valuations from Nonhypothetical Choice Experiments Different from Those of Experimental Auctions?” *American Journal of Agricultural Economics*, vol. 92, no. 5, pp. 1358-1373.
- Harrison, G. W. and J. A. List (2004): “Field Experiments.” *Journal of Economic Literature*, vol. 42, no. 4, pp. 1009-1055.
- Heerwagen, L. R.; T. Christensen and P. Sandøe (2013): “The Prospect of Market-Driven Improvements in Animal Welfare: Lessons from the Case of Grass Milk in Denmark.” *Animals*, vol. 3, iss. 2, pp. 495-512.
- Hensher, D. A.; J. M. Rose and W. H. Greene (2005): *Applied Choice Analysis. A Primer*. Cambridge, MA: Cambridge University Press.
- Hunt, L. M.; D. Gonder and W. Haider (2010): “Hearing Voices from the Silent Majority: A Comparison of Preferred Fish Stocking Outcomes for Lake Huron by Anglers from Representative and Convenience Samples.” *Human Dimensions of Wildlife*, vol. 15, iss. 1, pp. 27-44.
- Kahneman, D. and A. Tversky (1979): “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica*, vol. 47, no. 2, pp. 263-292.
- Kjærnes, U. and R. Lavik (2008): “Opinions on Animal Welfare and Food Consumption in Seven European Countries.” In U. Kjærnes, B. B. Block, E. Roe and J. Roex (Eds.) *Consumption, Distribution and Production of Farm Animal Welfare. Opinions and Practices within the Supply Chain*. Welfare Quality Reports no. 7. Cardiff: School of City and Regional Planning, Cardiff University. (Pp. 3–117). URL: <http://www.welfarequality.net/everyone/42462/7/0/22> [Date of reading: 08.08.2013]
- Kvakkestad, V.; K. Refsgaard and H. Berglann (2011): *Citizen and consumer attitudes to food and food production in Norway*. Discussion paper no. 2011-2. Oslo: Norwegian Agricultural Economics Research Institute. URL:

- http://www.nilf.no/publikasjoner/Discussion_Papers/2011/dp-2011-02.pdf [Date of reading: 16.01.2013]
- Lagerkvist, C. J. and S. Hess (2011): "A meta-analysis of consumer willingness to pay for farm animal welfare." *European Review of Agricultural Economics*, vol. 38, no. 1, pp. 55-78.
- List, J. A. (2011): "Why Economists Should Conduct Field Experiments and 14 Tips for Pulling One Off." *Journal of Economic Perspectives*, vol. 25, no. 3, pp. 3-15.
- List, J. A. (2007): "Field Experiments: A Bridge between Lab and Naturally Occurring Data." *The B.E. Journal of Economic Analysis & Policy*, vol. 5, iss. 2, pp. 1-47.
- Lobb, A. E. and M. Mazzocchi (2007): "Domestically produced food: Consumer preferences of origin, safety and the issue of trust." *Food Economics*, vol. 4, iss. 1, pp. 3-12,
- Louviere, J. J., T. N. Flynn and R. T. Carson (2010): "Discrete Choice Experiments Are Not Conjoint Analysis." *Journal of Choice Modelling*, vol. 3, no. 3, pp. 57-72.
- Lovdata (2009): Dyrevelferdsloven. [The animal welfare act.] URL: <http://lovdata.no/dokument/NL/lov/2009-06-19-97?q=dyrevelferdsloven> [Date of reading: 07.05.2014]
- Lusk, J. L. and J. A. Fox (2003): "Value Elicitation in Retail and Laboratory Environments." *Economic Letters*, vol. 79, iss.1, pp. 27-34
- Lusk, J. L. and J. A. Fox (2002): "Consumer Demand for Mandatory Labeling of Beef from Cattle Administered Growth Hormones or Fed Genetically Modified Corn." *Journal of Agricultural and Applied Economics*, vol. 34, no. 1, pp. 27-38.
- McFadden, D. (1976): "The Revealed Preferences of a Government Bureaucracy: Empirical Evidence." *The Bell Journal of Economics and Management Science*, vol. 7, no. 1, pp. 55-72.
- McFadden, D. (1968): *The Revealed Preferences of a Public Bureaucracy*. Berkeley, CA: Department of Economics, University of Berkeley.
- Ministry of Agriculture and Food (2011): *Matproduksjon for framtiden*. [Food production for the future.] Meld. St. 9 (2011-2012). Oslo: Ministry of Agriculture and Food. URL: <http://www.regjeringen.no/nb/dep/lmd/dok/regpubl/stmeld/2011-2012/meld-st-9-20112012/2.html?id=664982> [Date of reading: 29.07.2013]
- Napolitano, F.; G. Caporale; A. Carlucci and E. Monteleone (2007): "Effect of information about animal welfare and product nutritional properties on acceptability of meat from Podolian cattle." *Food Quality and Preference*, vol. 18, iss. 2, pp. 305-312.
- Nobelprize.org (2013a): *Daniel Kahneman – Facts*. Stockholm: Nobel Media. URL: http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2002/kahneman-facts.html [Date of reading: 02.05.2014]
- Nobelprize.org (2013b): *Vernon L. Smith – Facts*. Stockholm: Nobel Media. URL: http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2002/smith-facts.html [Date of reading: 10.04.2014]
- Norwegian Agricultural Authority (2013): *Produksjon og omsetning av økologiske landbruksvarer. Rapport for 2012*. [Production and sales of organic agricultural products. Report for 2012.] Rapport-nr.: 12/2013. Oslo: Norwegian Agricultural Authority. URL: <https://www.slf.dep.no/no/miljo-og-okologisk/okologisk-landbruk/om-okologisk-landbruk/publikasjoner/attachment/28778?ts=13d8c2bb720&download=true> [Date of reading: 28.07.2013]
- Norwegian Scientific Committee for Food Safety (2014): *Comparison of organic and conventional food and food production. Part II: Animal health and welfare in Norway*. Doc. no. 11-007-02. Oslo: Norwegian Scientific Committee for Food

- Safety. URL: <http://www.vkm.no/dav/9cdfb4c26c.pdf> [Date of reading: 09.05.2014]
- Norwood, F. B. and J. L. Lusk (2011): "A calibrated auction-conjoint valuation method: Valuing pork and eggs produced under differing animal welfare conditions." *Journal of Environmental Economics and Management*, vol. 62, iss. 1, pp. 80-94
- Onozaka, Y. and D. Thilmany McFadden (2011): "Does local labeling complement or compete with other sustainable labels? A conjoint analysis of direct and joint values for fresh produce claims." *American Journal of Agricultural Economics*, vol. 93, no. 3, pp. 693–706.
- Patton, M. (1990): *Qualitative evaluation and research methods*. Beverly Hills, CA: Sage.
- Pettersen, I.; J. Kjuus and R. Lavik (2011): "Stor prisforskjell - med naturlige, politiske og strukturelle forklaringer." [Large price difference – with natural, political and structural explanations.] In I. Pettersen and T. S. Gabrielsen (Eds.) *Dagligvarehandel og mat 2011. Perspektiver på verdikjedene for matvarer*. [Grocery trade and food 2011. Perspectives on value chains for food stuff.] Oslo: Norwegian Agricultural Economics Research Institute. (Pp. 87-112). URL: http://www.nilf.no/publikasjoner/Andre_publicasjoner/DagligvarehandelOgMat/2011/dagligvarehandel_og_mat_2011_endelig.pdf [Date of reading: 08.05.2014]
- Rotevatn, S. (2014): *Skriftlig spørsmål fra Sveinung Rotevatn (V) til landbruks- og matministeren*. [Written question from Sveinung Rotevatn (V) to the Minister of Agriculture and Food.] URL: <https://stortinget.no/no/Saker-og-publikasjoner/Sporsmal/Skriftlige-sporsmal-og-svar/Skriftlig-sporsmal/?qid=59297> [Date of reading: 09.05.2014]
- Rygh, A. (2007): *Ethical information and working conditions. A comparison of Fairtrade labelling and codes of conduct in coffee, apparel and personal computers*. Professional Report no. 8-2007. Oslo: National Institute for Consumer Research. URL: http://sifo.no/files/file72607_fagrappport2007-8.pdf [Date of reading: 09.04.2014]
- Schulz, L. L. and G. T. Tonsor (2010): "Cow-Calf Producer Preferences Regarding Individual Animal Traceability." *Journal of Agricultural and Applied Economics*, vol. 42, no. 4, pp. 659-677.
- Schumacher, T.; T. C. Schroeder and G. T. Tonsor (2012): "Willingness-to-Pay for Calf Health Programs and Certification Agents." *Journal of Agricultural and Applied Economics*, vol. 44, no. 2, pp. 191-202.
- Smith, V. L. (1991): *Papers in Experimental Economics*. Cambridge, MA: Cambridge University Press.
- Smith, V. L. (1962): "An Experimental Study of Competitive Market Behavior." *Journal of Political Economy*, vol. 70, no. 2, pp. 111–137.
- Statistics Norway (2013): *Table 10235: Expenditure per household per year, by commodity and service group*. Oslo: Statistics Norway.
- Teyssier, S.; F. Etilé and P. Combris (2012): *Social- and Self-Image Concerns in Fair-Trade Consumption: Evidence from Experimental Auctions for Chocolate*. Working paper no. 2012-33. Paris: Paris School of Economics. URL: <http://halshs.archives-ouvertes.fr/docs/00/72/25/92/PDF/wp201233.pdf> [Date of reading: 09.04.2014]
- Tonsor, G. T.; T. C. Schroeder and J. L. Lusk (2013): "Consumer valuation of alternative meat origin labels." *Journal of Agricultural Economics*, vol. 64, iss. 3, pp. 676–692.
- Tonsor, G. T. and N. J. Olynk (2011): "Impacts of Animal Well-Being and Welfare Media on Meat Demand." *Journal of Agricultural Economics*, vol. 62, iss. 1, pp. 59–72.
- Vangen, O.; N. H. Sæter and B. K. B. Nordtømme (2007): *Beitende husdyr i Norge*. [Grazing farm animals in Norway.] Oslo: Tun.

- Van Loo, E. J.; V. Caputo; R. M. Nayga Jr.; M. Canavari and S. C. Ricke (2012): "Organic Meat Marketing." In S. C. Ricke, E. J. Van Loo, M. G. Johnson and C. O'Bryan (Eds.) *Organic Meat Production and Processing*. Ames, IO: Wiley-Blackwell. (Pp. 67-85).
- Vecchio, R. and A. Annunziata (2011): "Italian consumer awareness of layer hens' welfare standards: a cluster analysis." *International Journal of Consumer Studies*, vol. 36, iss. 6, pp. 647-655.

Part II: Papers

Paper 1

Eliciting Consumer Preferences for Credence Attributes in an Haute Cuisine Restaurant

By Alexander Schjøll and Frode Alfnæs

Abstract

The choices faced by restaurant customers are very similar to the choices marketers provide in choice experiments. In this paper, we illustrate and discuss restaurant choice experiments in an haute cuisine restaurant in Oslo where the customers do not know they are part of an experiment. In our illustration, we explore how menu descriptions affect customer choices of an organic veal main dish. Holding everything else constant, we saw how sales fluctuated when we changed the price and the menu description of the dish. The sales data show that describing the veal dish as animal friendly increased sales of the item, but describing it as organic had no effect. When the veal dish had same price as the other main dishes on the menu, sales were highest. We conclude by giving some tips on how to conduct experiments in an operating restaurant.

Key words

Restaurant choice experiment, consumer preferences, animal welfare, organic meat

Introduction

The restaurant sector is a large and growing food market, but very little research on consumer preferences and choice in haute cuisine restaurants have so far been carried out. In this paper, we present a natural field experiment conducted at an haute cuisine restaurant in Norway. By manipulating the menu text and price for an organic veal main course, we investigate how information about production methods and the price affects customers' choices in the restaurant.

According to the National Restaurant Association in the U.S., almost one in 10 U.S. workers work in a restaurant, 48% of the food dollar is used in restaurants, and the U.S. restaurant industry's sales exceed \$600 billion in 2011 (National Restaurant Association, 2012; Steward, 2011). Europeans eat less food away from home than the Americans, and in Norway 24% of the money for food and drinks is used in restaurants, cafés, bars, and cantinas (Statistics Norway, 2013). European households spend 10% of the household budget in different type food establishments, which is less than what they spend on food and drinks at home (Eurostat, 2014).

Upper class restaurants often use niche products that are hard to find or highly priced in ordinary grocery stores and therefore not commonly bought by consumers for home

consumption. In Norway, around 1% of the food eaten was organic in 2013 (Norwegian Agricultural Authority, 2014), and meat is one of the food categories in which the organic share is lowest. Norwegian restaurants can exploit the scarcity of organic meat in grocery stores by using a product like organic veal to provide customers something they would unlikely eat at home.

The objective of this paper is to present how one can conduct a choice based natural field experiment in an haute cuisine restaurant without interrupting the daily running of the restaurant. We discuss the strength and weakness of the methodology, and the challenges we faced when working in a restaurant where customers pay with their own money. To make it easier for others to conduct similar choice experiments, we also discuss how we handled the practical research challenges. In our illustration, we explore how menu descriptions affect customer choices of an organic veal main course in an haute cuisine restaurant.

Food experiments in the field

In marketing research, a continual and heavily debated issue is the relationship between action and attitude (e.g. Gabor, Granger, & Sowte, 1970; Lusk & Fox, 2002). The choice of methodology is important in this discussion. A survey may conclude that a product is a certain success, but when the product moves on to the market, sales are lackluster. This may be because surveys measure only attitude, not action. Experiments with real economic consequences, which seek to imitate real purchase situations, may reduce this mismatch. However, in most experiments respondents are aware that they are taking part in an experiment. This may lead them to behave differently compared with their behavior in other settings.

Our experiment used a restaurant that was open to the public with participants that were ordinary customers walking in from the street. The participants did not receive any information about the experiment and made selections from the menu, as they would have on

any other day that they might have visited the restaurant. The menu items, the descriptions of the menu items products, and the prices of the menu items were in line with those normally presented by the restaurant. In Harrison & List's (2004) typology, our experiment is then a *natural field experiment* with a field context in all factors.

A natural field experiment has as its advantage that the setting is as natural as it gets. No instructors telling subjects they are part of an experiment; and if the experiment is ideal, subjects will have no knowledge that they are taking part in an experiment and that their behavior will be analyzed. As a result, we expect to see no difference between behavior in a natural field experiment and in the real world, because the natural field experiment is, in fact, the real world.

The main disadvantage of conducting experiments in the field is the lack of control over many factors that can affect the results. In experimental auctions, for example, it is standard procedure that the participants are not allowed to talk to each other. A similar "sterile" environment is not possible in a natural field experiment, and many uncontrollable factors will likely affect the results. The lack of control and additional noise will lead to a reduction in internal validity, and will typically widen the confidence intervals of parameters, and thereby reduce the goodness of fit and predictive power of the models. In the economic literature, this lack of control has long been seen as a big hurdle to overcome, and few field experiments have therefore been conducted.

Nevertheless, in recent years, there has been a rapid increase in field experiments in many areas within economics. For a good overview, of the use of field experiments in other parts of the economic literature, see John List's webpage www.fieldexperiments.com.

Why conduct a restaurant experiment?

There are several reasons why we chose to conduct our study in an haute cuisine restaurant. These reasons are also likely to be equally relevant in many other studies of consumer

behavior. First, the product of interest in our project is organic veal, and this is a typical restaurant product in Norway. One of the largest Norwegian producers of organic veal informed us that the restaurant market is his main sales channel. This is consistent with our observations of meat in grocery stores in Oslo, where we rarely find organic veal.

Secondly, we wanted to elicit revealed preferences using a method in which it was in the participant's best interest to follow, and thereby reveal, his or her true preferences (Alfnes & Rickertsen, 2011). Revealed preference refers to real decisions in real markets, unlike preferences from surveys, which are about hypothetical decisions (Jaeger & Rose, 2008).

Ding, Grewal, & Liechty (2005) compare the predictive abilities of a hypothetical stated-preference and an incentive-aligned choice experiment in a Chinese dinner restaurant. In predicting the customers' choices, the incentive-aligned choice experiment outperformed the hypothetical choice experiment. The authors argue that the problem with the hypothetical choice experiment is too low price sensitivity, too much risk taking, too high willingness to test new things, and a tendency to suffer from social desirability bias. According to the authors, this is because participants are not equally serious about answering hypothetical questions as compared to real purchase decisions.

Third, knowledge of being observed, and that researchers will analyze their behavior, often make study participants behave differently than they would in a similar situation without monitoring. For credence attributes with ethical dimensions, such as fair trade products and animal welfare friendly food, we would be concerned about social desirability bias in any method where the participants know they are observed (Fisher, 1993; Norwood & Lusk, 2011). Social desirability bias means that respondents are more likely to make ethically correct choices in an observed setting than in daily life. Laboratory experiments with economic incentives can reduce, to some extent, social desirability bias by making it costly to choose the socially correct product (Norwood & Lusk, 2011). However, the economic

incentives are not likely to totally alleviate the effect of researchers observing the participants' choices (Levitt & List, 2007).

Fourth, previous studies show that context is very important when studying human behavior. In what context the decisions have been made, should therefore always be an issue when interpreting results. For example, Benkahla, Boutonnet, & Fort (2005) show that benefits of protected-designation-of-origin labeling varied greatly between sales channels. Because context matters and we wanted point-of-purchase realism, it was natural to test the product in a context identical to the actual context in which it would most likely be sold.

There are of course also disadvantages when doing a natural field experiment in a restaurant. As in other natural field experiments, lack of control of external factors are also the main concern in natural field experiments conducted in restaurants. A second disadvantage of a natural field experiments where the researchers do not seek out the participants, but the participants self-select into the experiment, is the inability to screen, match, or randomize participants into the different treatments without revealing that an experiment is in progress. These internal validity problems might also affect the external validity of the results.

The restaurant setting also has its own limitations. Foremost, people do not come to restaurants in order to be tested; they come to enjoy a dining experience. Disturbance of the customers must therefore be minimized not to destroy the dining experience and to elicit preferences without informing consumers of the experiment. Otherwise, it would be hard to generalize one's results to non-monitored restaurants. This fact makes it difficult to collect good background information about subjects' motivations for their choices and other segmentation variables.

With these pros and cons in mind, it is easy to see that field experiments in restaurants can be a valuable addition to the surveys and lab-experiments conducted to elicit consumer preference for food. Surveys and lab-experiments have their own pros and cons. They are

both relatively easy to conduct, but are likely to be poor on context and suffer from social desirability bias. For surveys, we also have the infamous hypothetical bias in valuation (Cummings, Harrison, & Rutström, 1995).

Literature Review of Restaurant Experiments

Most of the recent studies using some sort of experiment in restaurants focus on nutritional labeling (see e.g. Howlett, Burton, Bates, & Huggins, 2009; Roberto, Larsen, Agnew, Baik, & Brownell, 2010). Few have studied the effect of other types of information, such as origin, organic, animal welfare, sustainability, and fair trade. Furthermore, as far as we know, none of them is a natural field experiment in an haute cuisine restaurant. Most restaurant experiments have been conducted in restaurants associated with universities, and in many of the studies, the customers are fully aware that they are part of a study, even though they usually do not know what the purpose of the study is. Here we will focus on the few papers manipulating the menu by either using culinary words, words informing about credence attributes other than nutrition, or informing about best sellers.

Menu descriptions are important in themselves. McCall & Lynn (2008) show how more complex menu descriptions increase perceptions of quality, the chance of liking the menu item, and its price expectation. Whether they priced the menu item or not, did not affect these results. It therefore seems that menu texts have their sole effect regardless of the ingredients in the dish or its price.

Another way of altering the description without altering the dish, is to use words with positive food associations. Wansink, Painter, & van Ittersum (2001) study the influence of descriptive labels like "Grandma's zucchini cookies" or "succulent Italian seafood filet". Such labels have many positive effects for the restaurant, such as increased sales and improved evaluations of the restaurant.

Few studies have investigated the effect of inserting words informing about credence attributes into the menu text. One example is Alfnes & Sharma (2010), which study consumers' willingness to pay for local food in restaurants. The study was conducted at a university training restaurant, and they manipulated the price of a local food set menu relative

to the price of non-local food set menu. They found a non-linear pricing effect, and conclude that a price premium is important in strengthening the quality signal sent by the local origin.

A study by Cai, Yuyu, & Fang (2007) investigates social learning in restaurants. The authors define social learning as "any mechanism through which individuals learn from each other" (p. 864). They test social learning through experiments in several restaurants serving traditional Chinese food in Beijing. These medium scale restaurants have around 60 hot dishes on the menu, so selecting what to eat can be quite challenging. This is where social learning may play a role. In their first treatment, they provide some of the tables in the restaurants with a plaque displaying the five best selling dishes. The idea behind this treatment is to test for "observational learning", i.e. herd instinct. Their second treatment tests for a "saliency" effect. The second treatment also uses a plaque describing five dishes on some of the tables. This time three dishes were among the five most selling, and they pick two randomly from the menu. The two treatments made it possible to test whether popularity or saliency influences guests' choice of dishes. The results show that the first treatment with information of the five most selling dishes increased the chance for selecting one of these five dishes with 13-18%. The second treatment with three popular dishes and two picked at random on the plaque did not increase sales of these five dishes. The authors therefore conclude that observational learning has effect on sales, but there is no "saliency" effect. In other words, people's choices are influenced by the choice of other restaurant guests.

Method illustration

Description of the Restaurant Choice Experiment

We conducted our restaurant choice experiment in "34 Restaurant & Bar" located at the 34th floor of Radisson Blu Plaza Hotel in Oslo, Norway, from June 11 to 26, 2010. This landmark in Oslo is the tallest hotel in Northern Europe, the second tallest building, and largest hotel in Norway. The restaurant is located at the top of the hotel, offering a magnificent view of the city. The restaurant has a French kitchen with an international touch, has a price level that is

in line with other haute cuisine restaurants (without a listing in the Michelin Guide) in Oslo, and has received favorable reviews in newspapers. The restaurant can seat 62 people.

It was very important to the restaurant manager that we did not affect the restaurant experience of the customers. This had several consequences. One was that the manager did not allow us to interview or conduct a survey among the customers while they were seated in the restaurant. At the same time, knowing why the guests chose as they did would be relevant for our study. The restaurant compromised by allowing the waiters to hand out a business card with a web address to an online questionnaire and guests were asked to visit this web site. The waiter handed out this when the guests were paying, so it would not influence their food choices.

Unfortunately, very few people responded to the web survey after coming home, so we do not have reliable background information on customers, beyond the general description provided by the restaurant manager. According to the executive head chef, a majority is guests at the hotel, and around 50% are foreigners. On weekdays, business people constitute most of the clients, while on weekends; there is mix of locals, business, and tourists. Being aware of this, we made weekends one of the attributes in our design.

Table 1 shows an example of the menu pages for main courses. We held prices for all non-veal main courses on the menu at constant NOK 285 (€36)¹ during the experimental period. For the main course, guests could choose from fish, meat, lobster, and a vegetarian alternative, in addition to the veal. Varying the veal prices was a part of the experimental design.

¹ At the time of experiment, NOK 100 equaled € 12.75.

Table 1. Example of the menu pages with the main dishes.

Today's fish	
Arctic char with chorizo lentils and fig sauce	NOK 285
Today's meat	
Trio of veal from Grøndalen farm	NOK 285
Lobster natural	
½ lobster served on a bed of toast	NOK 285
Lobster gratinated	
½ lobster served with aioli on a bed of salad	NOK 285
Today's vegetarian	
Potato ravioli filled with mushrooms. Served with spinach and sweet pepper sauce	NOK 285

We created a fractional factorial design using the SAS *%MktEx* macro (Kuhfeld, 2010). The design is based on a choice between two alternatives, where alternative 1 is kept constant (the other main courses) and alternative 2 (the veal dish) is varied according to the design.² The design included four factors. Two 2-level credence attributes (dummies for organic and positive animal welfare), a 3-level price attribute (NOK 265, 285, and 310), and a 2-level factor used to control for possible customer differences between weekends and other days. The restaurant required that we only changed the menu every second day, starting on a Friday and ending on a Saturday two weeks later. The length of the experiment was determined by the amount of organic veal the restaurant had bought; two full calves. With the restaurant

² The design is not a typical choice design, as only one of the alternatives is altered, and therefore only one alternative is included in the design. If one has the possibility to alter several dishes on the menu, one can create an efficient choice design using for example the design software *Ngene* or the SAS macro *%choiceeff*.

closed on Sundays, this gave us seven two-day periods, with three being Friday and Saturday nights. The software SAS reports a D-efficiency of 94.34 for the design.³

Table 2 describes the four menu texts used in the experiment. We had to write these texts in a way that would seem natural for our restaurant. Therefore, the head chef at the restaurant took part in discussion of the wording, and he had clear views of what he thought would work and what he could accept on his menu. Primarily, only positive descriptions could be on the menu. Second, the wording had to be non-technical. Third, the entry had to be relatively short to be in line with the rest of the menu. This posed no problem for organic, which is a well-established concept among consumers and commonly used in restaurant menus. However, there is no standard way of referring to enhanced animal welfare in restaurants. We ended up using the phrase "from happy calves that have received much care and exercise". The veal dish in itself was the same regardless of the price and menu text.

Table 2. Menu texts used in the experiment for the organic and animal welfare attributes.

Organic	Animal welfare	Menu text
No	No	Trio of veal from Grøndalen farm.
Yes	No	Trio of organic veal from Grøndalen farm.
No	Yes	Trio of veal from Grøndalen farm from happy calves that have received much care and exercise.
Yes	Yes	Trio of organic veal from Grøndalen farm from happy calves that have received much care and exercise.

³ See Table A1 in the appendix for the design.

We did not want the wait staff to affect the choice of the customers, and therefore instructed them to refrain from saying anything more about the dishes than what was available on the menu. In order to ensure that the experiment was conducted as planned, we visited the restaurant on two random occasions as regular guests. We had not met the waiters before (only the head chef), so we were unfamiliar to the waiters on those evenings. On both evenings, the experiment went as planned.

Econometric Model

In contrast to most survey-based choice experiments, our participants in the restaurant choice experiment made only one choice. Our 462 participants provided us with only 462 choice observations, and among these, 180 (39%) were veal dishes. Because of the relatively low number of choice observations, and the lack of panel features, we employ the binary logit model to analyze the choice data. The dependent variable in our logit model is purchase of veal. The three veal attributes are independent variables.

With only seven profiles in the design and one choice per respondent, we have chosen to estimate models with as few parameters as possible. However, to test if the estimated effects of organic production and animal welfare are robust to different specifications of the price effect, we include three price specifications. Hence, we have estimated three binary logit models that only differ in the price specification. The utility functions of the three models are specified as:

$$(1) U_i = \beta_0 + \beta_1 AW_i + \beta_2 Org_i + \beta_3 P_i + \varepsilon_i$$

$$(2) U_i = \beta_0 + \beta_1 AW_i + \beta_2 Org_i + \beta_3 PW_i + \beta_3 PS_i + \varepsilon_i$$

$$(3) U_i = \beta_0 + \beta_1 AW_i + \beta_2 Org_i + \beta_3 PL_i + \beta_3 PH_i + \varepsilon_i$$

The U_i is guest i 's utility from choosing the veal alternative on the menu. The constant term β_0 captures the utility of the veal dish sold without any information about its characteristics. AW_i is a dummy indicating that the description of the veal dish included an

animal welfare statement. Org_i is a dummy indicating that the description of the veal dish included the word "organic". P_i is an ordinal variable indicating the price level. In Model 2, PW_i measures the price level on weekdays (Monday–Friday). PS_i measures the price level on weekends (Saturday). In Model 3, PH_i and PL_i are dummies indicating the high (NOK 310) and the low (NOK 265) price for the veal dish, respectively. The reference point for all three models, with utility normalized to zero, is the utility of choosing something other than veal from the menu.

The other β s are the respective parameters indicating how the variables affect the likelihood of choosing the veal over the other alternatives on the menu. Finally, ε_i is an independently and identically distributed extreme value error term. We estimated the model in Stata 11.

The rationale behind the three different price specifications are as follows: Model 1 is the standard main effect model assuming a linear price effect. Model 2 opens for variation in the price sensitivity between weekdays and Saturdays, based on the restaurant manager's comments of less business people on Saturdays, and model 3 opens for ideal-point pricing. Our aim is not to test which of these three price specifications that best describe the observed choice, but to show that the relative size of the effect of organic and animal welfare is robust to the price specification we use in the model.

Estimation Results

In table 3, we present the estimation results from the three models.⁴ We first notice that the production attributes in our models have only a weak influence on customer choice. In all three models, both animal welfare and organic attributes have the expected positive signs. The effect of animal welfare is larger than the effect of organic. The latter finding is somewhat surprising, given that animal welfare is one of the attributes associated with organic farming.

⁴ See Table A2 in the appendix for the sales data.

We can also see that the results for organic and animal welfare are relatively stable over the various price specifications in the three models.

Table 3. Logit regressions for purchase of veal.

Independent variables	Model 1		Model 2		Model 3	
	Coef.	<i>p</i> value	Coef.	<i>p</i> value	Coef.	<i>p</i> value
Constant	-0.49	0.07	-0.50	0.07	-0.35	0.11
Animal Welfare	0.31	0.14	0.38	0.07	0.34	0.11
Organic	0.23	0.27	0.16	0.47	0.10	0.64
Price	-0.13	0.28				
PriceW			-0.96	0.47		
PriceS			-0.33	0.07		
PriceL					-0.38	0.09
PriceH					-0.78	0.00
N	462		462		462	
Log likelihood	-305.95		-304.63		-302.29	
<i>Pseudo R</i> ²	0.0095		0.001		0.02	
<i>p</i> value, χ^2	0.12		0.07		0.03	

As can be seen from the *p* values, organic and animal welfare statements were not important drivers of choice in our experiment. Hence, other factors determine what customers eat at this restaurant. If they want seafood, for example, they choose fish or lobster independent of the description of the meat dish. Similarly, if they are vegetarians, the description of the meat dish has no effect on their choices. The inclusion of organic and animal welfare seems to affect customers' choice only to a very limited degree. As mentioned

above, the organic food sales in Norway is very small, and hence the result for organic in the experiment is consistent with the sales of organic products in stores. Here we would like to note, that even though the sales of organic products are small, the concept of "organic" is familiar to Norwegians. For example, a survey by Kvakkestad, Refsgaard, & Beglann (2011) found that 60% of Norwegians means that the government should aim at increasing the sale of organic food.

For the price parameters, we obtain the expected negative signs in Models 1 and 2, but only the Saturday price parameter in Model 2 is significant at the 10% level. Model 3 has significant negative coefficients for both low and high prices. Compared with the medium price level, low price seems unappreciated. Here we probably have a price signaling effect: Low listed price on the menu may indicate an inferior item, or if you are part of a group, it can signal that you think that the higher priced menu items are too expensive. This may not be something you would like to show in front of your date or business partner, so you might select a more expensive main course. This result is consistent with Alfnes & Sharma (2010), who find that when they price their local dish equal to their regular dishes, there is no preference for the local dish. However, when the local menu item is the highest priced it is also preferred, as long as it is not too high priced.

In general, Kiefer, Kelly, & Burdett (1994) find low price elasticity for some menu items. Hence, increasing the price of a menu item does not necessarily mean reduced demand for that menu item.

Discussion on Design Issues in Restaurant Experiments

We would like to point out five design issues that are important to take into consideration when designing restaurant experiments.

Issue 1. Choice of restaurant is also choice of sample

Conducting a natural field experiment in a restaurant means that the sample will be the customers that walk into the restaurant. By choosing restaurant, you also choose what type of sample you will get. For example, conducting the experiment at a restaurant close to a university campus means that you will mainly get a sample from the university. Most restaurants target a specific part of the population. In order to get something resembling a representative sample, one needs multiple restaurants.

Issue 2. Restaurants seldom sell two products that differ only in credence attributes

Restaurants usually sell a relative small number of dishes and it is very rare to find two dishes on a menu that only differ in credence attributes. Hence, if one wants to see how introducing a credence attribute like organic affects the sale of a dish, one should compare the sales for the dish on days where the dish is described as organic with the days it is not. Alternatively, one can do like Alfnes & Sharma (2010), who always included both non-local and local food, but the two dishes differed. They changed the menu on a regular basis, in fact every day. One day the local food was a chicken dish and the non-local a pork dish, the next day the local was pork and the non-local chicken, and so on. The researcher can then analyze the choice data to see if the local had any effect on the choice of set menu.

One reason for having dishes only differing in credence attributes on the menu the same day, is if one want to study how including several close substitutes will affect sales. Would we for example sell more of a beef dish if the customers could choose between several countries of origins of the beef?

Issue 3. Time is an important factor in field experiment that goes over several weeks

Since restaurant customers typically make only one choice out of an often rather large choice set, we need many respondents to document changes in purchase frequency of a specific dish. As a result, restaurant experiments usually need several weeks to get enough data.

Over the time it takes to run the experiment, products might change due to seasonality, external factors, such as weather or media events, can change consumer preferences, or the customer base for the restaurant might change. For example, in case of the restaurant we studied, the head chef told us that they significantly changed the menu in the summer holiday due to more international tourist wanting Norwegian specialties. It is therefore important to make sure that one controls for time

differences in the design. This can either be done by using a diff-in-diff approach with treatment and control restaurants, or if one let several attributes change, by making time one of the attributes in the design so that there is no correlation between time and treatment attributes. Alternatively, one can of course do both of the above.

Issue 4. Keeping other factors in the restaurant constant

Many factors not related to what we study in the experiment can affect the choice probability of a menu item. Waiter recommendations, special offers on other products, and anything that change the atmosphere of the restaurant are likely to affect the choices (Bell, Meiselman, Pierson, & Reeve, 1994; Bouwen & Morris, 1995). For example, a special offer on a specific type of wine can make the restaurant sell more of products that go well with it. Therefore, a researcher must be careful instructing the restaurant manager and waiters to avoid unnecessary noise from factors not included in the experimental design.

Issue 5. Restaurant jargon

Restaurant jargon differs from the language typically used to elicit preferences in surveys. In surveys, we often try to use only objective descriptions of the attributes of interest. An example is survey instructions explaining the differences in animal welfare for two different production methods. In restaurants, the focus is on selling the products and, therefore, the product descriptions are all positive, like "Mamma's juicy Italian roasted chicken". Furthermore, there is almost no use of labels, and only limited use of brands describing dishes in restaurants. This makes it very challenging to describe production process attributes of the dishes.

Concluding remarks

The restaurant market is a very important market for agricultural goods, however close to nothing of the now very extensive literature on consumer preference for food products focus on this market. The exception is the relative new literature on nutritional labeling in restaurants. In this paper, we have described how one can conduct a restaurant choice experiment where the participants are ordinary customers walking in from the street and do

not know that they take part in an experiment. We have used standard choice methodology in both design (fractional factorial design for attributes) and in analysis (logit models). We have also discussed the pros and cons of field experiments in restaurants, and some design challenges.

The main advantage is observing the customer in the context of interest without affecting their choices by letting them know that they are part of a study. This points in the direction of high external validity. However, the lack of control and lack of background information is a serious problem for the internal validity. From the estimation results in our illustrative example, we can also see that the very limited amount of data coming out of the experiment is a disadvantage. A similar stated choice experiment conducted in a web survey would have given multiple answers per respondent, and therefore a much richer data set.

In our illustrative example, we manipulate the menu text and price of an organic veal dish in an haute cuisine restaurant. We find that animal welfare and organic attributes had very limited effect on the choices customers made. This is in line with the very low share of organic meat in Norwegian grocery stores, and indicates that most restaurants should emphasize other aspects of their products. However, there might of course be room for niche restaurants focusing entirely on organic food, and thereby recruiting customers specially interested in organic food. Survey results indicate that Norwegians are aware of the concept of organic foods, and that they would like to buy more of it. The main reason stated by Norwegian for eating organic food is to avoid pesticide remnants in the food (Kvakkestad et al., 2011). When Norwegians go to a restaurant it is because they want to cozy time, they do not what to think about possible harmful ingredients in the food. In addition, the word "organic" does not appeal so much in the restaurant setting, since very few Norwegians try to eat as much organic food as possible (op cit.). That is, the term does not trigger purchase.

Animal welfare related to food in general, and veal in particular, is not a big issue in Norway, and gets very limited media coverage. Hence, it is not strange that labeling the dish as animal friendly only seems to have a marginal effect on sales. We should be very careful with drawing strong conclusions based on our illustrative example, but the results indicates none or only limited interest for organic and animal friendly attributes from restaurant customers in Norway.

The experiences with this experiment calls for further research in to the possibilities of using natural field experiments in restaurants to complement other methods used in food economics. We currently have very limited knowledge about what influences customers' choices in restaurants. Hence, more research on this topic seems essential. As we see it, natural field experiments in restaurants as illustrated in this paper can be an important contribution to the toolbox for investigating consumer preferences for premium food products consumer choices in HORECA in general.

References

- Alfnes, F. & Rickertsen, K. (2011). Non-Market Valuation: Experimental Methods. In J. L. Lusk, J. Roosen & J. F. Shogren (Eds.), *The Oxford Handbook of the Economics of Food Consumption and Policy* (pp. 215-242). New York, NY: Oxford University Press.
- Alfnes, F., & Sharma, A. (2010). Locally Produced Food in Restaurants: Are the Customers Willing to Pay a Premium and Why? *International Journal of Revenue Management*, 4(3/4), 238–258. doi: 10.1504/IJRM.2010.035955
- Bell, R., Meiselman, H. L., Pierson, B. J., & Reeve, W. G. (1994). Effects of Adding an Italian Theme to a Restaurant on the Perceived Ethnicity, Acceptability, and Selection of Foods. *Appetite*, 22(1), 11–24. doi: dx.doi.org/10.1006/appe.1994.1002
- Benkahla, A., Boutonnet, J.-P. & Fort, F. (2005). *Enjeux de la certification d'origine et stratégies d'acteurs: Le cas de l'AOC pélaridon*. [Stakes of certification of origin and actors strategies: the case of the PDO pelardon.] Retrieved April 3, 2014 from http://epe.cirad.fr/fr2/doc/enjeux_certif_orig.pdf
- Bouwen, J. T. & Morris, A. J. (1995). Menu Design: Can Menus Sell? *International Journal of Contemporary Hospitality Management*, 7(4), 4–9. doi: 10.1108/09596119510091699
- Cai, H., Yuyu, C. & Fang, H. (2009). Observational learning: Evidence from a Randomized Natural Field Experiment. *American Economic Review*, 99, 864-882. doi: 10.1257/aer.99.3.864
- Cummings, R. G., Harrison, G. W., & Rutström, E. E. (1995). Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive-Compatible? *American Economic Review*, 85, 260–266.
- Ding, M., R. Grewal & Liechty, J. 2005. Incentive-Aligned Conjoint Analysis. *Journal of Marketing Research*, 42, 67–82. doi: <http://dx.doi.org/10.1509/jmkr.42.1.67.56890>
- Eurostat (2014). *Harmonised indices of consumer prices – Item weights*. Retrieved April 2, 2014, from: <http://epp.eurostat.ec.europa.eu/portal/page/portal/hicp/data/database>
- Fisher, R. J. (1993). Social Desirability Bias and the Validity of Indirect Questioning. *Journal of Consumer Research*, 20, 303–315. doi: 10.1086/209351
- Gabor, A., Granger, C. W. J., & Sowte, A. P. (1970). Real and Hypothetical Shop Situations in Market Research. *Journal of Marketing Research*, 7, 355–359.
- Harrison, G. W. & List, J. A. (2004). Field Experiments. *Journal of Economic Literature*, 42, 1009–1045. doi: 10.1257/0022051043004577
- Howlett, E. A., Burton, S., Bates, K., & Huggins, K. (2009). Coming to a Restaurant Near You? Potential Consumer Responses to Nutrition Information Disclosure on Menus. *Journal of Consumer Research*, 36, 494–503. doi: 10.1086/598799
- Jaeger, S. R., & Rose, J. M. (2008). Stated Choice Experimentation, Contextual Influences and Food Choice: A Case Study. *Food Quality and Preference*, 19, 539–564. doi: <http://dx.doi.org/10.1016/j.foodqual.2008.02.005>
- Kiefer, N. M., Kelly, T. J., & Burdett, K. (1994). Menu Pricing: An Experimental Approach. *Journal of Business & Economic Statistics*, 12, 329–337. doi: 10.1080/07350015.1994.10524548
- Kuhfeld, W. F. (2010). *Marketing Research Methods in SAS: Experimental Design, Choice, Conjoint, and Graphical Techniques*. Cary, NC: SAS Institute. Inc.
- Kvakkestad, V. Refsgaard, K., & Berglann, H. (2011). *Citizen and consumer attitudes to food and food production in Norway*. (Discussion paper no. 2011-2). Oslo: Norwegian

- Agricultural Economics Research Institute. Retrieved from nilf.no:
http://www.nilf.no/publikasjoner/Discussion_Papers/2011/dp-2011-02.pdf
- Levitt, S. D., & List, J.A. (2007). What do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? *Journal of Economic Perspectives*, 21(2), 153–174. doi: 10.1257/jep.21.2.153
- Lusk, J. L., & Fox, J. A. (2002). Value Elicitation in Retail and Laboratory Environments. *Economic Letters*, 79, 27–34. doi: [http://dx.doi.org/10.1016/S0165-1765\(02\)00284-7](http://dx.doi.org/10.1016/S0165-1765(02)00284-7)
- McCall, M., & Lynn, A. (2008). The Effects of Restaurant Menu Item Descriptions on Perceptions of Quality, Price, and Purchase Intention. *Journal of Foodservice Business Research*, 11, 439–445. doi: 10.1080/15378020802519850
- National Restaurant Association (2012). *2013 Restaurant Industry Pocket Factbook*. Retrieved April 4, 2014 from: http://www.restaurant.org/Downloads/PDFs/News-Research/Factbook2013_LetterSize.pdf
- Norwegian Agricultural Authority (2014). *Produksjon og omsetning av økologiske landbruksvarer. Rapport for 2013*. [Production and sales of organic foods. Report for 2013.] (Report no. 8/2014). Oslo: Norwegian Agricultural Authority- Retrieved from slf.dep.no (SLF): https://www.slf.dep.no/no/miljo-og-okologisk/okologisk-landbruk/om-okologisklandbruk/_attachment/36328?ts=1450336feb8&download=true
- Norwood, F. B., & Lusk, J. L. (2011). Social Desirability Bias in Real, Hypothetical, and Inferred Valuation Experiments. *American Journal of Agricultural Economics*, 93, 528-534. doi: 10.1093/ajae/aaq142
- Roberto, C. A., Larsen, P. D., Agnew, H., Baik, J., & Brownell, K. D. (2010). Evaluating the Impact of Menu Labeling on Food Choices and Intake. *American Journal of Public Health*, 100, 312-318. doi: 10.2105/AJPH.2009.160226
- Statistics Norway (2013). *Table 10235: Expenditure per household per year, by commodity and service group*.
- Steward, H. (2011). Food Away from Home. In J. L. Lusk, J. Roosen, & J. F. Shogren (Eds.), *The Oxford Handbook of the Economics of Food Consumption and Policy* (pp. 647-666). New York, NY: Oxford University Press.
- Wansink, B., Painter, J., & van Ittersum, K. (2001). *Cornell Hotel and Restaurant Administration Quarterly*, 42(6), 68–72. doi: [http://dx.doi.org/10.1016/S0010-8804\(01\)81011-9](http://dx.doi.org/10.1016/S0010-8804(01)81011-9)

Paper 2

Country-of-origin preferences for organic food

By Alexander Schjøll

Abstract

International agreements about organic products have led to harmonization of regulations allowing organic food produced in one country to be sold as organic in another. However, mandatory country-of-origin labelling for food products makes country-of-origin preferences a potential barrier to trade. This study uses an online choice experiment with conventional and organic veal to investigate country-of-origin preferences for organic meat. The results show that despite the fact that the organic label assures consumers that the meat has been produced under the same organic regulations, the willingness to pay for organic meat depends greatly on the country of origin.

Key words

Organic food, country-of-origin labelling, domestic bias, choice experiment.

1. Introduction

International food trade has increased by 514 per cent measured in current prices since 1980 (World Trade Organization, 2014). Two of the major drivers of the increase are reductions in structural barriers, such as the removal of trade tariffs and import quotas, and increases in wealth and population in many countries. Lowering of the structural barriers has also led to agreements about common product standards and cross-classifications of products between countries (Grane and Boys, 2011). An example of the former is the common European standard for organic production signified by the ‘Euro-leaf’ logo for organic food (European Commission, 2007). An example of the latter is the agreement on cross-classifications between the EU and the US, making food sold as organic in the EU automatically classified as organic in the US, and vice versa (USDA, 2012).

While structural trade barriers have been reduced, many countries have strengthened preference barriers by implementing mandatory country-of-origin (COO) labelling for most types of unprocessed food products (European Commission, 2011; USDA, 2009). This allows consumers to make choices based on the place of origin of the products. The literature

investigating COO preferences finds strong country-of-origin preferences, especially for meat, but also for other food products. In the US, for example, Ehmke, Lusk and Tyner (2008) and Tonsor, Schroeder and Lusk (2013) found a clear preference and higher willingness to pay (WTP) for domestic beef, similar to Alfnes and Rickertsen (2003) and Realini *et al.* (2013) in Europe. The consumers' WTP for premium domestic meat operates as a preference barrier, reducing the potential for international trade in meat (Lim *et al.*, 2013).

There are several reasons for the preference for domestic food. Some consumers buy domestic products because they want to support domestic producers or their own country. Others believe that the place of origin of the food influences its quality and safety (Alphonse and Alfnes, 2012). For meat, scandals such as mad cow disease in the 1990s and the recent European 'horse meat scandal' have placed a focus on the place of origin of meat, and therefore likely strengthened preferences for domestic food in most countries.

Quality labels have juridical and technical standards, which should ensure that food carrying the labels is of similar quality regardless of place of origin. For imported food products, such labels might be of special interest if consumers believe that the label means that the imported and domestic products are of similar quality. This takes us to the purpose of this paper.

This paper investigates whether imported organic meat, using the same organic label as domestic meat, faces the same preference barriers as non-organic meat products. The focus is on Norwegian consumers' WTP for organic and non-organic veal produced in Norway, Denmark and Poland. The data come from an online choice experiment (CE) with a representative sample of 953 Norwegian consumers.

The remainder of this paper is organized as follows. The next section provides background on imported and organic meat in Norway. Section 3 describes the literature on

both COO labelling and organic foods. Section 4 describes the data and methods used. Section 5 presents the results, while section 6 discusses them. Section 7 concludes.

2. Background on organic and imported meat in Norway

As a member of the European Economic Space, Norway has adopted the EU's legislation on organic foods, and organic food produced in the EU can carry the Norwegian organic label when sold in Norway. The consumption of organic food in Norway is low, and for meat, the market share is less than one per cent (Norwegian Agricultural Authority, 2013).

Annual meat consumption in Norway is around 51 kg per capita (Animalia, 2013), and of this, about 16 per cent is imported. In recent years, beef has had the highest import share and chicken the lowest (Statistics Norway, SSB, 2014). Adding to the quantity of imported meat is the significant border trade in meat with neighbouring Sweden and Denmark, both EU members with lower meat prices (Lavik, 2006; SSB, 2013a).

3. Literature on imported organic foods

The literature on consumers' perceptions of and WTP for COO and organic food is very large, but there are limited studies combining these two credence attributes, and most of them were conducted in the US. This review focuses on consumer studies that measure both COO and organic food preferences. The studies are grouped according to geographic location.

3.1 North American studies

3.1.1 Meat studies

There are several US meat studies that include both COO and organic production, but not imported organic meat in their designs. Hence, they investigate the importance of COO and the importance of organic meat, but do not investigate COO preferences in relation to organic meat. Examples of this are Umberger *et al.* (2003), Loureiro and Umberger (2005) and Mennecke *et al.* (2007).

Umberger *et al.* (2003) use both a survey and an experimental auction to investigate the importance of different credence attributes among US beef buyers. The survey shows that consumers prefer beef labelled as organic/natural, and they are less interested in COO. The auction finds the opposite result, with a significantly positive effect on WTP of a 'USA Guaranteed' label, but little effect of an organic label.

Using a mail survey in the US, Loureiro and Umberger (2005) map US consumers' ranking of meat attributes. In the study, consumers prefer domestically produced meat, while organic production is not important. The consumers think domestic meat is safer than meat from Canada and Denmark, and their preference for domestic meat is partly due to food safety concerns.

Mennecke *et al.* (2007) find similar preferences for place of origin by using a conjoint study of meat attributes. Their respondents rate 'region of origin' as the most important attribute, while they rate 'organic certification' the sixth most important attribute for steaks. These three US meat studies include both COO and organic meat production, but not imported organic meat in their designs. Hence, they do not investigate COO preferences for organic meat.

3.1.2 Other North American studies

North American studies that investigate COO and organic food in general include the US study of May (2004) and the Canadian study of Lu, Cranfield and Widowski (2013). In a nationwide US survey, 68 per cent of consumers would pay more for food grown in the US than abroad, while 60 per cent would pay a premium for food grown organically (Mey, 2004). Lu, Cranfield and Widowski (2013) investigate Canadian consumers' ranking of different attributes related to food in general, and find that Canadian consumers see COO as more important than organic production.

There are also studies that look at US consumer preferences for COO and organic bread, fruits, berries and baby foods. Lusk and Briggeman (2009) investigate consumer preferences for bread through a best–worst scaling experiment in a mail survey. Among the 11 attributes tested, COO was one of the six that influenced WTP significantly, while organic production was not. Shi, Gao and House (2011) test consumer preferences for blueberries and find that consumers see COO as more important than organic production.

Peterson and Li (2011) study consumer preferences in the US for baby foods using a survey and a CE. When asked about the importance of different attributes, consumers stated that COO of ingredients was on average much more important than organic production. However, in the CE, the authors find identical WTP estimates for the organic production and US-produced attributes. This discrepancy indicates that the participants who said that organic production was important acted on it to a greater degree when asked to make a choice between baby food alternatives.

The two studies that are closest to our analysis of COO preferences for organic meat consider fruit and vegetables. Both include imported organic products in their CE designs. Pozo, Saak and Peterson (2009) use a CE to test US consumers' preferences for COO in relation to organic apples. They include apples from South America, Australia, Europe and China. In the CE, consumers differentiate between the places of origin of organic food, with domestically produced organic food the preferred choice. A similar result is found by Onozaka and Thilmany McFadden (2011), who look at imported apples and tomatoes. Organic imported produce is preferred to other imported produce, but the WTP for imported organic produce is significantly lower than the WTP for domestic organic produce.

3.2 European studies

European studies looking at both COO and organic production are rare. There are only a few examples, but none of them investigates consumer preferences for COO of organic food

products. In Stolz (2011), German and Swiss consumers rank domestic and organic production as being of equal importance. Illichmann and Abdulai (2013) use a CE to measure WTP for organic apples, milk and beef in Germany. For the average German consumer, place of origin is more important than organic production.

To the author's knowledge, no studies have been published on consumer preferences regarding imported organic meat in European countries.

3.3 Asian studies

As with the European studies, Asian studies including both COO and organic production are rare. Two examples are Kim, Suwunnamek and Toyoda (2008) and Zheng, Li and Peterson (2013).

Kim, Suwunnamek and Toyoda (2008) investigate Japanese consumers' WTP for noodles using a conjoint analysis. In this study, as in most studies from the US and Europe, consumers prefer domestic food, and place of origin is more important than organic production.

Zheng, Li and Peterson (2013) map Chinese consumers' preferences for the attributes of soybeans used to produce soymilk. In contrast to the studies mentioned above, they find that organic production is more important than place of origin. In addition, Chinese consumers seem to prefer organic soybeans imported from the US compared with domestic organic soybeans. These preferences might be a result of domestic food scandals in China in recent years.

3.4 Knowledge gap

The papers mentioned above show that there has been very little research on consumer preferences towards imported organic products, and no research on imported organic meat. The closest are two US studies of imported organic apples and tomatoes (Pozo, Saak and

Peterson, 2009; Onozaka and Thilmany McFadden, 2011). Their findings indicate that consumers have a significantly higher WTP for imported organic food than for imported conventional food. However, there is still the question of to what degree the result holds for meat, and whether it holds in Europe, where there is a common labelling scheme. This study attempts to provide new insights into this area by investigating whether organic labelling can overcome the strong preference for domestic production found in most meat studies.

4. Data description

4.1 Sample

The sample consists of 953 Norwegian consumers recruited from TNS Gallup's Norwegian online survey panel in July 2013. The summer send-out resulted in a response rate of around 40 per cent, which is on the lower side for this kind of online panel in Norway. All the participants stated that they were the main purchasers of food in their household and ate meat. Table 1 presents an overview of the descriptive statistics for the sample.

Table 1. Description of the full survey sample and the sample of organic buyers

Socio-demographic characteristics		Sample	Norwegian population
Age	25–29	13	11
	30–34	5	11
	35–39	6	12
	40–44	8	13
	45–49	9	12
	50–54	12	11
	55–59	14	11
	60–64	13	10
	65–70	20	9
Gender	Female	52	50
Region	Eastern Norway	51	51
	Southern Norway	5	5
	Western Norway	25	26
	Mid Norway	9	9
	Northern Norway	10	10
Education level	Primary	6	24
	High school	61	44
	Higher education ≤ 4 years	24	24
	Higher education > 4 years	8	8
Income	< NOK 200,000	0	10
	NOK 200,000 – NOK 399,999	0	24
	NOK 400,000 – NOK 999,999	0	47
	≥ NOK 1,000,000	0	18
	No answer	8	–
	N	953	5,051,275



Source: SSB (2013b, 2013c, 2013d).

As Table 1 indicates, the full sample has a good spread in relation to age, gender, region, education and income. However, the sample has the typical weaknesses of online panels, with an under-representation of the oldest consumers, those with only primary education and those belonging to the highest income groups. Moreover, a significant group (8 per cent) were not willing to state their income.

Table 2 shows the frequency of organic purchases. As seen from the table, 17 per cent state that they ‘always’ or ‘quite often’ choose organic alternatives. In the estimations, these

respondents will be called ‘organic buyers’ and their answers will be compared with those of the full sample.

Table 2. Attributes and attribute levels used in the choice experiment

Attribute	Levels	Labels
Country of origin	Norway, Denmark, Poland	
Price in NOK	45, 50, 55, 60, 65, 70, 75, 80, 85	
Production method	Conventional	
	Friland	
	Organic	

4.2 Description of the choice experiment

The last part of the Web survey was a CE with minced veal meat. This product is very similar to popular products such as minced beef, pork or chicken meat, but is difficult to find in Norwegian grocery stores (Animalia, 2013). Hence, one could expect consumers to be willing to purchase this meat, but be unaware of market prices.

The CE had three alternative labels on the minced meat packages to provide information about the products. The labels were identical for all three products, and all packages had a net weight of 400 grams. They also had barcodes to make the labels as similar to typical store labels as possible. Figure 1 shows examples of the labels used on the meat packages.



Fig. 1. Examples of front labels used on the minced veal packages

As shown in Figure 1 and described in Table 2, the meat differed in three aspects: COO (Norway, Denmark and Poland), production method (conventional, organic and Friland) and price (9-level attribute from NOK 45 to NOK 85).

Poland and Denmark are Europe’s second and fourth largest producers, respectively, of meat from bovine animals (Eurostat, 2013). Denmark is a Scandinavian neighbour with closely related culture and language where many Norwegians shop for food (Lavik, 2006; Synovate MMI, 2012). The cultural and economic differences between Norway and Poland are much larger than those between the Scandinavian countries. These differences are likely to affect how Norwegians view food from the two countries.

It was not possible to find prices for minced veal at the time of the experiment, so the price range is based on minced beef prices and the general price level of veal compared with beef. The price of minced beef varies substantially between producers, production methods and stores. Veal prices are typically higher than the prices of comparable cuts of beef. This is likely partly because veal is a niche product in Norway.

Norway does not have a label for animal welfare, but Friland products have recently been making their way into the Norwegian food market. Friland-labelled food has the same animal welfare standards as organic food, but does not use organic fodder. We used the Danish Friland label as seen in Figure 1 and Table 2, and explained the different production

methods to the participants before they commenced the CE. Figure 2 describes the three production methods.

-
- *Conventional veal is produced following standard rules for space, access to outdoor areas, fodder and the waiting period between medication and slaughter.*
 - *Organic veal is produced following organic production rules including more space, regular access to outdoor areas, organic fodder and an extended waiting time between medication and slaughter.*
 - *Friland veal is produced following the same rules for animal welfare as organic veal, but without organic fodder and an extended waiting time between medication and slaughter.*
-

Fig. 2. Information about production methods given to the participants

The design has $3 \times 3 \times 9 = 81$ possible combinations of attributes and levels, using the full factorial to create the choice sets (giving a D-efficiency of 100 for the factorial design). From the full factorial, 27 choice sets were created using the SAS %ChoiceEff macro (Kuhfeld, 2010). During the Web survey, six of these choice sets were randomly drawn for each respondent, giving each respondent a unique combination of choice sets.

For each choice set, the respondents identified the most and least preferred alternatives, resulting in a full ranking of the three alternatives in each choice set (Hensher, Rose and Greene, 2007: 176). Following ranking CEs, such as Mueller, Lockshin and Louviere (2010) and Eckert, Louviere and Islam (2012), the CE was constructed without an opt-out alternative. That is, in each choice set, respondents had to indicate their most and least preferred alternative. Whereas the above studies have a follow-up question asking whether the respondent would actually buy his/her preferred alternative, the present study does not. This means that the data are limited to estimation of the price premiums of the product attributes, and do not give a total WTP estimate for a specific package of minced veal meat.

4.3 Estimation method

The ranking data are estimated using the following random utility function:

$$\begin{aligned}
U_{ijt} = & \beta_{1i}OrganicNO_j + \beta_{2i}FrilandNO_j + \beta_{3i}OrganicDK_j + \beta_{4i}FrilandDK_j \\
& + \beta_{5i}UnlabelledDK_j + \beta_{6i}OrganicPL_j + \beta_{7i}FrilandPL_j \\
& + \beta_{8i}UnlabelledPL_j + \beta_{9i}Price_j + \varepsilon_{ij},
\end{aligned}$$

where U_{ijt} is the utility of individual i when choosing veal package j in choice situation t . OrganicNO, OrganicDK and OrganicPL refer to organic veal from Norway, Denmark and Poland, respectively. The Friland and Unlabelled variables follow the same structure, and refer to veal with the Friland label and with no label, respectively, from the three countries. Price is per 400 grams of veal, and is measured in NOK.

The utility function is estimated by NLOGIT 5.0 using a mixed logit model with 1,000 Halton draws (Train, 2009: 230), a panel structure and freely correlated random parameters. All non-price parameters are random, following the standard normal distribution. The base alternative is the Norwegian unlabelled package.

The marginal WTP for the product attributes is estimated by dividing the label and place of origin parameters in a mixed logit model by the negative of the price parameter and multiplying by 2.5 to obtain the WTP per kilogram.

5. Results

The respondents were asked how often they choose the organic alternatives if available. As can be seen from Table 3, the majority answered that they sometimes choose the organic alternatives. The organic purchase frequency responses are used to group the respondents into three segments, which are used in the analysis below. The first group is the 133 respondents who said they never buy organic, the second group is the 631 who said they sometimes buy organic, and the third group is the 163 respondents who said they often or always buy organic (called 'typically organic' below).

Table 3. Frequency of organic purchases

Description	Per cent	N
I always choose the organic alternatives	1	8
I quite often choose the organic alternatives	16	155
I sometimes choose the organic alternatives	66	631
I never choose the organic alternatives	14	134
Do not know	3	25
Sum	100	953

5.1 Response to key food attributes

Before the CE, the participants were asked how important a series of attributes are in their choice of food, using a 6-point scale from 1 = ‘not important’ to 6 = ‘very important’. Those who said they bought organic products at least once in a while were also asked how they evaluate organic food compared with conventionally produced food using the same series of attributes, this time on a 6-point scale from 1 = ‘organic much worse than conventional’ to 6 = ‘organic much better than conventional’. Tables 4 and 5 present ranked lists of attributes from the two questions. Table 4 presents the results for the whole sample and for three subsamples based on how often they said they buy organic. Table 5 presents results for the ‘sometimes organic’ and ‘typically organic’ subsamples.

Table 4. When you buy food on a normal day, how important are the following attributes for your choice of food? 1 = ‘not important’, 6 = ‘very important’

Attributes	Full sample		Never organic		Sometimes organic		Typically organic	
	Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank
Taste	5.48	1	5.46	1	5.44	1	5.62	1
Freshness	5.24	2	5.19	2	5.19	2	5.46	3
Smell	5.12	3	5.11	3	5.08	3	5.26	4
Pesticides/medicine rests	4.91	4	4.52	7	4.85	4	5.49	2
Looks	4.79	5	4.76	5	4.80	5	4.73	10
Durability	4.76	6	4.93	4	4.79	6	4.55	11
Healthiness	4.71	7	4.48	8	4.64	7	5.19	5
Price	4.53	8	4.63	6	4.61	8	4.12	14
Country of origin	4.41	9	4.37	9	4.39	9	4.51	12
Animal welfare	4.24	10	3.58	12	4.15	10	5.17	6
Conditions for workers	4.14	11	3.61	11	4.07	11	4.84	8
Environmentally friendly	4.10	12	3.30	14	4.01	13	5.15	7
Short travel distance	4.10	13	3.50	13	4.04	12	4.81	9
Vitamin content	3.97	14	3.68	10	3.93	14	4.35	13
Number of participants	953		134		631		163	

Note: ‘Do not know’ answers are excluded. The percentage of this answer varies between 0% and 3.3%.

As seen from Table 4, taste, freshness and smell are very important for all subsamples when buying food. This seems reasonable, because these attributes are very important for the sensory experience of the food. At the lower end of the ranking, we mainly find attributes not easily connected to the sensory experience, such as COO (ranked 9 out of 14 in the full sample). It is important to note that these questions ask about food in general, and not about meat in particular. The ranking for meat may therefore differ from this general food ranking.

The largest differences in ranking in Table 4 are between the ‘never organic’ and the ‘typically organic’ subsamples. The ‘typically organic’ subsample ranks low levels of pesticides and longer medicine rests, animal welfare and environmentally friendly production much higher than the ‘never organic’ subsample, whereas the ‘never organic’ subsample

ranks looks, durability and price much higher than the ‘typically organic’ subsample. The latter finding is as expected, because the buyers of only conventional products are accustomed to greater durability and less blemishes on produce than regular buyers of organic products. Furthermore, in most food categories, the organic alternatives are more expensive than conventional foods (Oikos, 2012), so the buyers of organic products must be less concerned about price than typical Norwegian consumers, who buy 60% of their food at soft discount stores⁵ (Dagligvarehandelen, 2014).

Table 5. How do you evaluate the following attributes for organic food in general compared with conventional food? 1 = ‘organic much worse than conventional’, 6 = ‘organic much better than conventional’

Attributes	Sometimes organic		Typically organic	
	Mean	Rank	Mean	Rank
Pesticides/medicine rests	5.11	1	5.70	1
Animal welfare	4.80	2	5.53	2
Environmentally friendly	4.74	3	5.52	3
Healthiness	4.44	4	5.33	4
Short travel distance	4.30	5	5.04	5
Conditions for workers	4.27	6	4.99	6
Country of origin	4.24	7	4.93	7
Vitamin content	4.13	8	4.93	8
Taste	4.09	9	4.85	9
Freshness	4.02	10	4.82	10
Smell	3.98	11	4.67	11
Looks	3.74	12	4.64	12
Durability	3.37	13	4.36	13
Price	2.83	14	3.65	14
Number of participants	631		163	

Note: ‘Do not know’ answers are excluded. The percentage of this answer varies between 13.9% and 29.9%.

⁵ Soft discount stores are grocery stores not only competing on price, but other factors like placement of store as well.

Table 5 shows how the organic food buyers evaluate organic food compared with conventional food. The ranking is identical for the two groups, but the ‘typically organic’ group score organic food higher on all attributes. The factors often promoted as the main reasons to buy organic (low levels of pesticides and longer medicine rests, animal welfare and environmentally friendly production), are the factors these consumers rank highest. At the other end of the scale are looks, durability and price. In Norway, organic alternatives are typically more expensive than conventional food (Oikos, 2012), so the price result mirrors actual market features. From Table 4, we can see that looks, durability and price are not important for the ‘typically organic’ consumers, but somewhat important for the ‘sometimes organic’ consumers. The importance of these three attributes is likely to explain some of the differences in purchase practices between the ‘typically organic’ and ‘sometimes organic’ consumers.

5.2 Estimation results

Table 6 presents the estimation results from the mixed logit model. The first panel presents results for the full sample, while the second panel presents results for the organic buyers (‘sometimes organic’ + ‘typically organic’). Unlabelled meat from Norway is the base alternative in the estimation, and the preferences and WTP estimates are relative to the base.

Table 6.

Mixed logit estimates with unlabelled Norwegian veal as the base alternative

Full sample		Mixed logit		WTP per kg	
		Coeff.	SE	Coeff.	SE
Random					
OrganicNO	Mean	1.47***	0.21	44.62***	6.03
	SD	(3.00***)	(0.21)		
FrilandNO	Mean	1.12***	0.18	33.26***	5.36
	SD	(1.87***)	(0.18)		
OrganicDK	Mean	-0.84***	0.18	-25.11***	5.37
	SD	(4.12***)	(0.31)		
FrilandDK	Mean	-1.10***	0.17	-33.32***	4.97
	SD	(3.38***)	(0.31)		
UnlabelledDK	Mean	-2.56***	0.18	-76.18***	5.03
	SD	(2.85***)	(0.35)		
OrganicPL	Mean	-4.19***	0.27	-124.78***	7.89
	SD	(6.26***)	(0.41)		
FrilandPL	Mean	-4.59***	0.25	-136.60***	7.11
	SD	(5.58***)	(0.35)		
UnlabelledPL	Mean	-4.41***	0.26	-131.35***	7.70
	SD	(4.54***)	(0.43)		
Fixed					
Price	Mean	-0.08***	0.00		
Organic buyers					
Random					
OrganicNO	Mean	4.59***	1.18	253.56***	67.10
	SD	(5.60***)			
FrilandNO	Mean	3.68***	0.84	203.31***	52.41
	SD	(3.69***)			
OrganicDK	Mean	3.24***	0.93	178.74***	58.03
	SD	(6.32***)			
FrilandDK	Mean	0.15	0.72	8.15	39.84
	SD	(5.18***)			
UnlabelledDK	Mean	-3.99***		0.85	-220.51***
	SD	(4.49***)			50.05
OrganicPL	Mean	-0.29	0.87	-15.97	47.86
	SD	(9.09***)			
FrilandPL	Mean	-3.20***	0.86	-177.01***	49.33
	SD	(7.06***)			
UnlabelledPL	Mean	-6.61	4.86	-365.26	263.84
	SD	(6.20***)			
Fixed					
Price	Mean	-0.045***	0.01		
		Full sample	Organic buyers		
Number of choices observations		5,706	978		
Number of respondents		953	163		
Log likelihood		-5,800.85	-877.76		

P value χ^2 (45 df)	0.000	0.000
McFadden Pseudo R^2	0.078	0.174
AIC	11,691.70	1,845.50

***, ** and * denote significance at 1%, 5% and 10%, respectively.

Table 6 shows a clear preference for domestic meat for both the full sample and the organic buyers. For the full sample, all estimates for foreign meat, both labelled and unlabelled, are negative. COO seems to be much more important than production method when selecting meat. For the organic buyers, the production method is more important than for the full sample. They are willing to pay a significant premium for Danish organic meat compared with unlabelled Norwegian meat; however, the same does not hold for Polish organic meat.

For Norwegian meat, the full sample is willing to pay a premium for Norwegian organic and Friland meat compared with conventional meat. Danish meat shows similar results. However, the full sample is not willing to pay a premium for Polish organic and Friland meat compared with conventional Polish meat. Looking at the organic buyers only, the same results appear for Danish meat (Organic: Wald = 399.26; P value = 0.00; Friland: Wald = 228.66; P value = 0.00) and Polish meat (Organic: Wald = 349.29; P value = 0.19; Friland: Wald = 188.26; P value = 0.69).

Consumers seems to pay little attention to the difference between organic and Friland products. A test between Norwegian organic and Friland-labelled meat for the full sample shows no significant difference in WTP between the two products (Wald = 10.36; P value = 0.09). Similar results are found for the Danish meat (Wald = 8.21; P value = 0.09) and the Polish meat (Wald = 11.82; P value = 0.10). However, the organic buyers have a significantly higher WTP for Danish organic meat compared with Friland meat from Denmark (Wald = 170.60; P value = 0.00) and Poland (Wald = 161.03; P value = 0.00), but not for Norwegian meat (Wald = 50.25, P value = 0.46). The Norwegian organic buyers are thus willing to pay a

premium for Danish organic meat, but not for Danish Friland meat or for any meat from Poland.

6. Discussion

Consumers often find it challenging to assess the quality of foreign food products. Labels stating that the products have been produced in accordance with a strict set of regulations are meant to reduce uncertainty about quality. However, in order to achieve the labels' intended effect, consumers must trust the labelling and feel comfortable that it guarantees superior quality. In relation to imported meat, the organic and Friland labels are not fully trusted by Norwegian consumers.

Consumers' high level of trust in domestic meat is not peculiar to Norway or to this study. In a recent survey, 70 per cent of Norwegians agreed with the statement 'Norwegian meat is safer than foreign meat' (Animalia, 2013). In a study of consumers in seven European countries, Kjærnes, Poppe and Lavik (2005) found a higher proportion of consumers agreeing with the statement 'Domestic beef is very safe' in Denmark and Great Britain than in Norway.

Most Norwegians are able to identify the Debio label (the organic label used in this study) as a label identifying organic food (Heidenstrøm, Jacobsen and Borgen, 2011). The Friland concept has only recently been introduced on a few products in Norway, and was likely unknown to most participants before the study. The lack of a difference in WTP in relation to these two production methods among consumers in general indicates that sellers of organic food face a challenge in convincing most consumers that it is superior to other premium products. For consumers who typically buy organic food, domestic Friland meat is almost as good as domestic organic meat, and can be an important competitor in this segment. However, with regard to imported meat, organic buyers differentiate between organic and Friland. This indicates a higher level of trust in domestic fodder and medication practices. In relation to the imported meat, consumers want to be on the safe side by buying meat produced

with organic fodder and that has experienced an extended waiting time between medication and slaughter.

In the survey part of this study, which asked about food in general, COO was not very important when choosing food, while in the CE it was very important. This could be because of the two different question formats, or because the first question was about food in general, while the second question was specifically about meat. Norwegian consumers are more used to imported produce and processed food products than imported animal products such as milk, eggs and meat. The strong preferences for domestic products indicated in the CE are therefore possibly a result of conducting the CE on meat.

As in most studies, this study has limitations that require further research. Most important is the fact that it uses non-consequential survey questions, and therefore is likely to suffer from a hypothetical bias in WTP values. As discussed by Lusk, Fields and Prevat (2008), the hypothetical bias is likely to inflate the WTP values, but not change their relative size. Second, including multiple countries would have produced results that are more robust and provided the opportunity to see how the WTP values depend on the amount of meat import and the number of organic buyers in a country.

7. Conclusion

Consumers are willing to pay a premium for organic meat, but knowing where the meat is produced is important to them. A common organic label is not enough to overcome the country-of-origin preferences seen for other meat products. This is despite the fact that the common organic label is supposed to provide a guarantee that the products have been produced under the same strict organic regulations.

The implications of these results for agribusiness seem clear. A common organic label, well established in both the exporting and importing countries, is not enough to overcome COO preferences. The Norwegian consumers' preference for organic meat from Denmark

over organic meat from Poland indicates that organic meat with few food miles and coming from a country that is seen as relatively similar to the importing country is best suited for import. In the Norwegian case, therefore, the best options would be organic meat from fellow Scandinavian countries such as Sweden and Denmark, while in the US, for example, the best option would likely be organic meat from Canada. A second result that should be of interest to agribusiness is that products that have some but not all the characteristics of organic meat do quite well compared with organic meat. As organic fodder is an important cost driver for organic meat, the producers of Friland meat are likely to accept a lower price than producers of organic meat, and thereby take market share from both conventional and organic meat.

A common international organic label is important for increasing international trade in organic foods. However, in order to ensure such trade, consumers need to trust both the label and the country of origin. If policymakers want to increase trade, they have to implement policies that increase trust in production methods across borders and reduce the importance of place of origin preferences. At present, production location is a more important choice attribute for consumers than production method.

As discussed above, this study has some limitations that should motivate future research. Revealed preference studies, field experiments, studies in more countries and on different food products are important to test the validity of our results in other countries, on other products and most importantly, in real purchase contexts.

References

- Alfnes, F. and Rickertsen, K. (2003). European consumers' willingness to pay for US beef in experimental auction markets. *American Journal of Agricultural Economics* 85: 396–405.
- Alphonse, R. and Alfnes, F. (2012). Consumer willingness to pay for food safety in Tanzania: an incentive-aligned conjoint analysis. *International Journal of Consumer Studies* 36: 394–400.
- Animalia (2013). Kjøttets tilstand 2013. Status i norsk kjøtt- og eggproduksjon. [Situation for the meat 2013. Status in Norwegian meat and egg production.] Animalia. Oslo: Animalia.
- Dagligvarehandelen (2014). Lavpris mot 60% – Rema og Kiwi fosser frem. [Discount segment towards 60% - Rema and Kiwi surging forward.] <<http://www.dagligvarehandelen.no/lavpris-mot-60-rema-og-kiwi-fosser-frem/>> (accessed 21.03.2014).
- Eckert, C., Louviere, J. J. and Islam, T. (2012). Seeing the forest despite the trees: brand effects on choice uncertainty. *International Journal of Research in Marketing* 29: 256–264.
- Ehmke, M. D., Lusk, J. L. and Tyner, W. (2008). Measuring the relative importance of preferences for country of origin in China, France, Niger, and the United States. *Agricultural Economics* 38: 277–285.
- European Commission (2011). Council Regulation No. 1169/2011 on the Provision of Food Information to Consumers, amending Regulations No. 1924/2006 and No. 1925/2006 of the European Parliament and of the Council, and repealing Commission Directive 87/250/EEC, Council Directive 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and Commission Regulation No 608/2004. Brussels: European Commission.
- European Commission (2007). Council Regulation (EC) No. 834/2007 of 28 June 2007 on Organic Production and Labelling of Organic Products and repealing Regulation (EEC) No. 2092/9.1. Brussels: European Commission.
- Eurostat (2013). Meat Production and Foreign Trade. Luxembourg: Eurostat.
- Grane, J. H. and Boys, K. A. (2011). Agricultural trade and the GATT/WTO: does membership make a difference? *American Journal of Agricultural Economics* 94: 1–24.
- Heidenstrøm, N., Jacobsen, E. and Borgen, S. O. (2011). Seleksjon og ignorering. Forbrukerstrategier for å manøvrere i merkemangfoldet. [To select or ignore: Consumer strategies for manoeuvring in the label diversity.] Oppdragsrapport no. 2-2011, National Institute for Consumer Research (SIFO). Oslo: SIFO.
- Hensher, D., Rose, J. M. and Greene, W. H. (2007). *Applied Choice Analysis. A Primer*. Cambridge: Cambridge University Press.
- Illichmann, R. and Abdulai, A. (2013). Analysis of Consumer Preferences and Willingness-to-Pay for Organic Food Products in Germany. In Contributed Paper presented at the AAEA Annual Meeting, Washington DC, August 2013.
- Kim, R., Suwunnamek, O. and Toyoda, T. (2008). Consumer attitude towards organic labeling schemes in Japan. *Journal of International Food & Agribusiness Marketing* 20: 55–71.
- Kjærnes, U., Poppe, C. and Lavik, R. (2005). Trust, Distrust and Food Consumption. Project Report no. 15-2005, National Institute for Consumer Research (SIFO). Oslo: SIFO.

- Kuhfeld, W. F. (2010). *Marketing Research Methods in SAS: Experimental Design, Choice, Conjoint, and Graphical Techniques*. Cary, NC: SAS Institute Inc.
- Lavik, R. (2006). Grensehandelen 2006-enkelte utviklingstrekk. [Border trade 2006-some developments.] Hurtigstatistikk fra SIFO-survey 2006. Prosjektnotat nr. 4-2006, National Institute for Consumer Research (SIFO). Oslo: SIFO.
- Lim, K. H., Hu, W., Maynard, L. J. and Goddard, E. (2013). U.S. consumers' preference and willingness to pay for country-of-origin-labeled beef steak and food safety enhancements. *Canadian Journal of Agricultural Economics* 61: 93–118.
- Loureiro, M. L. and Umberger, W. J. (2005). Assessing consumer preferences for country-of-origin labeling. *Journal of Agricultural and Applied Economics* 37: 49–63.
- Lu, Y., Cranfield, J. and Widowski, T. (2013). Consumer Preference for Eggs from Enhanced Animal Welfare Production Systems: A Stated Choice Analysis. In Contributed Paper at the AAEA Annual Meeting, Washington DC, August 2013.
- Lusk, J. L. and Briggeman, B. C. (2009). Food values. *American Journal of Agricultural Economics* 91: 184–196.
- Lusk, J. L., Fields, D. and Prevatt, W. (2008). An incentive compatible conjoint ranking mechanism. *American Journal of Agricultural Economics* 90: 487–498.
- Mennecke, B. E., Townsend, A. M., Hayes, D. J. and Lonergan, S. M. (2007). A study of the factors that influence consumer attitudes toward beef products using the conjoint market analysis tool. *Journal of Animal Science* 85: 2639–2659.
- Mey, B. J. V. (2004). The globalization of food and how Americans feel about it: results of two surveys. *Journal of Food Distribution Research* 35: 6–17.
- Mueller, S., Lockshin, L. and Louviere, J. J. (2010). What you see may not be what you get: asking consumers what matters may not reflect what they choose. *Marketing Letters* 21: 335–350.
- Norwegian Agricultural Authority (2013). Produksjon og omsetning av økologiske landbruksvarer. Rapport for 2012. [Production and sales of organic food. Report for 2012.] Rapport nr. 12/2013, Norwegian Agricultural Authority (SLF). Oslo: SLF.
- Norwegian Competition Authority (2005). Betaling for hylleplass. Virkninger for konkurransen i dagligvaremarkedet i Norge. [Payment for shelf space. Effects for the competition in the Norwegian grocery market.] Konkurransetilsynets skriftserie 2/2005, Norwegian Competition Authority. Oslo: Norwegian Competition Authority.
- Oikos (2012). Økologisk mat stadig rimeligere! [Organic food increasingly cheaper!] Oslo: Oikos.
- Onozaka, Y. and Thilmany McFadden, D. (2011). Does local labeling complement or compete with other sustainable labels? A conjoint analysis of direct and joint values for fresh produce claims. *American Journal of Agricultural Economics* 93: 693–706.
- Peterson, H. H. and Li, X. (2011). Consumer preferences for product origin and processing scale. The case of organic baby food. *American Journal of Agricultural Economics* 93: 590–596.
- Pozo, V., Saak, A. and Peterson, H. H. (2009). Product Origin and Reputation for Quality: the case of Organic Foods. In Contributed Paper presented at the AAEA Annual Meeting, Milwaukee, July 2009.
- Realini, C. E., Font i Furnolsa, M., Sañudob, C., Montossic, F., Olivera, M. A. and Guerreroa, L. (2013). Spanish, French and British consumers' acceptability of Uruguayan beef, and consumers' beef choice associated with country of origin, finishing diet and meat price. *Meat Science* 95: 14–21.
- Shi, L., Gao, Z. and House, L. A. (2011). Consumer WTP for Blueberry Attributes: A Hierarchical Bayesian Approach in the WTP Space. In Contributed Paper presented at the AAEA Annual Meeting, Pittsburgh, July 2011.

- Statistics Norway (2014). Table 08811: External Trade in Goods, By Commodity Group, Country and Mode of Transport, Statistics Norway (SSB). Oslo: SSB.
- Statistics Norway (2013a). Table 08460: Cross Border Trade. Number of Same Day Trips and Expenditure, Statistics Norway (SSB). Oslo: SSB.
- Statistics Norway (2013b). Table 07459: Population, By Sex and One-year Age Groups, Statistics Norway (SSB). Oslo: SSB.
- Statistics Norway (2013c). Table 08921: Persons 16 Years and Older, By Sex, Age and Level of Education, Statistics Norway (SSB). Oslo: SSB.
- Statistics Norway (2013d). Table 07752: Households, By Type of Household and Total Household Income, Statistics Norway (SSB). Oslo: SSB.
- Stolz, H. (2011). Attitudes and Behaviour Towards Food Attributes and Organic Food – A Triangular Methodological Investigation of Occasional Organic Consumers. Dissertation at the University of Kassel, Faculty of Organic Agricultural Sciences. Kassel: University of Kassel.
- Synovate MMI (2012). Norske Spisefakta 2012. [Norwegian facts about eating practices 2012.] Oslo: Synovate MMI.
- Tonsor, G. T., Schroeder, T. C. and Lusk, J. L. (2013). Consumer valuation of alternative meat origin labels. *Journal of Agricultural Economics* 64: 676–692.
- Train, K. (2009). *Discrete Choice Methods with Simulation*. 2nd edn. New York: Cambridge University Press.
- Umberger, W. J., Feuz, D. M., Calkins, C. R. and Sitz, B. M. (2003). Country-of-origin labeling of beef products: U.S. consumers' perceptions. *Journal of Food Distribution Research* 34: 103–116.
- United States Department of Agriculture (2012). European Union and United States Agree to Historic New Partnership on Organic Trade, News Release no. 0051.12, United States Department of Agriculture (USDA). Washington: USDA.
<http://www.usda.gov/wps/portal/usda/usdahome?contentid=2012/02/0051.xml&> Accessed 21 March 2014.
- United States Department of Agriculture (2009). Country of Origin Labeling, United States Department of Agriculture (USDA). Washington: USDA.
<http://www.ams.usda.gov/AMSV1.0/cool> Accessed 21 March 2014.
- World Trade Organization (2014). Merchandise Trade by Commodity, World Trade Organization (WTO). Geneva: WTO.
- Zheng, Y., Li, X. and Peterson, H. H. (2013). In pursuit of safe foods: Chinese preferences for soybean attributes in soymilk. *Agribusiness* 29: 377–391.

Paper 3

Animal welfare in stores: natural field experiments with positive and negative labeling

By Alexander Schjøll, Frode Alfnes and Svein Ole Borgen

Abstract

Producers label their products to send positive quality signals to consumers and tempting them to buy their products. In order to differentiate products within the same product category, producers typically add one or more positive labels for each step in the quality ladder. Alternatively, the products could be differentiated by using labels informing about problematic aspects of the product or production process. In this paper, we report the results of natural field experiments involving the sale of eggs in stores. We find that negative animal welfare labeling on eggs produced under the least animal-friendly conditions (battery eggs) has a significant negative impact on sales, while positive animal welfare labeling on eggs produced under the most animal-friendly conditions (organic eggs) has no effect on sales. Hence, the strongest impact on sales does not come from positive encouragement (positive product labeling), but rather from the revelation of more problematic information (negative labeling).

Keywords

Animal welfare, Natural field experiment, Labeling, Eggs

Highlights

- Natural field experiments were conducted with animal welfare labeling of eggs.
- We compare the effects of positive and negative animal welfare labels.
- Negative labeling decreases sales of the least animal-friendly alternative.
- Positive animal-friendly labeling has no effect on sales.

1. Introduction

Many consumer studies find their subjects are highly engaged in animal welfare issues (see, e.g., Carlsson et al., 2007a; Dentoni et al., 2011; Gracia et al., 2011; Grimsrud et al., 2013; Kehlbacher et al., 2012; Kjærnes and Lavik, 2008; Lagerkvist and Hess, 2011; Norwood and Lusk, 2011a, 2011b; Vecchio and Annunziata, 2011). For example in Norwood and Lusk (2011b), US consumers are found to be willing to pay twice the price for organically produced eggs than for standard battery eggs produced by hens living in small cages. Of course, organic eggs are more expensive to produce than battery eggs, but if US consumers on average were willing to pay twice as much for the former, they would presumably have a larger market share than the current two percent (USDA, 2013). Some questions then arises;

are consumers really concerned about animal welfare; if so, is positive animal welfare claims an efficient way of conveying the animal welfare information?

This paper investigates how positive and negative animal welfare labeling affects the egg purchasing behavior of consumers. Positive labeling is letting the eggs carry a label indicating a higher level of care for animal welfare than the average product in the product category. For example, positive labels could declare that the eggs are free-range. In contrast, negative labeling is letting the eggs carry a label indicating a lower level of care for animal welfare than the average product in the product category. An example of the latter is a label on the product declaring that the contents are battery eggs. Producers and retailers wish to sell their products and therefore avoid grading products using negative labels on inferior products. In fact, most instances of negative labeling of products come in the form of mandatory warnings on unsafe products or as mandatory classifications. The case we investigate in this paper is a negative animal welfare label for eggs, voluntarily initiated and orchestrated by a Norwegian retail chain.

In general, mandatory warnings on food products are uncommon, but consumers in most Western countries are nonetheless accustomed to seeing warning texts or alarming pictures on the front of cigarette packages (Thrasher et al., 2007). It is also generally considered that the negative cigarette labeling has had its intended impact and increased the willingness of smokers to quit (Larsen et al., 2005; Zacher et al., 2014). More recently, there have been attempts to put sugar warnings on soft drinks in California (California State Senate, 2014). If implemented, sugar-containing soft drinks in California would from July 2015 display the following warning: “STATE OF CALIFORNIA SAFETY WARNING: Drinking beverages with added sugar(s) contributes to obesity, diabetes, and tooth decay.” The US beverage industry is of course challenging this proposal.

Mandatory classifications are more common on food products. The European Union's egg labeling scheme including the mandatory labeling of battery eggs is an example. Another example is the labeling signifying food safety in restaurants. In Denmark, for example, a smiley scheme was introduced in 2001, which comprised smileys ranging from a big smile to a sad face depending on the most recent public food inspections. Food enterprises are obliged to publish their smiley at their entrances, making it possible for customers to make informed choices (Nielsen, 2006). This public display of food safety standards has resulted in improved standards in Danish food enterprises (Danish Veterinary and Food Administration, 2011). Similar grading from best to worst can also be found regarding the energy efficiency of home appliances (Kallbekken et al., 2013; Sammer and Wüstenhagen, 2006) and car safety (Seeck et al., 2003).

However, most of the clearly negative information about food products is not presented at the point of purchase, but is available online and in other types of media. Pressure groups such as the World Wide Fund for Nature (WWF) and the Royal Society for the Prevention of Cruelty to Animals (RSPCA) publish substantial amounts of negative information about food products they recommend people avoid. At the same time, these organizations also support the positive labeling of those products they consider better. The best example of the latter is the WWF-supported Marine Stewardship Council label for sustainable fisheries. Other sources of negative information can be consumer test panels used by magazines or online review sites such as TripAdvisor, which publish customer evaluations of restaurants and hotels. Overall, it is likely that information presented at the point of purchase, such as the customer grading of hotels on online booking sites, is a more powerful instrument affecting consumer decisions than information presented elsewhere.

We wish to add to the literature on positive and negative labeling with the results of three natural field experiments concerning animal welfare in Norway. The second largest

retail chain in Norway itself initiated two of the experiments, and we instigated the third in two of the largest stores of the retail chain. Animal welfare in egg production is the common topic of all three experiments. We use egg sales over 19 months from all stores in the retail chain to investigate the effects of the store-initiated experiments and sales data over five weeks from two large stores for the third experiment.

The main experiment was conducted from August to December 2011. In August, the retail chain ceased labeling battery eggs as “Farm Eggs” (with no information concerning the production method) and voluntarily introduced a carton with “Eggs from Battery Hens” in large bold capital letters. A retail chain voluntarily putting such negative information on one of their products is a very rare occurrence, and therefore worth studying in itself. As part of the negative information, the store announced that it would not sell battery eggs after January 1, 2012. The second experiment concerns the withdrawal of battery eggs from the stores’ shelves; restraining the choice of their customers to only eggs with a higher level of animal welfare. In the third experiment, conducted in November 2011, we used positive animal welfare information in the form of posters and stickers to promote organic eggs.

According to Harrison and List’s (2004) taxonomy of field experiments, our three experiments are all natural field experiments. They are conducted in the natural setting of the behavior to be studied, and participants do not know an experiment is going on. The respondents’ lack of awareness implies they will not deviate from their “true behavior”. For normative attributes like farm animal welfare, natural field experiments is especially appropriate because of the problem of social desirability bias in studies where respondents know they are being observed (Fisher, 1993). According to Norwood and Lusk (2011c), this bias is difficult to eliminate even when using lab experiments with real products and sales. In addition, Johansson-Stenman and Svedsäter (2012) warn that the use of real money in experiments does not remove all biases, and it can be difficult to generalize results to other

valuation contexts such as stores. Hence, a natural field experiment seems appropriate given our interest in studying consumer behavior as it relates to animal welfare. Our study also supplements the many surveys and lab experiments conducted to map consumer perceptions of animal welfare (Burrell and Vrieze, 2003; Carlsson et al., 2007b; Gerhardy and Ness, 1995; Kjærnes and Lavik, 2008; Norwood and Lusk, 2011b).

The contextual background for this study is European Union (EU) Directive 1999/74/EC (European Commission, 1999) effective January 1, 2012. This directive prohibits conventional small battery cages for hens and allows for larger “enriched cages.” The new cages must have at least 750 cm² per hen, and each cage must be furnished with a roost, bedding, and a claw-shortening device. Facing this new regulation, the second largest retail chain in Norway decided to abandon battery eggs altogether, and asked its egg suppliers to invest in barn production systems instead. Vertical integration has given the retail chain full control over the egg-packaging firm and the information on the packages. The retail chain’s corporate social responsibility strategy and wish to enhance their reputation probably motivated the decision to abandon battery eggs completely. Thus, the rather unconventional decision to label the battery eggs negatively five months before their removal provided good publicity about its decision to eventually abandon battery eggs.

Discussing the huge discrepancy between the stated interest in animal welfare in many consumer studies and the low sales of animal welfare-labeled products, Andersen (2011) concludes that the stated interest of consumers in animal welfare is mostly “cheap talk.” Alternative explanations to the low sales are that consumers are uninformed about animal welfare differences, and that they have stopped paying attention to many positive product claims. If this hypothesis holds, increased positive labeling on products with the best animal welfare credentials can be expected to have a limited effect on consumer choice, while negative labeling on the worst animal welfare credentialed products can be expected to have a

significant effect. Our study therefore investigates the effect of both positive and negative animal welfare labeling.

2. Material and methods

We use sales data from Norway's second largest retail chain REMA 1000 to learn more about consumer attitudes to animal welfare. This chain first introduced a new egg carton declaring its battery egg status in bold capital letters, and later removed battery eggs from their stores nationwide. To our knowledge, our study is the first to report a natural field experiment concerning the negative labeling of animal welfare concerns. Furthermore, no previous paper has investigated how customers react when products with the lowest level of care for animal welfare (such as battery eggs) have been removed. In addition to the two chain initiated experiments, we conducted an experiment over five weeks with positive animal welfare information labeled on organic eggs.

2.1 Egg terminology

Eggs are especially suitable for studying consumer attitudes and behavior relating to animal welfare because consumers can select eggs based upon different levels of animal welfare (see the meta-analysis by Lagerkvist and Hess, 2011). The following hierarchy of animal welfare for eggs is evident in Norwegian stores, starting at the lowest level of animal welfare: (1) Battery system, referring to hens living in cages. (2) Barn system, implying hens can move freely indoors, either in aviaries or on the floor. (3) Organic systems, implying hens can move freely both indoors and outdoors (Freedom Food, 2014).

In Norway, "free-range eggs" is the commonly used term for indoor free-range hens, corresponding to the "barn system" in the EU regulations (European Commission, 2003: L 340/31). In contrast, in the EU, the term "free-range eggs" applies when hens have continuous daytime access to open-air runs. In this paper, we keep to the EU definition and use the three

terms detailed earlier. However, the reader should note that the common term for barn eggs in Norway is free-range eggs, which may provide more positive animal welfare associations than the EU term. For organic eggs, Norway follows the EU regulation (European Commission, 1991), namely, organic hens must have access to outdoor areas. Each hen should have at least 6 m² of area indoors and 4 m² outdoors in which to move freely. All fodder must be organic.

2.2 Experiment 1: Negative labeling of battery eggs

The main experiment was conducted by the Norwegian retail chain REMA 1000 in August 2011 when they introduced a new carton for their battery eggs. Fig. 1 illustrates the change.

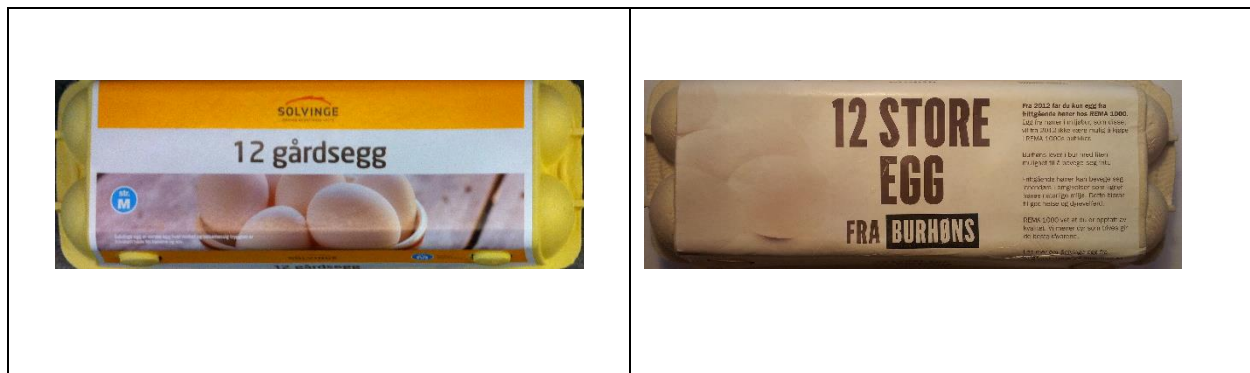


Fig. 1. The original (left-hand panel) and new (right-hand panel) battery egg cartons.

Prior to the change in the battery egg carton, consumers had no information about animal welfare on the battery egg cartons sold in REMA 1000 (or in fact in any other Norwegian retail chain). The original carton (left-hand panel in Fig. 1) merely says (in Norwegian) “12 farm eggs.” The new cartons have “BATTERY HENS” printed on them in large white letters on a black background. The design of the new carton resembles the warning text on packets of cigarettes in Europe.

On the right side of the new carton, consumers also found the following information: “From 2012, you can only buy eggs from barn systems at REMA 1000. Eggs from hens in cages, as these, are not available from 2012 for purchase in REMA 1000 stores. Battery hens live in cages with little opportunity to move freely. Hens living in barn systems can move

indoors in environments similar to their natural environment. This contributes to good health and welfare. REMA 1000 knows you are concerned about quality. We believe animals that thrive provide the best ingredients.”

The negatively labeled eggs were exclusive to this chain and produced by an egg producer owned by the chain. There were no other battery eggs in the store, but there were barn eggs available from two other producers, including the leading national brand, along with organic eggs. Other store chains changed from conventional to enriched battery eggs without making either public announcements or any form of associated marketing.

2.3 Experiment 2: Removal of battery eggs from the stores

As noted, the retail chain officially stopped selling battery eggs from January 1, 2012. We refer to this change as Experiment 2. By abandoning all types of battery eggs, the retail chain went a step further than the new EU Directive, which allows eggs from enriched battery egg systems. An interesting question concerns what purchasers of battery eggs would do. Would they choose other eggs (barn or organic) or start buying their eggs from other stores? Furthermore, would purchasers of barn eggs purchase the eggs now having the lowest animal welfare level in the store (the barn eggs) or would they start buying organic eggs? We use nationwide sales data from the retail chain from January to March 2012 to investigate what happened to the egg customers when the store removed the battery eggs.

2.4 Experiment 3: In-store experiment with positive information

We conducted the third experiment in two large REMA 1000 stores in Oslo before the ban on conventional battery systems came into effect. In this experiment, we used posters and labels with positive information concerning the animal welfare of organic eggs, which are the eggs with the highest level of associated animal welfare (Freedom Food, 2014). As there is currently no animal welfare label in Norway, we created one for the experiment, comprising a picture of a hen and the text “better animal welfare.” The label was blue and white, 3.5 cm

wide and 1.5 cm high, and was used on both the poster and as stickers on the organic egg packages. Fig. 2 depicts the poster (left-hand panel) and the sticker (right-hand panel).



Fig. 2. Poster (left-hand panel) and sticker (right-hand panel) used to provide in-store information to consumers.

The text (in Norwegian) on the poster reads “Organic eggs – Happier hens. REMA 1000 is not selling battery eggs after January 1, 2012. If you would like even better animal welfare, choose organic eggs.”

Our research collaborator, the Norwegian Animal Protection Alliance, put their name under the “better animal welfare” label in the lower right corner of the poster as a third-party supporter of the animal welfare message. The impending removal of battery eggs was the poster’s main message, but it also stated that organic eggs had the highest level of animal welfare involved in their production. The sticker only declared the high animal welfare of organic eggs.

We applied four levels of information over a five-week period in November–December 2011. Weeks 1 and 5 were reference weeks with no information; Week 2 had the poster hanging on the egg shelf; Week 3 had the stickers on the organic egg cartons; and Week 4 had both the poster and the stickers. We regularly inspected the stores and made sure the amount and placement of eggs on the shelf were similar for the five weeks of the experiment.

2.5 Order of experiments

In general, Experiments 2 and 3 are dependent on Experiment 1. When the battery eggs were removed (Experiment 2), a five-month period had elapsed with negative labeling and information declaring that the retail chain would cease selling the battery eggs after January 1, 2012. This may have affected customer attitudes and behavior toward animal welfare concerns. We could raise similarly confounding arguments about Experiment 3, which occurred after the introduction of the new carton, but before the removal of the battery eggs. Whereas Experiment 1 focused on the lowest level of animal welfare, Experiment 3 concerned the highest level of animal welfare. The battery egg consumers who were most sensitive to animal welfare likely changed their egg-buying behavior in August when the negatively framed carton appeared. Thus, when we introduced the positive information, some of the customers had already “moved up” one animal welfare category. Furthermore, at the time of Experiment 3, consumers were likely already paying increased attention to animal welfare in the retail chain’s egg shelves.

Throughout the experiments, battery eggs were sold only in 12 egg packages, barn eggs were sold both in 6 and 12 egg packages, and organic eggs were only sold in 6 egg packages. The small packages are typically more expensive per egg than the larger packages, which may dissuade some customers who might have bought the organic eggs if available in large packages from their purchase. However, according to an in-store survey of Norwegian egg buyers (Schjøll et al., 2013), the number of eggs in the carton is not an important attribute for consumers when buying eggs, especially not for organic egg buyers.

3. Literature on positive and negative information and labeling on food products

A large pool of scholarly literature shows consumers react more to negative information than positive information. According to Fox et al. (2002: 77), such “...behavior is consistent with several (economic) models – reference risk effects, loss aversion, status quo bias, ambiguity

aversion, alarmist reactions – in which people seem to focus more on the downside than the upside.” Here, we focus on some insights from the psychology literature and provide some examples of positive and negative information and the labeling of food in the economic literature.

3.1 Positive and negative information in the psychology literature

Baumeister et al. (2001) review the empirical research on negative vs. positive events in the psychology literature. In most cases, their review concludes that negative aspects have a greater influence than positive; they provide several explanations for this finding. The first relates to fear. In nature, fear (something negative) is more important than joy (something positive), because ignoring danger can harm or kill you, while missing an opportunity for pleasure has no dramatic consequences. The second is that negative information typically signals a need for action, while positive information typically signals a steady course. Finally, humans consider negative information as more informative than positive because negative characteristics deviate from the norm.

Psychologists have studied the reaction to positive and negative stimuli in many contexts. Examples are Fiske (1980), Ito et al. (1998), and Yzerbyt and Leyens (1991). In Fiske (1980), participants viewed pictures of behaviors varying from positive to negative and were found to spend significantly more time viewing the negative pictures than the positive pictures. Elsewhere, Ito et al. (1998) studied brain responses to pictures, which were affectively neutral, positive, or negative to the participants. The experiment found the largest brain response to negative stimuli. In other words, humans appear to pay more attention to negative than positive information. Skowronski and Carlston (1989) provide an early overview of theories on how the brain reacts to positive and negative information more generally.

Studying how much information people need to make a judgment, Yzerbyt and Leyens (1991) provided information about an actor to a group of participants and asked them to judge if the actor was suitable for a role. Respondents could freely decide when they had sufficient information to make their judgment. Less negative information was necessary to disqualify an actor for a likeable role, than positive information to disqualify an actor from an unlikable role. The evidence indicated that people seemed to make faster decisions with less information when they received negative information than positive information. As summarized in the review by Baumeister et al. (2001), psychologists often identify the disproportionate effects of negative and positive stimuli in a wide range of contexts.

3.2 Positive and negative information about food products

Much of the food information literature focuses on products where there is some potential risk associated with the product, and different groups send out conflicting information about the products. For instance, Fox et al. (2002) used experimental auctions to examine the effects of alternative descriptions of food irradiation on the willingness to pay for a pork sandwich irradiated to control for foodborne pathogens. As expected, a favorable (unfavorable) description of irradiation increased (decreased) the willingness to pay for the pork sandwich. Of particular interest for the present study, in this experiment when subjects had access to both pro- and anti-irradiation information, the anti-irradiation information dominated and the willingness to pay decreased almost as much as when there was only anti-irradiation information.

Few existing studies address how positive and negative information affects choices in a natural field setting. One interesting exception is Lusk's (2010) study of egg sales in California in the months leading up to the vote on a proposition to ban the use of cages in egg production in the state in 2008. Lusk (2010) argues that most consumers were hitherto unknowledgeable of egg production practices, and therefore the proposition and the

campaigns related to it served as an information shock to consumers. Both sides conducted information campaigns to persuade consumers to vote in favor of or against the proposition. Hence, Californian consumers had both positive and negative information available about egg production practices. It was found that actual sales of eggs associated with higher animal welfare standards (barn systems and organic) increased over time and in response to newspaper articles concerning the proposition. Conversely, the demand for cage eggs fell. This indicates that also in natural settings, negative information will outweigh positive information when one presents both types to consumers.

We have not been able to find any studies that find the strongest effect for positive information, but some studies find equal effects of positive and negative information. In Akaichi et al. (2012), positive and negative information about organic food had the equal positive and negative effects on consumers' willingness to pay. This contradicts the results in Fox et al. (2002) and Lusk (2010) where negative information had the strongest effect.

3.3 Positive and negative labeling on food products

In Grankvist et al. (2004), labeling a product as better (positive labeling) or worse (negative labeling) for the environment than the average product in a given product category was found to have little impact on consumers with limited interest in environmental issues. For consumers with an intermediate interest in environmental issues, the negative label had a greater effect than the positive label. Consumers who already were concerned about environmental issues, and therefore bought environmentally friendly products regularly, displayed the same probability of choosing products with positive and negative labels.

Balcombe et al. (2010) used a hypothetical choice experiment to see how respondents reacted to baskets of food with a traffic light label indicating the levels of salt, sugar, fat, and saturates, respectively. One result from this study throws light on negative labeling, as there was evidence of higher willingness-to-pay estimates for respondents moving from red to

amber than from amber to green. This indicates consumers are more interested in avoiding unhealthy food than they are in selecting healthy alternatives.

A second traffic light study is Levy et al. (2012). They studied traffic light labeling to promote healthy food choices in a US hospital cafeteria. This study followed the purchases of individual cafeteria guests (i.e., hospital employees) over time while conducting manipulations. This was possible as a card automatically drew money from the guest's salary as the payment method. The findings indicated that employees decreased their red-labeled purchases more than they increased green-labeled purchases.

A final color-based labeling study worth mentioning is Vanclay et al. (2011), which tested color-based CO₂ labeling on a series of food products over a three-month period in an Australian convenience store. Green (below average), yellow (near average), and black (above average) footprints reflected the carbon emissions embodied in grocery products. They found evidence of a small overall change in purchasing patterns after labeling, with the sales of black-labeled products decreasing 6% and sales of green-labeled products increasing 4%. A clearer warning sign of high CO₂ emissions, such as the use of red instead of black labeling, may have exerted a greater impact for the negative label.

4. Results

4.1 Results from Experiments 1 and 2

REMA 1000 offers many different egg packages, which we placed into three categories based on their level of animal welfare (battery, barn, and organic). Fig. 3 presents the market shares (in percentages) for egg packages across the three categories sold from August 2010 to March 2012, representing some 20 million egg packages.

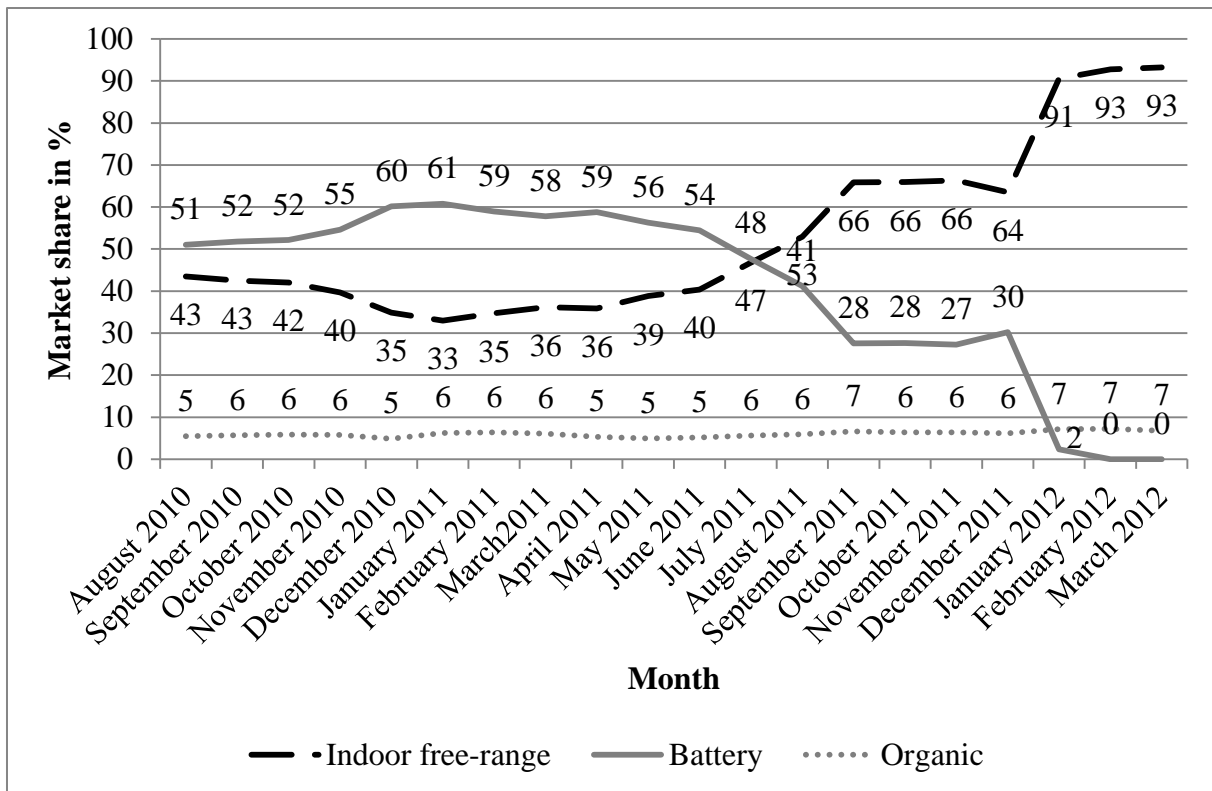


Fig. 3. Market share for the three animal welfare categories of eggs in REMA 1000. N = 19,652,797.

In mid-July, REMA 1000 introduced its new battery egg package in some stores, and had a nationwide rollout of the new package in August 2011. As shown, the market share of battery egg packages fell from 54% in June to 28% in September 2011, after varying between 51% and 61% in the month before the introduction of the new package. From September 2011, the sales of battery eggs were relatively stable until they were finally withdrawn from the market at the end of the year.

On January 1, 2012, the retail chain removed battery eggs from its shelves. Some of the stores sold out the battery eggs they had left over from 2011. According to the sales data, the stock of battery eggs was completely gone from the stores a few weeks into 2012. As illustrated in Fig. 3, the market share for barn eggs increased by the same amount battery eggs fell in both August and January. Most consumers therefore see these two types of eggs as close substitutes. Furthermore, there was no negative shift in the total sales of eggs in REMA

1000. This suggests that REMA 1000 did not lose egg customers by first relabeling and then later removing the battery eggs.

The market share for organic eggs remained constant both when the negative battery egg cartons were released and when the battery eggs were removed from the stores. The added focus on animal welfare from the negative battery egg cartons did not imply that a larger share of customers bought the eggs with the highest animal welfare. Neither did the removal of the battery eggs, in practice now making the barn eggs the product with the lowest level of animal welfare on the egg shelf.

The apparent lack of interest of consumers in organic eggs may have multiple causes. First, consumers may be unaware of the stricter animal welfare regulations for organic systems when compared with barn production, and therefore do not discern any animal welfare-related reason to choose organic eggs. We investigate this possible reason in Experiment 3. Second, consumers may consider that organic eggs are too highly priced.

4.2 Results from Experiment 3

We conducted Experiment 3 in two REMA 1000 stores over five weeks in the fall of 2011, after the introduction of the new packages, but before the removal of the battery eggs. We analyzed the sales data over these five weeks, comprising approximately 20,000 egg package sales. Fig. 4 summarizes the market shares for battery, barn, and organic egg sales in the five weeks. Table 1 presents the results of a logit model testing if the poster or the stickers affected the likelihood of buying organic eggs.

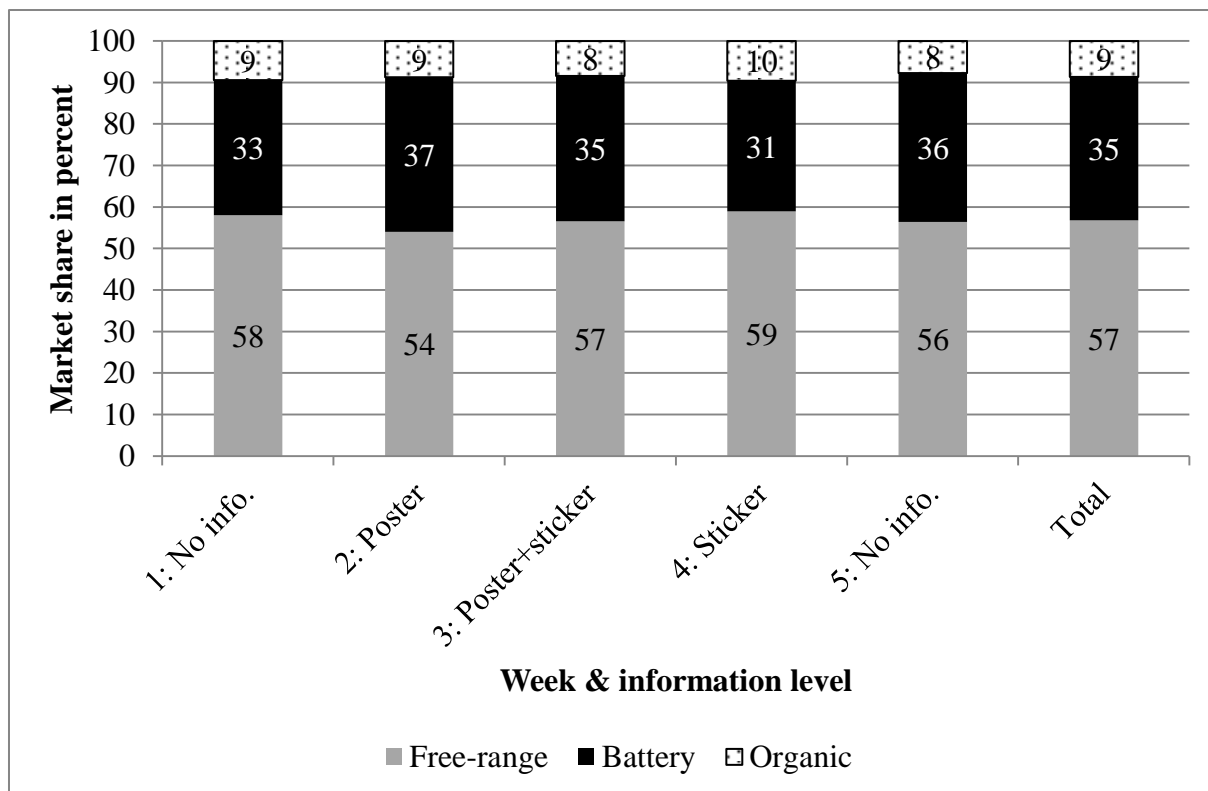


Fig. 4. Market shares for the three egg categories during the five weeks of manipulation. N = 19,717.

Table 1. Logistic regression of organic egg sales as a function of information.

Variable	Coefficient	SE	p-value
Constant	-2.238	0.058	0.000
Poster	-0.110	0.083	0.187
Poster+sticker	-0.146	0.084	0.080
Sticker	0.112	0.123	0.361
Organic low price	-0.157	0.126	0.213
Week 5	-0.018	0.030	0.534
Log-likelihood	-5,825		
R ²	0.001		
p-value χ^2	0.006		
N	19,717		

As shown in Fig. 4, the market share for organic eggs ranged around 9% over the five-week period, with no major changes resulting from the poster or the stickers. The highest market share of 10% for organic eggs was in the week with the stickers. The logit regression

in Table 1 confirms the very stable result from the figure. There are no significant effects of any of the treatments and no significant differences between week 1 and 5, as seen from the non-significant coefficient from the dummy variable indicating week 5.

Furthermore, the small price decrease on organic eggs had no significant effect on sales of organic eggs. This can be seen from the non-significant dummy for days where the organic eggs had a 10% lowered price. The low market share for organic eggs is not due to lack of knowledge about the superior animal welfare of organic egg production. Even with posters and stickers informing about organic being the best animal welfare alternative, nine out of ten customers still bought eggs from battery or barn systems. To summarize, the added positive information and labeling had almost no effect on the average egg consumer.

5. Conclusions and implications

This paper explores how consumers react to farm animal welfare labeling. We investigate the effects of both positive and negative labeling schemes. With positive labeling, a product claims to have higher animal welfare than other products in the product segment. In contrast, negative labeling signifies the product is associated with lower animal welfare than the alternatives. From our three natural field experiments conducted in Norway's second largest retail chain, we find no effect on sales from positive animal welfare information and the labeling of organic eggs. However, negative labeling on the least animal welfare-friendly eggs resulted in a major shift in consumer choice. Our finding of a disproportionate effect of positive and negative animal welfare labeling is consistent with most of the literature on information effects.

The implications for animal welfare organizations are clear. If they wish to trigger a demand shift toward more animal welfare-friendly production, negative labeling and information on problematic situations has a much stronger effect than the promotional effect of the positive labeling of premium animal-friendly qualities. However, a negative labeling

strategy can be difficult to implement, as few food producers would voluntarily label their products with information about low animal welfare. Hence, animal welfare organizations should work toward improved animal welfare regulations, including the mandatory labeling of battery eggs and other products with low levels of animal welfare involved in their production, and make use of information channels allowing negative information, such as online information campaigns and media coverage of animal welfare issues.

For producers and retailers, the results of Experiment 1 show the strength of negative information, especially if presented at the point of purchase. Negative information can change the market composition of an otherwise very stable product category. Experiments 1 and 2 show that when there are more than two animal welfare levels, altering or removing the product with the lowest animal welfare level only affects those products with a medium animal welfare level and not those products with the highest animal welfare level. The results of Experiment 3 reveal the limited effect of positive labeling of attributes to which most consumers paid little attention. Furthermore, most consumers are not willing to pay a large premium to obtain the best animal welfare alternative. This implies that the shift to higher welfare alternatives will take some time.

References

- Akaichi, F., Nayga, R.M. Jr, Gil, J.M., 2012. Assessing consumers' willingness to pay for different units of organic milk: evidence from multiunit auctions. *Can. J. Agr. Econ.* 60 (4), 469–494.
- Andersen, L.M., 2011. Animal welfare and eggs: cheap talk or money on the counter? *J. Agr. Econ.* 62 (3), 565–584.
- Balcombe, K., Fraser I., Di Falco, S., 2010. Traffic lights and food choice: a choice experiment examining the relationship between nutritional food labels and price. *Food Policy* 35 (3), 211–220.
- Baumeister, R.F.; Bratslavsky, E.; Finkenauer, C.; Vohs, K.D., 2001. Bad is stronger than good. *Rev. Gen. Psychol.* 5 (4), 323–370.
- Burrell, A., Vrieze, G., 2003. Ethical motivation of Dutch egg consumers. *Tijdschr. voor sociaalwetenschappelijk onderzoek van de landbouw* 18 (1), 30–42.
- California State Senate, 2014. Senate bill No. 1000. <http://www.leginfo.ca.gov/pub/13-14/bill/sen/sb_0951-1000/sb_1000_bill_20140213_introduced.pdf> (accessed 19.03.14).
- Carlsson, F., Frykblom, P., Lagerkvist, C.J., 2007a. Consumer willingness to pay for farm animal welfare: mobile abattoirs versus transportation to slaughter. *Eur. Rev. Agric. Econ.* 34 (3), 321–344.
- Carlsson, F., Frykblom, P., Lagerkvist, C.J., 2007b. Farm animal welfare testing for market failure. *J. Agric. App. Econ.* 39 (1), 61–73.
- Danish Veterinary and Food Administration, 2011. Smileys keep food safety high in Denmark. <<http://www.findsmiley.dk/en-US/Forside.htm>> (accessed 19.03.14)
- Dentoni, D., Tonsor, G., Calantone, R., Peterson, H.C., 2011. “Animal welfare” practices along the food chain: how does negative and positive information affect consumers? *J. Food Prod. Mark.* 17 (2–3), 279–302.
- European Commission, 1991. Council regulation 2092/91 of 24 June 1991 on organic production of agricultural products and indications referring thereto on agricultural products and foodstuffs. <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:1991:198:0001:0015:EN:PDF>> (accessed 19.03.14)
- European Commission, 1999. Council directive 1999/74/EC of 19 July 1999 laying down minimum standards for the protection of laying hens. <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:1999:203:0053:0057:EN:PDF>> (accessed 19.03.14)
- European Commission, 2003. Commission regulation (EC) No 2295/2003 of 23 December 2003 introducing detailed rules for implementing council regulation (EEC) no. 1907/90 on certain marketing standards for eggs. <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:340:0016:0034:EN:PDF>> (accessed 19.03.14)
- Fox, J.A., Hayes, D. J., Shogren, J.F., 2002. Consumer preferences for food irradiation: how favorable and unfavorable descriptions affect preferences for irradiated pork in experimental auctions. *J. Risk. Uncertainty* 24 (1), 75–95.
- Fisher, R.J., 1993. Social desirability bias and the validity of indirect questioning. *J. Consum. Res.* 20 (2), 303–315.
- Fiske, S.T., 1980. Attention and Weight in Person Perception: The Impact of Negative and Extreme Behavior. *J. Pers. Soc. Psychol.* 38 (6), 889–906.
- Freedom Food. 2014. Shopping guide for eggs. <<http://www.freedomfood.co.uk/wheretobuy/egg-guide>> (accessed 19.03.14)

- Gerhardy, H., Ness, M.R., 1995. Consumer preferences for eggs using conjoint analysis. *World Poultry Sci. J.* 51 (2), 203–214
- Gracia, A., Loureiro, M.L., Nayga, R.M., 2011. Are valuations from non-hypothetical choice experiments different from those of experimental auctions? *Amer. J. Agr. Econ.* 93 (5), 1358–1373.
- Grankvist, G., Dahlstrand U., Biel, A., 2004. The impact of environmental labelling on consumer preference: negative vs. positive labels. *J. Consum. Policy* 27 (2), 213–230.
- Grimsrud, K., Nielsen, H.M., Navrud, S., Olesen, I., 2013. Households' willingness-to-pay for improved fish welfare in breeding programs for farmed Atlantic salmon. *Aquaculture* 372–375, 19–27.
- Harrison, G.W., List, J.A. 2004. Field experiments. *J. Econ. Litt.* 42 (4), 1009–1045.
- Ito, T.A, Larsen, J.T., Smith, N.K., Cacioppo, J.T., 1998. Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *J. Pers. Soc. Psychol.* 75 (4), 897–900.
- Johansson-Stenman, O., Svedsäter, H., 2012. Self-image and valuation of moral goods: stated versus actual willingness to pay. *J. Econ. Behav. Organ.* 84 (3), 879–891.
- Kallbekken, S., Sælen, H., Hermansen, E.A.T. 2013. Bridging the energy efficiency gap: a field experiment on lifetime energy costs and household appliances. *J. Consum. Policy* 36 (1), 1–16.
- Kehlbacher, A., Bennett, R., Balcombe, K., 2012. Measuring the consumer benefits of improving farm animal welfare to inform welfare labelling. *Food Policy* 37 (6), 627–633.
- Kjærnes, U., Lavik, R., 2008. Opinions on animal welfare and food consumption in seven European countries, in: Kjærnes, U., Block, B.B., Roe, E., Roex, J. (Eds.), *Consumption, Distribution and Production of Farm Animal Welfare. Opinions and Practices within the Supply Chain. Welfare Quality Reports no. 7.* School of City and Regional Planning, Cardiff University, Cardiff, pp. 3-117.
- Lagerkvist, C.J., Hess, S., 2011. A meta-analysis of consumer willingness to pay for farm animal welfare. *Eur. Rev. Agric. Econ.* 38 (1), 55–78.
- Larsen, E., Lund, M., Lund, K.E., 2005. Evaluering av helseadvarslene på tobakkspakkene. [Evaluation of health warnings on tobacco packages.] SIRUS skrifter no. 6/ 2005. Norwegian Institute for Alcohol and Drug Control, Oslo.
- Levy, D.E., Riis, J., Sonnenberg, L.M., Barraclough, S.J., Thorndike, A.N., 2012. Food choices of minority and low-income employees: a cafeteria intervention. *Am. J. Prev. Med.* 43 (3), 240–248.
- Lusk, J.L., 2010. The effect of proposition 2 on the demand for eggs in California. *J. Agric. Food Ind. Organ.* 8 (1), article 3.
- Nielsen, A., 2006. Contesting competence: change in the Danish food safety system. *Appetite* 47 (2), 143–151.
- Norwood, F.B., Lusk, J.L., 2011a. *Compassion, by the Pound. The economics of farm animal welfare.* Oxford University Press, New York, NY.
- Norwood, F.B., Lusk, J.L., 2011b. A calibrated auction-conjoint valuation method: valuing pork and eggs produced under differing animal welfare conditions. *J. Environ. Econ. Manag.* 62 (1), 80–94.
- Norwood, F.B., Lusk, J.L., 2011c. Social desirability bias in real, hypothetical, and inferred valuation experiments. *Amer. J. Agr. Econ.* 93 (2), 528–534.
- Sammer, K. and Wüstenhagen, R., 2006. The influence of eco-labelling on consumer behaviour—results of a discrete choice analysis for washing machines. *Bus. Strat. Environ.* 15 (3), 185–199.

- Schjøll, A., Borgen, S.O., Alfnes, F., 2013. Consumer preference for animal welfare when buying eggs. Professional report no. 1-2013. National Institute for Consumer Research, Oslo.
- Seeck, A., Coxon, C., Wani, K., 2003. Global NCAP harmonization: initial position, requirements and prospects. *Int. J. Crashworthiness* 8 (1), 5–15.
- Skowronski, J.J., Carlston, D.E., 1989. Negativity and extremity biases in impression formation: a review of explanations. *Psychol. Bull.* 105 (1), 131–142.
- Thrasher, J.F., Rousu, M.C., Anaya-Ocampo, R., Reynales-Shigematsu, L.M., Arillo-Santillán, E., Hernández-Ávila, M., 2007. Estimating the impact of different cigarette package warning labels polices: the auction method. *Addict. Behav.* 32 (12), 2916–2925.
- USDA, 2013. Table 3. Certified organic and total U.S. acreage, selected crops and livestock, 1995–2011 <<http://www.ers.usda.gov/data-products/organic-production.aspx#.UwxrVv15NjR>> (accessed 19.03.14)
- Vanclay, J.K., Shortiss, J., Aulsebrook, S., Gillespie, A.M., Howell, B.C., Johanni, R., Maher, M.J., Mitchell, K.M., Stewart, M.D., Yates, J., 2011. Customer response to carbon labelling of groceries. *J. Consum. Policy* 34 (1), 153–160.
- Vecchio, R., Annunziata, A., 2011. Italian consumer awareness of layer hens' welfare standards: a cluster analysis. *Int. J. Consum. Stud.* 36 (6), 647–655.
- Yzerbyt. V.Y., Leyens, J.-P., 1991. Requesting information to form an impression: the influence of valence and confirmatory status. *J. Exp. Soc. Psychol.* 27 (4), 337–356.
- Zacher, M., Bayly, M., Brennan, E., Dono, J., Miller, C., Durkin, S., Scollo, M., Wakefield, M., 2014. Personal tobacco pack display before and after the introduction of plain packaging with larger pictorial health warnings in Australia: an observational study of outdoor café strips. *Addiction* 109 (4), 653–662.

Paper 4

From:

Hansson, L. ; Holmberg, U. ; Brembeck, H. (Eds.)(2013).

Making sense of consumption. Selections from the 2nd Nordic Conference on Consumer Research 2012.

Göteborg: Centre for Consumer Science, University of Gothenburg. ISBN: 978-91-974642-6-0.

Available online at <http://hdl.handle.net/2077/34508>

Attitude vs. action for farm animal welfare: What can we learn from natural field experiments?

Alexander Schjøll, Frode Alfnes & Svein Ole Borgen

The discrepancy between consumers' attitudes and their actions is a well-known theoretical and methodological challenge in consumer studies (Grebitus et al. 2012). Surveys and focus groups can convincingly reveal consumers' stated preferences. However, consumers' stated preferences often turn out to correspond poorly with their actual behaviour in the market. The traditional methods in consumer research – like focus groups and surveys – measure attitudes, not actions. For ethical issues, such as fair trade, organic farming, sustainability, and animal welfare, the divergence between action and attitude is of special concern. Consumers may endorse high farm animal welfare standards, but nonetheless be unwilling to pay a premium price for farm animal welfare-friendly food. The purpose of this chapter is to discuss how natural field experiments can contribute to more valid and reliable studies of consumer behaviour. First, we will briefly discuss the strengths and weaknesses of the natural field experiment, as compared to methodological approaches that are delimited to mapping consumers' attitudes. Thereafter, we illustrate these characteristics of a natural field experiment by presenting results from an experiment conducted among consumers buying eggs in 2011. Finally, we present the learning content from our experiment.

Characteristics of natural field experiments

Non-market evaluations of market behaviour are plagued with biases. One example is the social desirability bias, as discussed by Fisher (1993). Animal welfare is an ethical issue where respondents are more likely to make ethically 'correct' choices in an observed setting than in daily life. This is one of several biases that can be expected when people know they are being observed. In general, people tend to behave differently in situations with monitoring than in similar situations without monitoring. A classic example is the experiment conducted at the Hawthorne Works in 1927, in which factory workers improved their efficiency in spite of worse working conditions (Landsberger 1958). Some participants might also want to please the researcher and give answers that are in line with what they think the researcher would like them to do, and not according to their own convictions. The infamous electroshock experiments undertaken by Milgram (1974) revealed how far people are willing to go in order to please the researcher. The very presence of an authority figure (such as the researcher) made people behave differently from what they would otherwise do. They tend to choose the path of least resistance and stick to conformity.

Natural field experiment is oriented towards revealing the respondents' true behaviour; what people actually do is of more interest than what they say they do. Such experiments are conducted in the normal setting of the decision of interests. The consumers are usually in situations that are well-known to them, like for instance buying everyday food in a store or dining in a restaurant. The participants make their choices as they are used to in their daily life without trying to please a researcher or being worried about what someone might think about or infer from their actions. Consumers do not deviate from their normal behaviour due to a very simple fact: They are unaware of their participation in the experiment.

Levitt and List (2009: 9) define a natural field experiment as an experiment where the subjects naturally undertake the tasks given and where the subjects do not know they are part of an experiment. More specifically, Harrison and List (2004: 1012) propose six factors that can be used to determine the field context of an experiment: (1) The nature of the subject pool, (2) the nature of the information that the subjects bring to the task, (3) the nature of the commodity, (4) the nature of the task or trading rules applied, (5) the nature of the stakes, and (6) the nature of the environment that the subjects operate in. The appealing feature of a natural field experiment is that no systematic difference is expected between participants' behaviour in the experiment and

their behaviour in the real world, due to the simple fact that the experiment is conducted in the real world, as perceived by the participants.

The flip side of this coin is that natural field experiments are conducted in a 'non-sterile' environment. The acting experimenter cannot control and 'freeze' everything, contrary to what the researcher can do in the laboratory. As a result, a significant amount of noise can occur. The researcher must try to identify these distracting factors prior to the experiment, and take them into account when designing the experiment and interpreting the results. In practice, a pragmatic approach is inevitable in order to adjust data to the actual contextual setting of the data collection. Some rules of thumb apply for this pragmatic approach. First, the experiment needs to be relatively simple, implying that the manipulation is introduced into the environment without changing other important characteristics of the environment. Thereby, it will be easy to assess the impact of the manipulation. Second, other exogenous factors that may influence the dependent variable should be controlled, or at least recorded, to the largest possible extent and controlled statistically. For instance, if sales of organic eggs are measured in one grocery chain and a competitor cuts his egg prices during the experiment, the results will be flawed unless one includes control stores. In some instances, the experimenter is in a position to identify these external shocks. In other cases, shocks cannot possibly be identified. The experimenter therefore needs to set up a control sample as well as using design methods, like e.g. orthogonal arrays or Latin squares, in order to control for the effect of external factors. The full factorial design with repetition of the first profile and use of control stores – which we used in the case to be presented below – illustrates the point. To mitigate most of the effect of various types of noise, strict follow-up of the experiment is mandatory. Natural field experiments can therefore be a rather time-consuming exercise for the researcher.

Design of the sample can represent another challenge in natural field experiments. It is difficult to screen participants without revealing that an experiment is going on. Without screening, the researcher has limited control of who will constitute the sample. By choosing the location (store or restaurant, for example) in which to conduct the experiment, the researcher indirectly also chooses a population from which the sample will be drawn. However, since the researcher does not control who decides to enter the store or restaurant on the days of the experiment, the sample design is not controlled by the researcher. Subsequently, the sample properties can differ from treatment to treatment. The internal validity can thereby be reduced. Alternatively, the

sample might differ from the target population, which can lead to diminished external validity.

Another challenge is that the sample of buyers might differ from the general population. This is not critical if we want to predict purchases, but is more problematic if we want to say something about the general population. In the case to be presented below, we studied purchasers' behaviour when buying eggs. Most Norwegians eat eggs, but we lack certain knowledge about who purchases eggs in store. Most likely, one or two persons in each household are responsible for purchasing everyday food. Hence, the population is not 'all Norwegians who eat eggs', but rather 'the people who buy eggs'. This distinction is not necessarily critical with respect to the interpretation of results, but needs to be taken into account in order to increase reliability.

The context of the experiment location must also be taken into account when consumers' perceptions are measured. Jervell and Borgen (2004) show that context matters when consumers evaluate food. With respect to field experiments, data collection takes place in the natural context for the activity. Consumers seem to think differently about potatoes when they are in the grocery store than when holding discussions in focus groups or responding to surveys. Buying potatoes directly from the farmer appears different from buying potatoes in a supermarket, even though the potatoes are identical. A focus group is a retrospective setting where the role of consumer is combined with other roles (e.g. citizen).

Experimental economists have introduced laboratory experiments with economic incentives to mimic real market behaviour. Introducing economic incentives can reduce some of the biases seen in surveys, such as the social desirability bias or the hypothetical willingness-to-pay bias, by making it costly to choose the socially 'correct' alternative. However, even laboratory experiments can suffer from the fact that the researcher observes the subjects' choices (Levitt and List 2007). Obviously, all methods in consumer research have inherent strengths and weaknesses that must be taken into account. Our contribution in the following is to discuss the potential and pitfalls of conducting natural field experiments, by extracting the basic learning content from an in-store study among purchasers of eggs.

Case study: Using a natural field experiment to reveal consumers' actual preferences

Background

Our empirical illustration is drawn from a natural field experiment conducted among egg buyers in two Norwegian grocery stores. Consumers in this market can select between three categories of eggs that are produced using different animal welfare quality standards (battery, indoor free-range, and organic).

Farm animal welfare is a much studied consumer issue, but most studies use focus groups or surveys to study this phenomenon. Consumers seldom reveal their real behaviour in relation to ethical issues such as animal welfare in a survey situation. A method that focuses on revealed preferences is therefore preferred as compared to methods that only capture stated preferences. The specific purpose of our natural field experiment was to reveal consumers' actual behaviour in relation to eggs from hens that benefit from better welfare than the standard production method. What did we learn? Let us start with a few remarks on the contextual background for the experiment. Studies of consumer attitudes can lead to the quick conclusion that a huge sales potential exists for organic products in European markets. Actual market sales do not support this assumption. On average, each European spends only €26 each year on organic food (Schaak 2010). Few Norwegians eat organic food regularly. The overall market share for organic food is only 1 per cent (Norwegian Agricultural Authority, SLF, 2013). The most successful organic products in Norway are baby food, soy and rice milk and eggs. Compared to Sweden and Denmark, organic food is a small niche in Norway. Approximately 162,000 hectares are used for organic farming in Denmark, and 480,000 hectares are used for the same purpose in Sweden. In comparison, only 56,000 hectares are used in Norway. In per cent of total agricultural hectares, this corresponds to 6.1 per cent in Denmark, 15.4 per cent in Sweden and 5.4 per cent in Norway (op cit.). Per capita consumption of organic foods in Denmark and Sweden are quadrupled and double the amount in Norway respectively (SLF 2013).

These measures serve as an informative reference point for our natural field experiment. Another reference point is the results from a survey about consumers' perception of farm animal welfare (Kjærnes and Lavik 2008). Figure 1 summarizes the results from this survey, consisting of a sample of 1,365 Norwegian respondents. This survey concluded that high animal

welfare is among the most important factors when consumers buy eggs. That the eggs are organically produced is reported by the informants to be equally important as low price.

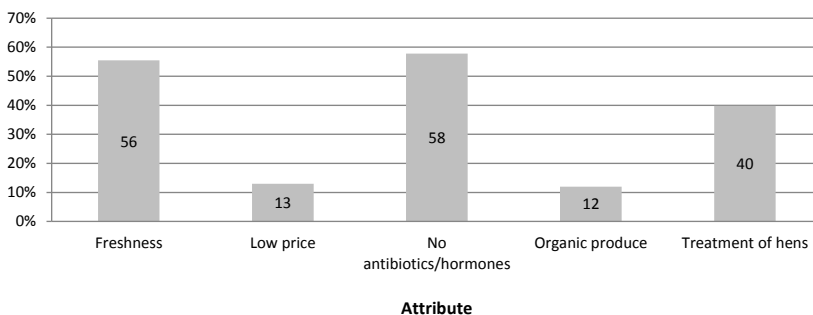


Figure 1: Percentage reporting 'very important' for different attributes when buying eggs among Norwegians. Source: Welfare Quality® survey. N = 1,365. The questions asked were 'When you buy eggs, how important is freshness, low price etc.'

An important motivator for our study is that consumers' revealed attitudes do not materialize in the market. The cheapest egg alternative (egg from caged hens) totally dominates. In 2012, eggs produced with the highest animal welfare standards had a market share of only 5.3 per cent (SLF 2013).

That actual market shares appear to be inconsistent with attitudes does not necessarily reflect that consumers ignore animal welfare as such. People may feel powerless (Kjærnes et al. 2009). Consumers who believe that their actions have little or no influence on animal welfare issues dissimulate their interest in this topic. A consumer cannot choose farm animal welfare-friendly products that are not available. According to a study by Roe and Marsden (2007), Norway has the smallest selection of farm animal welfare-friendly food among the seven European countries studied. However, when it comes to eggs, most stores have several levels of farm animal welfare. Consumers who are concerned about the standard of animal welfare in their country may optimistically believe that the welfare is improving because 'someone else' takes care of it, consistent with what Norwood and Lusk (2011) refer to as 'The California Egg Paradox'. In November 2008, 64 per cent of the voters in California voted for a ban on battery eggs. Alternatives to battery eggs have a very low market share in this state, so a discrepancy between attitude and action seems to exist among Californians. Norwood and Lusk claim that the major reason for the discrepancy is not free-riding or hypothetical answering,

which would be the normal reasoning in economics, but the different roles humans play. In a referendum, people act as citizens, which means they think in a more holistic manner. The same may hold in a survey, where consumers think more about their general views on society than their actual purchasing behaviour. As citizens, people act normatively. As shoppers, they think about what to eat, not about solving complex societal problems. Our study mapped what consumers see as important when they actually buy eggs, and came to conclusions that deviated significantly from the survey results. The survey and the experiment measured different roles and tasks ('responsible citizen' and 'feed the household'), and the roles are not very consistent.

What is the core point? Since people play different roles in different settings, the effect of the setting must be clearly comprehended by the researcher. The setting is an integral part of the analysis. The relevant setting when it comes to animal welfare may be either consumers' choices in stores, or the politicians' responsibility to demand sufficiently high farm animal welfare standards from food suppliers. Consumer research dealing with ethical issues should try to map people's viewpoints in their roles as both consumers and citizens.

Experiment

Our natural field experiment was conducted during five weeks in November-December 2011. We explored whether consumers choose animal friendly products when products are more clearly labelled as animal friendly. We manipulated the information given to egg buyers in two large grocery stores in Oslo, Norway. We exposed the participants to four levels of information: (1) No information given, (2) an A3 format poster hanging on the egg shelf stating that organic hens have higher animal welfare compared to free-range hens, (3) a sticker on organic egg cartons about animal welfare, and (4) both the poster and the sticker. A professional designer designed both the poster and the sticker. The poster read 'Organic eggs – Happier hens. REMA 1000 will stop selling battery eggs on January 1, 2012. If you want even better animal welfare, choose organic eggs'. The sticker had a drawing of a hen and read 'Improved animal welfare'. The sticker was also on the poster to strengthen the connection between the poster and the organic eggs. The poster and sticker are presented in figure 2 on next page.



Figure 2: Poster (left panel) and sticker (right panel) used to provide in-store information to consumers.

In the first and fifth weeks of the experiment, we did not manipulate the information, and we used these to control for the effect of the sticker and poster. We also had one control store to pick up changes in demand for the organic eggs unrelated to the experiment. The control store was chosen based on: (1) Belonging to the same chain, (2) the size of the store, and (3) similarities in market share for organic eggs. Compared with both monthly sales data for all the stores in the chain and the two test stores, the control store had significantly lower sales of battery eggs. This is most likely due to the location of the control store in a high-income part of Oslo with little variation in income.

During the experimental period, we observed the selection of eggs for all egg buyers in the test stores and the control store. The researchers went to the stores on a daily basis to make sure all alternatives were available to the customers. Sales data from the control store where no manipulation took place shows that there were no major changes in the composition of eggs purchased during the experiment period, and we conclude that there were no external shocks affecting the market share for organic eggs in the test period.

Results

Figure 3 shows the different market shares in our two test stores for the three production methods for eggs during the experimental period.

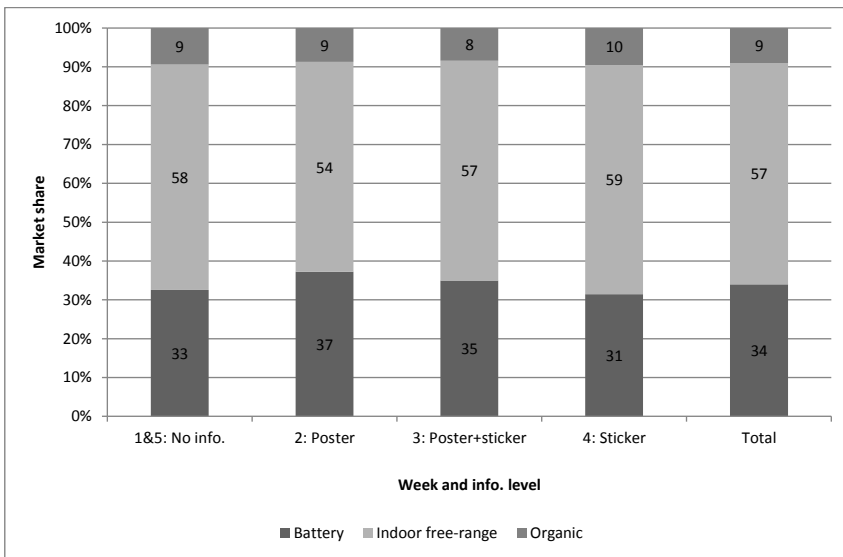


Figure 3: Percentage sales for different egg types in our two test grocery stores during five weeks of manipulation. N = 14,366.

Figure 3 shows no overall effect of our manipulation. The organic eggs have the highest market share when we gave no information about animal welfare. ‘Nudging’ of egg buyers does not seem to affect their decisions in a more animal-friendly direction.

Discussion of results

Organic eggs were the most expensive category of eggs on the shelf. It is not surprising that most customers chose one of the other alternatives even when informed or reminded about the animal welfare differences. The poster and sticker were probably not sufficient to overcome the price barrier. Based on these results, it is not easy to see any significant effect of in-store information about animal welfare on consumers’ choices. In an in-store setting, consumers do not seem to care very much about this credence quality, even though the stores communicate the message effectively. These core results from our natural field experiment appear to be much closer to the actual, aggregated market shares than the impression offered by attitude-based methods.

When visiting a store the consumer is bombarded with teasing information. Every product shouts 'buy me' and there is a lot of in-store marketing to attract consumers to purchase selected items. In the eyes of the buyers, our experiment was probably perceived of as just another in-store marketing trick. Nonetheless, the information was highly visible. For example, we placed the poster on the door that opens the fridge containing eggs. It was impossible for the egg buyers not to see it. We did not support our information campaign with an advertising campaign, which is normal thing to do when a new label is introduced. It is therefore possible that an animal welfare label introduced with an advertising campaign could have achieved a better result than our information campaign. The experiment can only say something about consumers' disposition to choose animal welfare-friendly food based on in-store marketing. The motivation behind this experiment was to see whether animal welfare is an issue that is already present in consumers' minds when they buy eggs. Just as the Welfare Quality® survey wanted to measure consumers' spontaneous *attitudes*, we wanted to measure their spontaneous *actions*. In that sense, a highly profiled advertisement campaign, supporting our treatment, would in fact harm the natural field experiment because the consumers would be more concerned about animal welfare in the first place. If they see advertisements promoting high animal welfare everywhere, they would most likely start to think about the issue and form an opinion about it.

In marketing, forced exposure is the most common way to study the effect of treatments (Nordfält 2007). In studies using forced exposure, the study objects are aware of the treatments they are exposed to, but they are unaware of the purpose of the study. When studying a natural situation, this technique may not be optimal. Incidental exposure would be a better solution, i.e. study objects that are not aware of the treatments taking place. Not only does this division of exposure have practical implications, it also influences the choice of the underlying psychological model (op. cit.). Using forced exposure, people know monitoring takes place, hence the brain focuses on the treatment. A placebo effect from the treatment itself may result, similar to the fact that the placebo effect in a medical experiment seems to treat complaints like headaches etc. In our experiment, we wanted to clarify whether information about hens' welfare could influence the market share of organic eggs. The previous survey (forced exposure) concluded that animal welfare was important for consumers. In this survey, respondents were forced to think about animal welfare. They knew that the purpose of the study was to measure interest in attributes relevant to eggs, including animal welfare. In

our natural field experiment, conducted in-store, consumers did not know it was a study. The poster and sticker that were used as treatments looked like any other in-store marketing material. Our natural field experiment was set up to deal practically with incidental exposure. In our experiment, consumers' brains did not focus on animal welfare. For consumers who already view animal welfare as an important choice parameter, the treatment might act as a reminder of buying animal welfare-friendly products. For most consumers however, animal welfare is not a latent issue, and no effect of the treatments could be observed.

For food marketers, the story often goes like this: Pre-tests of a new product show a huge potential for the new product. In focus groups, participants were enthusiastic about the new product and thought it tasted good. Surveys reveal many consumers who would like to eat it on different occasions. But when the product is launched, it does not sell and is withdrawn from the shelves before one year has passed. Obviously, something is wrong with the way marketers measure perceptions. The setting could be wrong. Consumers have trouble thinking how they may potentially use a new product. The same holds true for food products with an ethical dimension like animal welfare. Consumers say they want to buy it, but when they are in the store, the price is too high. They go for the cheapest alternative with the lowest animal welfare instead. In the store, no-one observes consumers as they do in a focus group. In addition, since consumers must actually pay for the food product in the store, the willingness to pay question is not hypothetical as it is in a survey. Consumer research methods that do not take into account the setting and bias problems are most likely to overestimate the interest in food products. Consumers have nothing to lose if they show too much interest, but they can lose face towards other consumers in focus groups or towards the researchers in surveys if they are too sceptical. This is especially likely to happen if the product deals with ethical concerns.

Validity is the extent to which a measuring instrument measures what the researchers intended it to measure. It is common to distinguish between internal and external validity. Internal validity is the ability of the design used to test the hypothesis the researchers intended to test. We believe our internal validity is high when it comes to natural field experiments. In this case study, we wanted to discern how consumers react to information about animal welfare in store when they select eggs. Of course, the most valid way to do this is in store because the experimental setting is identical to the setting we want to study.

External validity is the ability to extend the results beyond the limited research setting. External validity is thus about generalizing. The question is how generalizable our experimental results are. We used only two test stores and one control store, all located in Oslo. Oslo is the only metropolis in Norway, so it can be hard to extend the results to other parts of the country. However, we picked very large stores that attract many consumers each day. The stores we used probably have a much broader range than the average REMA 1000 store. In addition, this chain is the market leader in Norway. In this sense, it is possible to have high external validity by picking the most typical location for the natural field experiment. In the experiment, we used eggs to measure consumers' perceptions of animal welfare. Eggs are just one food product where animal welfare is important. The other is meat. Animals used for meat need to be slaughtered. This process can be painful for the animal. Egg-laying hens will be slaughtered as well, even though the meat is seldom used as human food in Norway. However, consumers probably do not have the slaughter process in mind when they think about welfare among hens. In this sense, we have low external validity since we did not choose a product that covers all aspects of the welfare of production animals. As a rule, one should choose a product to study in an experiment that covers all aspects of the field of study. If that is not possible one cannot generalize about other products. Hence, our study of consumers' perceptions of animal welfare for egg-laying hens is a study of this perception only.

The reliability of a measure is the measure's ability to produce similar results when one conducts repeated measurements under identical conditions. To our knowledge, our case study is the first in-store experimental study of animal welfare. It is therefore difficult to say whether we have obtained the same results as other researchers. We would expect the same result if we were to conduct our experiment again, since our method was not special, it was only carried out on a new topic. Market researchers manipulate in-store information every day in order to reveal the optimal in-store marketing. Their goal is to sell more products, which was identical to ours. We expect our reliability to be high because we used standard methods for measuring in-store marketing. However, repeating the same experiment with the same consumer sample would remind consumers about animal welfare. This learning effect would most likely influence them and affect sales. It is difficult to be more precise about the reliability of natural field experiments since the method is quite new in consumer research and not many studies have been repeated so far.

Concluding discussion of the method of natural field experiments

The question of selecting the proper method is essential in consumer research, since potential biases can harm the results. Traditional methods in consumer research – like surveys and focus groups – focus on consumers' attitudes, and can reveal consumers' stated purchase intentions or willingness to pay for specific product qualities. Nevertheless, these methods struggle to reveal consumers' real-life actions. The impact of complex ethical considerations on consumers' choices has proven difficult to conceptualize and measure properly. In many domains and situations, where a discrepancy exists between consumers' attitudes towards complicated ethical issues and their real-life actions, natural field experiment is a particularly useful option. Natural field experiments are conducted in people's normal contexts. Participants are unaware of being part of the experiment. The characteristics of this method have been discussed by extracting learning content from an in-store natural field experiment about consumers' willingness to buy products with enhanced animal welfare. Our case was selected because animal welfare is an increasingly important credence quality attribute in food markets. As compared to results from recent research on consumers' attitudes, the results from our experiment seem more in line with the actual, aggregated market demand. We claim that including a natural field experiment improves the external validity of consumer studies investigating credence attributes with an ethical dimension. Nonetheless, the inherent problems and pitfalls of natural field experiments must be properly understood in order to reap the full benefits from using the method.

A merit of natural field experiments is that participants make real choices in the context where they usually make such choices, and that the participants are unaware that an experiment is going on. Consumers' real behaviour is revealed in a valid way. The major problem of natural field experiment is to gather background information about the participants in order to understand the drivers of their choices. Natural field experiments have the opposite challenge compared to surveys and focus groups: Whereas traditional methods in consumer research (like surveys and focus groups) measure attitude and not action, natural field experiments measure action but not attitude. These problems can be overcome by interviewing the participants after they have made their actual choice. If interviewing does not influence further behaviour or the behaviour of other (potential) participants, the experiments can produce data based on both action and attitude. The natural field experiment is a useful method to complement more traditional methods in consumer

research like focus groups and surveys. Methodological triangulation is normally preferable, because there is no such thing as a perfect and unbiased consumer research method. The weakness of one method is often the strength of another. For all practical purposes, a reflective eclectic approach is required in consumer research, not only theoretically and empirically, but also methodologically.

References

- Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research*, 20, 303-315.
- Grebitus, C., Colson, G., & Menapace, L. (2012). A Comparison of Hypothetical Survey Rankings with Consumer Behavior. *Journal of Agricultural and Applied Economics*, 44, 35-47.
- Harrison, G. W., & List, J. A. (2004). Field Experiments. *Journal of Economic Literature*, 42, 1009–1045.
- Jervell, A. M., & Borgen, S. O. (2004). New marketing channels for food quality products in Norway. *Food Economics - Acta Agriculturae Scandinavica*, 1, 108-118.
- Kjærnes, U., & Lavik, R. (2008). Opinions on Animal Welfare and Food Consumption in Seven European Countries. In U. Kjærnes, B. B. Block, E. Roe, & J. Roex (Eds.), *Consumption, Distribution and Production of Farm Animal Welfare. Opinions and Practices within the Supply Chain. Welfare Quality Reports no. 7* (pp. 3–117). Cardiff: School of City and Regional Planning, Cardiff University. Retrieved from <http://www.welfarequality.net/everyone/42462/7/0/22>
- Kjærnes, U., Jacobsen E., & Lavik, R. (2009). Ansvarlige forbrukervalg som politikk. Eksempelet dyrevelferd. In K. Asdal & E. Jacobsen (Eds.), *Forbrukerens ansvar* (pp. 83-112). Oslo: Cappelen Akademisk Forlag.
- Landsberger, H. A. (1958). *Hawthorne revisited: management and the worker, its critics, and developments in human relations in industry*. Itacha, New York: Cornell University Press.
- Levitt, S. D., & List, J. A. (2007). What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? *Journal of Economic Perspectives*, 21, 153-174.
- Levitt, S. D., & List, J. A. (2009). Field Experiments in Economics: The Past, The Present, and The Future. *European Economic Review*, 53, 1-18.

- Milgram, S. (1974). *Obedience to Authority*. New York: Harper & Row.
- Nordfält, J. (2007). *Marknadsföring i butik. Om forskning och branschkunskap i detaljhandeln*. Malmö: Liber.
- Norwegian Agricultural Authority (SLF) (2010, March 16). *Produksjon og omsetning av økologiske landbruksvarer – rapport for 2009*. Rapport nr. 7/2010. Retrieved from https://www.slf.dep.no/no/kampanje/_attachment/165?_ts=12824833450&download=true
- Norwegian Agricultural Authority (SLF) (2013, March 15). *Produksjon og omsetning av økologiske landbruksvarer. Rapport for 2012*. Rapport nr. 12/2013. Retrieved from https://www.slf.dep.no/no/miljo-og-okologisk/okologisk-landbruk/om-okologisk-landbruk/_attachment/28778?_ts=13d8c2bb720&download=true
- Norwood, F. B., & Lusk, J. L. (2011). *Compassion, by the Pound. The economics of farm animal welfare*. New York: Oxford University Press.
- Roe, E., & Marsden, T. (2007). Analysis of the Retail Survey of Products that Carry Welfare-claims and Non-retailer Led Assurance Schemes whose Logos Accompany Welfare-claims. In U. Kjærnes, M. Miele, & J. Roex (Eds.), *Attitudes of Consumers, Retailers and Producers to Farm Animal Welfare. Welfare Quality Reports no. 2*. (pp. 6-65). Cardiff: School of City and Regional Planning, Cardiff University. Retrieved from <http://www.welfarequality.net/downloadattachment/37097/20179/WQR2.pdf>
- Schaak, D. (2010, February 20). *Organic area and sales in Europe 2008*. Presentation held at BioFach Nuremberg. Retrieved from <http://www.organic-world.net/fileadmin/documents/Common-contents/publications/schaack-2010-biofach.pdf>