2013 | Faculteit Bedrijfseconomische wetenschappen

DOCTORAATSPROEFSCHRIFT

Valuing the external effects of soil contamination: An empirical analysis of a polluted area in Belgium

Proefschrift voorgelegd tot het behalen van de graad van doctor
in de toegepaste economische wetenschappen, te verdedigen door

Eloi Schreurs

Promotor: prof. dr. Theo Thewys

D/2013/2451/65
ISBN-NUMMER : 9789089130266

Woorden van dank

Een niet-kwantificeerbare hoeveelheid kennis, ervaring en fijne tijden,
471376 karakters inclusief spaties,
400988 karakters exclusief spaties,
73298 waardevolle woorden,
54564 documenten,
3466 mappen,
1507 zonsondergangen,
68 Word-documenten met overbodige tekst,
7 hoofdstukken,
4 verjaardagen,
3 buitenlandse congressen,
2 grote onderzoeksthema's,
en 1 doctoraat later is er een einde gekomen aan mijn jaren als doctoraatsonderzoeker aan de Universiteit Hasselt. Ik heb er ongelofelijk veel bijgeleerd, veel toffe mensen leren kennen en fijne ervaringen opgedaan. Ik weet nu al zeker dat ik er later nostalgisch op terug zal kijken met het gevoel dat het "toen toch allemaal veel beter was". Maar dit is dus (voorlopig) het eindpunt en hopelijk ook het toppunt van mijn periode aan de universiteit. Daarom wil ik deze opportuniteit aanwenden om enkele mensen te bedanken zonder wie dit allemaal niet mogelijk was geweest.

Eerst en vooral denk ik hierbij aan mijn promotor, prof. dr. Theo Thewys. Ook al raadde je me regelmatig aan niet té perfectionistisch te zijn, toch heeft vooral uw neus voor details ervoor gezorgd dat alles altijd tot in de puntjes nagelezen en verbeterd werd. Ik wil daarom mijn dankbaarheid uiten voor de goede verstandhouding en de vrijheid die ik heb mogen genieten de laatste vier jaren. Ik hoop tegelijkertijd dat je trots bent op het onderzoekswerk dat ik onder uw hoede heb verricht. En wie weet waar we elkaar nog tegenkomen later…

Steven, je maakte me wegwijs in de academische wereld van tijdschriften, publicaties en congressen. Voornamelijk de hoogst gerankte/meest

kwaliteitsvolle dan natuurlijk :-) Het valt moeilijk te omschrijven hoeveel je me hebt bijgebracht, ook buiten het puur academische aspect. Als ik de helft van al je tips heb onthouden, zal ik al een pak wijzer en strategischer zijn dan ervoor. Blijf vooral doen wat je doet, even enthousiast en begripvol als altijd, en het zal er zekèr niet slechter op worden in de onderzoeksgroep milieueconomie.

De twee andere leden van mijn doctoraatscommissie, Irina en Jaco, ben ik dankbaar voor de steun en om mij de juiste richting uit te sturen tijdens mijn doctoraat. Irina, ik vrees dat ik af en toe vragen heb gesteld die vanzelfsprekend en mogelijk wat belachelijk kunnen overkomen voor een gezondheidseconome. Bedankt voor de steevast wijze raad en om mij te inspireren om dieper na te denken over niet voor de hand liggende thema's. Jaco, bedankt om mij te gidsen in de wondere wereld van fytotechnologieën.

Ook de overige juryleden, Tim, Liesbet en prof. Maddison, wil ik graag bedanken voor de waardevolle opmerkingen, ieder in jullie eigen vakgebied. Ik ben ervan overtuigd dat, moest het meetbaar zijn, er een significant verschil in kwaliteit te merken zou zijn tussen de eerste en de finale versie van mijn doctoraat. Met een positief teken welteverstaan.
Furthermore, I would like to thank the other members of the jury, Tim, Liesbet and prof. Maddison, for their valuable comments, each in their own research domain. I am convinced that, in case it would be quantifiable, a significant difference in quality can be found between the first and the final version of my PhD. With a positive sign, of course.

Daarnaast zijn er enkele organisaties die mij geholpen hebben bij de uitwerking van mijn doctoraat. Ik richt me dan in eerste instantie op de personen binnen deze organisaties aan wie ik veel te danken heb. Voor het economische gedeelte van de gezondheidshoofdstukken rekende ik op het IMA, en niet in het minst op Ragna Préal en Bram Peters. Het was een lange, soms vermoeiende weg naar een correcte analyse, maar ik denk dat we uiteindelijk fier mogen zijn op het resultaat. Bedankt om mijn vragen en ook de daarop volgende vragen met zo veel geduld en zorgvuldigheid aan te pakken.

Verder wil ik de longartsen Koenraad Demuynck en Luc Spaas (Virga Jesse Ziekenhuis, Hasselt) en vooral Michiel Thomeer (Ziekenhuis Oost-Limburg, Genk) graag bedanken voor de noodzakelijke hulp bij het afbakenen van de sample longkankerpatiënten en mij wegwijs te maken in de zeer uitgebreide lijsten van nomenclatuurnummers en ATC-codes. De onderzoekers van de Bone and Cartilage Research Unit van Université de Liège, en dan met name Mickaël Hiligsmann, hebben me dan weer goed op weg geholpen bij alle materie betreffende osteoporotische fracturen.

Voor de hedonistische prijsanalyse wil ik graag de administratie van het Kadaster bedanken voor de goodwill met betrekking tot de verschaffing van landbouwgrondprijzen in de Kempen. Speciale vermelding hierbij gaat uit naar Rob Hoogmartens, vroeger één van mijn thesisstudenten, nu collega in de onderzoeksgroep. Zonder jou was het mij waarschijnlijk niet gelukt om dit hoofdstuk toe te voegen aan mijn doctoraat, bedankt voor je volharding in de zoektocht naar landbouwgegevens. Stan Forier en Johan Laeremens van de Vlaamse Landmaatschappij (VLM) ben ik erkentelijk voor de informatie omtrent de landbouwgevoeligheidsanalyse van VLM, die mij in staat heeft gesteld het hedonistisch model te perfectioneren.

Aangezien de tijd om het laatste hoofdstuk uit te voeren vrij beperkt was, ben ik ook hier dankbaarheid verschuldigd aan verschillende personen. Eerst en vooral Sebastien natuurlijk, voor mijn eindeloze reeks vragen over keuze-experimenten te beantwoorden. Zonder jou had ik het zeker niet binnen het vooropgestelde tijdsbestek klaargespeeld. Daarnaast ook een welgemeende dank u aan Koen Vanheukelom en Liesbeth Franssen van Boerenbond Limburg, voor jullie kennis en expertise ter beschikking te stellen en de vlotte samenwerking tijdens dit onderzoek. Thesisstudenten Dries Smeets en Lene Cillen hebben mee geholpen bij de dataverzameling. Gelukkig was de landbouwpopulatie in de Limburgse Kempen heel bereidwillig om in te gaan op onze vragen, wat het ons zeker heeft vergemakkelijkt om onze doelstellingen te bereiken. Bedankt, beste landbouwers.

Ook een big up aan de collega PhD-studenten en postdocs over de grenzen van de vakgebieden en faculteiten heen, maar natuurlijk voornamelijk in de onderzoeksgroep milieueconomie. In alfabetische volgorde, om zeker niemand voor te trekken of achter te stellen: Dries, Ellen, Frederic, Janka, Miet, Nele, Rob, Sarah, Sebastien, Silvie, Tine, Tom en Yann, bedankt voor de gezellige middagpauzes en momenten van afleiding. Ik ben er zeker van dat de groep nog mooie tijden tegemoet gaat en hoop jullie allemaal nog terug te zien, zowel binnen als buiten het professionele leven. Een speciale vermelding hierbij gaat uit naar ex-collega Thomas, die vaak als persoonlijke hulplijn heeft gediend bij de GIS-analyses die de basis vormen van een groot deel van mijn onderzoek. Bedankt, Thomas.

Zoals de traditie het voorschrijft, wordt het dankwoord afgesloten met de personen die de grootste algemene invloed hebben gehad. Ouders, bedankt voor de steun tijdens de laatste vier jaren (en natuurlijk ook de 23 voorgaande jaren). Als het niet op het inhoudelijke vlak van mijn doctoraat was, dan wel op alle andere gebieden waar ouders belangrijk in zouden moeten zijn. Misschien het meest geschikte zinnetje hierbij: 'Ik heb het toch getroffen he.' Wat qua overgangszin ideaal uitkomt, omdat het ook van toepassing is op de laatste persoon in het rijtje. Dexy, ook al vind je zelf dat je geen bijdrage hebt geleverd aan dit doctoraat, ik vind zelf dat je de grootste bijdrage van iedereen hebt geleverd. Door er te zijn, door mij te laten lachen, door samen fijne dingen te doen, door mijn gedachten te verzetten. Bedankt om al meer dan 8 jaar mijn lief, mijn beste vriendin en mijn toeverlaat te zijn. Het is al een fantastische tijd geweest met u, maar toch weet ik dat het in de toekomst alleen nog maar beter gaat worden. De rest van mijn speech zal ik bewaren voor onze trouw :-)

Samenvatting

De natuurlijke omgeving heeft het vermogen om tot een bepaalde hoogte afvalstromen te absorberen, te transformeren en op te slaan. Menselijke activiteiten hebben echter vaak te veel beroep gedaan op deze dienst, wat heeft geleid tot milieudegradatie en –vervuiling. Hoewel er op een aantal vlakken vooruitgang is geboekt sinds de erkenning van de internationale gemeenschap dat het milieu geïntegreerd moet worden in economische doelstellingen, blijft milieuvervuiling een aanzienlijk probleem in veel landen over de hele wereld. Milieueconomisch onderzoek richt er zich daarom op de schaarse milieugoederen en economische middelen optimaal te alloceren met het oog op het bekomen van een zo hoog mogelijk welvaartsniveau. Op deze manier zou men een beter begrip moeten krijgen van de effectieve waarde die het natuurlijke systeem toevoegt aan de maatschappij.

Deze thesis focust zich op het achterhalen van de economische impact van langdurende milieuverstoringen op het maatschappelijke niveau. We hebben hierbij de mogelijkheid om verder te werken op een geval van milieuvervuiling dat het onderwerp is geweest van verschillende interdisciplinaire onderzoeksprojecten in de laatste decennia, namelijk de Belgische Kempen. De metallurgische industrie in de regio heeft er een uitgebreid gebied in België en Nederland diffuus verontreinigd met zware metalen, voornamelijk cadmium (Cd). Hoewel de uitstoot van vervuilende elementen is beëindigd in de jaren ’80, zijn er op verschillende plaatsen in de regio nog steeds bodemconcentraties aan Cd die hoger zijn dan toegestaan. Dit doet de vraag rijzen of de vervuiling nog steeds een probleem vormt vanuit een economisch perspectief. Deze thesis combineert daarom elementen van biologisch, epidemiologisch en gezondheids-economisch onderzoek om met behulp van waarderingstechnieken voor niet-vermarktbare goederen te komen tot een inschatting van de economische schade die is ontstaan als gevolg van de milieuvervuiling.

Voordat het economisch onderzoek wordt opgestart, is het belangrijk om eerst een goed overzicht te krijgen van de reikwijdte van de vervuiling. In **hoofdstuk 2** worden er daarom geostatistische voorspellingstechnieken gehanteerd om voorspellingskaarten te creëren van de bodemvervuiling in de regio. Dit wordt gedaan met behulp van een databank van bodemstalen die zijn afgenomen in eerdere onderzoeksprojecten. De resulterende kaarten geven aan dat, hoewel de vervuiling een uitgebreid gebied bestrijkt, de vervuilingsniveaus in het algemeen slechts in geringe mate boven de bodemnormen uitstijgen. Alleen in de directe omgeving van de zinksmelters worden er (zeer) hoge Cd concentraties opgetekend. Tegelijkertijd laten de voorspellingskaarten ons toe om te ontdekken hoeveel landbouwgebied in de regio vervuild is. Hieruit blijkt dat ongeveer 3000 hectare landbouwgrond te maken heeft met bodemconcentraties aan Cd die de streefwaarde van de Vlaamse overheid overschrijden. Meer dan 1200 hectare kan als vervuild worden beschouwd omdat de bodemsaneringsnorm voor landbouwgrond overschreden wordt. Aan de positieve kant van dit verhaal komt wel naar voren dat 60% en 80% van de vervuilde percelen kan teruggebracht worden tot onder de streefwaarde binnen respectievelijk 21 en 42 jaren door middel van fytoremediatie.

In de twee volgende hoofdstukken ligt de focus op gezondheidseconomische gevolgen van de milieuvervuiling. In **hoofdstuk 3** is de doelstelling het aantal ziektes te bepalen dat toe te schrijven is aan de milieuvervuiling. De schadefunctie benadering werd geselecteerd als de meest geschikte methode om dit doel te bereiken. Dit onderzoek is gebaseerd op het eerdere epidemiologische onderzoek in die regio dat uitwees dat mensen die blootgesteld zijn aan verhoogde omgevingswaardes aan Cd een verhoogd risico hebben om osteoporotische breuken of longkanker op te lopen. Deze relatieve risico's worden geëxtrapoleerd naar het populatieniveau in een evaluatie van gezondheidsrisico's. Als er wordt vanuit gegaan dat de relatieve risico's constant zijn gebleven doorheen de jaren, zijn er ieder jaar ongeveer 22 gevallen van longkanker, 8.5 heupbreuken in mannen en 32 in vrouwen toe te schrijven aan de vervuiling in de Kempen. Maar aangezien het vrij aannemelijk is dat de relatieve risico's gedaald zijn sinds de afsluiting van de epidemiologische studies, kan het verwacht worden dat deze aantallen in de toekomst verder zullen dalen.

Het economische aspect van deze ziektes wordt nader bekeken in **hoofdstuk 4** zodat de gezondheidseconomische schade als gevolg van de milieuvervuiling geschat kan worden. In deze context ligt de focus op de effectieve kostenstromen die voortkomen uit deze aandoeningen, niet op de totale bereidheid om te betalen om de ziektes te vermijden. Aangezien de voornaamste medische kosten van heupbreuken in België recent zijn geschat door onderzoekers aan de Universiteit van Luik, worden in dit hoofdstuk enkel de directe medische kosten van longkanker in beschouwing genomen. Deze kosten worden afgeleid uit de permanente steekproef van het Intermutualistisch Agentschap door middel van een longitudinaal gematched case-controle ontwerp. Deze databank omvat de uitgaven aan de gezondheidszorg van een representatief staal van 300.000 Belgische burgers. De resultaten van de analyse tonen aan dat longkankerpatiënten tussen €25,000 en €30,000 aan incrementele kosten oplopen ten gevolge van de ziekte. Als deze schattingen gecombineerd worden met de resultaten van het vorige hoofdstuk, wordt er gevonden dat er ieder jaar meer dan €500,000 in longkankerkosten en €400,000 heupbreukkosten toe te schrijven zijn aan de milieuvervuiling.

Overgaand naar het tweede grote onderzoeksthema geven de resultaten van hoofdstuk 2 aanleiding om dieper in te gaan op de economische gevolgen voor landbouwers met vervuilde percelen. Daarom wordt het effect van bodemvervuiling op de waarde van landbouwgrond onderzocht in **hoofdstuk 5**. De hedonistische prijsmethode is een gereveleerde voorkeursmethode die zich baseert op de veronderstelling dat de prijs van gedifferentieerde producten zoals landbouwgrond bestaat uit de waarde die ieder kenmerk toevoegt aan het product. Op deze manier kan de marginale impliciete prijs voor ieder productkenmerk afgeleid worden, meer bepaald bodemvervuiling in dit geval. Er wordt een hedonistisch model opgesteld dat gebruik maakt van de verkoopprijzen van landbouwgrond in de regio en een vector van onafhankelijke variabelen waarvan verwacht wordt dat ze de prijszetting van landbouwgrond beïnvloeden. De regressieresultaten wijzen uit dat Cd concentraties in de bodem geen significante invloedsfactor zijn voor prijzen van landbouwgrond. Aangezien er verschillende mogelijkheden zijn om dit resultaat te verklaren, wordt een nieuw onderzoek opgestart en beschreven in het volgende hoofdstuk waarbij op zoek wordt gegaan naar de bevestiging van sommige hypotheses.

In **hoofdstuk 6** is de doelstelling om aan de hand van een discreet keuze-experiment de voorkeuren van landbouwers te ontdekken bij aankoopbeslissingen van vervuilde landbouwpercelen. Deze uitgedrukte voorkeursmethode laat toe om de voorkeuren van consumenten te ontdekken door hen de keuze te laten het meest waardevolle alternatief te selecteren uit keuzesets die bestaan uit verschillende (hypothetische) alternatieven. Deze alternatieven worden beschreven aan de hand van verschillende attributen met variërende attribuutniveaus. De data voor dit onderzoek worden verzameld door middel van persoonlijke interviews met 200 landbouwers in de regio die allemaal lid zijn van de grootste landbouworganisatie in Vlaanderen. Een multinomiaal logit model wordt geschat dat aangeeft dat de aanwezigheid van bodemvervuiling en de bijhorende teeltbeperkingen het nut van landbouwgrond significant verlaagt in vergelijking met landbouwgrond zonder gebruiksbeperkingen. Wanneer er echter interactie-effecten worden geïntroduceerd, wordt er gevonden dat gespecialiseerde melkveehouders, de belangrijkste landbouwactiviteit in de regio, blijkbaar significant minder aandacht hebben voor de gebruiksbeperkingen die worden opgelegd door bodemvervuiling.

Als de resultaten van de twee laatste hoofdstukken worden gecombineerd kan de situatie met betrekking tot vervuilde landbouwgrond verduidelijkt worden. Een aantal mogelijke verklaringen voor het gebrek aan prijsdepreciatie in vervuilde landbouwgronden worden naar voren geschoven. Ten eerste blijkt dat wanneer landbouwers expliciet geïnformeerd worden over de bodemvervuiling en de resulterende gebruiksbeperkingen effectief worden toegepast, de gemiddelde landbouwer minder waarschijnlijk een vervuild perceel zal aankopen. Dit kan een indicatie zijn dat kopers van landbouwgrond momenteel niet over de noodzakelijke informatie beschikken bij transacties van landbouwgrond. Daarnaast is het mogelijk dat melkveehouders vervuilde percelen niet als ‘beperkt’ beschouwen aangezien er geen beperkingen zijn met de betrekking tot de teelt van voedergewassen of het gebruik als weiland. Daarom is hun bereidheid om te betalen voor deze percelen mogelijk ook niet beïnvloed, waardoor het gebrek aan een negatief effect in landbouwgrondprijzen verklaard kan worden. Ten derde geven de resultaten van het discreet keuze-experiment aan dat landbouwers voornamelijk bezorgd zijn over hun mestafzetcapaciteit bij het maken van grondaankoopbeslissingen. Landbouwers zijn daardoor vooral geïnteresseerd in land zonder bemestingsbeperkingen, waardoor de potentiële risico’s van de vervuiling eerder genegeerd worden. De schaarste aan (landbouw)grond in Vlaanderen versterkt dit argument enkel maar.

Als conclusie kunnen we stellen dat de milieuvervuiling in de Kempen een langdurende last is die tot op de dag van vandaag een rol speelt. Meer dan 25 jaar nadat de zinksmelters gestopt zijn met het uitstoten van zware metalen wordt de regio nog steeds geconfronteerd met ziektes die te wijten zijn aan de vervuiling, terwijl landbouwers moeten omgaan met de beperkingen die opgelegd worden door de vervuiling van hun gronden. Dit voorbeeldgeval zou overheden in de wereld, en dan vooral in industrialiserende landen, ertoe moeten aanzetten om milieuaspecten in rekening te brengen bij het ontwerpen en uitstippelen van een economisch beleid. De implementatie van een strikte milieuwetgeving, gecombineerd met een van nabij opgevolgd monitoringssysteem kan ervoor zorgen dat vervuiling vermeden wordt en deze lasten niet gedragen moeten worden door toekomstige generaties. Hoewel deze strategie relatief hoge opstartkosten vereist, past het beter binnen het vermogen van de natuurlijke omgeving om afvalstromen te absorberen en wordt het milieu hierdoor beter beschermd tegen verstoringen. Het heeft bovendien het potentieel om de (externe) kosten op lange termijn te verminderen en mogelijk te elimineren, waardoor het uiteindelijk ook economisch voordeliger kan uitdraaien.

Summary

The environment has the capacity to absorb, transform and store waste products to some extent. However, human activities have often overextended this service, leading to environmental degradation and pollution. Although some progress has been made on a number of aspects since the global community has acknowledged the need to integrate the environment into the economic development agenda, environmental pollution continues to be a substantial problem in countries all across the world. In this setting, environmental economic research aims to determine the optimal allocation of environmental and economic resources in view of achieving a maximum amount of welfare. This way, the aspiration is to obtain a better understanding of the true value resource-environmental systems provide to society.

This dissertation focuses on finding out to what extent long-lasting environmental disturbances have an economic impact on a societal level. Conveniently, we have the opportunity to advance on a case of environmental pollution that has been the subject of many interdisciplinary research projects in the last couple of decades, i.e. the Campine region in Belgium. The metallurgic industry in the area has diffusely contaminated an extensive area in Belgium and the Netherlands with various heavy metals, particularly cadmium (Cd). Although the zinc smelters have stopped emitting anomalous elements in the 1980s, soil Cd concentrations remain higher than allowed in a number of places throughout the area. This raises the question whether the contamination still presents a problem from an economic point of view. This dissertation combines elements from biological, epidemiologic and health economic research in order to value the economic damage as a result of environmental pollution using nonmarket valuation techniques.

Before the economic research is initiated, it is important to first get a better view on the extent of the pollution. In **chapter 2** geostatistical prediction techniques are employed to create prediction maps for the soil

contamination in the area using a database of soil sample measurements that have been collected in former research projects. The resulting maps indicate that, although the contamination covers an extensive area, the contamination levels are generally just slightly elevated in reference to soil thresholds. Only in the direct vicinity of the zinc smelters (very) high Cd concentrations are registered. At the same time, the prediction maps allowed us to find out how much farmland in the area has been affected by soil contamination. The results show that soil Cd levels in approximately 3000 hectares of agricultural land exceed guide values set by the Flemish government. More than 1200 hectares can considered to be contaminated because they exceed soil thresholds. On the upside of these results, it appears that 60% and 80% of the contaminated parcels can be brought back to below guide values in respectively 21 and 42 years using phytoremediation.

In the following two chapters the focus is on the health economic consequences of the environmental pollution. In **chapter 3**, the objective is to establish the amount of illnesses which are attributable to environmental pollution. The damage function approach is selected as the most appropriate methodology for achieving this goal. This research is based on the former epidemiologic research which identified the extra risk of incurring osteoporotic fractures and lung cancer when people are exposed to elevated environmental Cd levels. These relative risks are extrapolated to the population level in a health risk assessment. Assuming the relative risks have remained constant throughout the years, there are each year approximately 22 cases of lung cancer, 8.5 hip fractures in men and 32 in women attributable to the pollution in the Campine region. However, since it is quite plausible that the relative risks have decreased ever since the epidemiologic studies have been finalized, it can be expected that the pollution-attributable illnesses will decrease further in the future.

The economic consequences of the pollution-attributable illnesses are estimated in **chapter 4** in order to asses to health economic damage as a result of the environmental pollution. In this setting, the focus is on the resource costs produced by these illnesses, not on the complete willingness to pay for avoiding them. Since the predominant medical costs of hip fractures in Belgium have recently been estimated by researchers at the University of Liège, solely the direct medical costs of lung cancer are explored in this chapter. Lung cancer costs are extracted from the Intermutualistic Agency's permanent sample, a database comprising health care expenditures of a representative sample of 300,000 Belgian citizens, using a longitudinal matched case-control design. The results of the analysis show that lung cancer patients incur incremental costs in the range of €25,000 - €30,000 per patient due to the illness. If these estimates are combined with the results from the previous chapter, each year more than €500,000 in lung cancer costs and €400,000 in hip fracture costs can be attributed to the environmental pollution.

Transferring to the second major research theme, the results in chapter 2 provide a rationale for proceeding on the economic consequences for farmers with contaminated parcels. Therefore, the impact of soil contamination on farmland values is investigated in **chapter 5**. Hedonic pricing analysis is a revealed preference technique that is based on the assumption that prices of differentiated products such as farmland consist of the value that each characteristic appends to the product. This way, the marginal implicit price of each product characteristic, i.e. soil contamination in this case, can be derived. A hedonic model is created using a sample of farmland sales in the area and a vector of independent variables that are expected to affect farmland price determination. The regression results indicate that soil Cd levels are insignificant determinants for farmland prices. Since there are a number of hypotheses apt for explaining this result, another research is initiated and described in the next chapter in search of confirmation for some of these hypotheses.

In **chapter 6**, the objective is to investigate farmer's preferences in land purchase decisions in case of soil contamination by introducing a discrete choice experiment. This stated preference technique allows for a revelation of consumer preferences by giving respondents the opportunity to select the most valuable alternative from choice sets consisting of multiple (hypothetical) alternatives. These alternatives are described by different attributes with varying attribute levels. Data for this research are obtained from in-person interviews with 200 farmers, all members of the largest farmer association in Flanders. A multinomial logit model is estimated, showing that the presence of soil contamination and the resulting crop restrictions significantly reduces farmland utility in comparison with farmland without land use restrictions. However, introducing interaction effects in the model points out that specialist dairy farmers, the largest farming type in the area, seem to pay significantly less attention to the land use restrictions imposed by soil contamination.

Combining the results from the two final chapters can clarify the situation with regard to soil contamination in farmland. A number of potential explanations for the lack of price depreciation in contaminated farmland parcels are put forward. Firstly, in case farmers are explicitly informed on the soil contamination and the resulting land use restrictions are effectively applied, the average farmer is unlikely to purchase contaminated parcels. This might be an indication that farmland buyers are currently lacking the necessary information in farmland transactions. Secondly, since there are no restrictions with regard to the production of fodder crops and usage as pasture in contaminated parcels, dairy farmers might not experience these parcels as 'restricted'. Hence, their willingness to pay for contaminated farmland might be unaffected, which can explain the lack of depreciating effect in farmland prices as well. Thirdly, the results from the discrete choice experiment suggest that farmers are mainly concerned about their manure disposal capacity when making land purchase decisions. Therefore, farmers might be particularly interested in land without fertilizing restrictions, encouraging them to ignore potential contamination risks. Land scarcity in Flanders merely reinforces this argument.

In conclusion, the environmental pollution in the Campine region has generated a long-term burden, which is still at play up to this day. More than 25 years after the zinc smelters have stopped emitting heavy metals, the population is still facing illnesses which are attributable to the pollution, while farmers have to deal with the constraints imposed by the contamination of their land. This exemplary case must be a strong incentive for governments all across the world, but especially in industrializing countries, to take environmental aspects into account in designing and carrying out economic policies. The implementation of stringent environmental regulations combined with a closely controlled monitoring system can prevent the pollution from occurring and save future generations from bearing this burden. Although this strategy calls for relatively high start-up costs, it is more likely to fit the environment's capacity to absorb waste products and a better way to protect the environment from disturbances. Moreover, it has the potential to reduce or possibly eliminate the long-term (external) costs associated with environmental pollution, which might turn out to be also economically more beneficial in the end.

Table of contents

List of Tables

List of Figures

List of abbreviations

AF	attributable fraction
ALV	Flemish Agency for Agriculture and Fishery
ASC	alternative specific constant
ASE	average standard error
ATC	anatomical therapeutic chemical
BCR	Belgian Cancer Registry
BFM	budget of financial means
BONK	Exposure Research Nothern Campine
CAP	Common Agricultural Policy
CBA	cost-benefit analysis
Cd	cadmium
CE	choice experiment
CI	confidence interval
cm	centimeter
COI	Cost Of Illness
COPD	chronic obstructive pulmonary disease
CPI	consumer price index
CRAB	central reference address file
CV	contingent valuation
DCE	discrete choice experiment
DFA	damage function approach
€	euro

EPS	permanent sample (of IMA's database)
ESDA	exploratory spatial data analysis
FFS	fee for service
FPS	Federal Public Service
g	gram
GDP	gross domestic product
GIR	ground information register
GIS	geographic information system
GP	general practitioner
ha	hectare
HEA	high exposure area
HPA	hedonic pricing analysis
ICD	International Classification of Diseases
i.e.	id est, that is
IIA	independence of irrelevant alternatives
IID	independently and identically distributed
IMA	Intermutualistic Agency
km	kilometer
£	pound sterling
LEA	low exposure area
LR	likelihood ratio
m	meter
MAC	marginal abatement cost
MDC	marginal damage cost

MC_P	marginal private cost
MC_S	marginal social cost
ME	mean error
Mg	megagram
min	minute
MNL	multinomial logit
MOC	multidisciplinary oncologic consult
MSE	mean standardized error
n	amount of observations
NSCLC	non-small cell lung cancer
OLS	ordinary least squares
OVAM	Public Waste Agency of Flanders
PA	pollution-attributable
ppm	parts per million
QALY	quality-adjusted life year
QR	quantile regression
RIZIV-INAMI	National Institute for Health and Disability Insurance
RMSE	root-mean-square error
RMSSE	root-mean-square standardized error
RP	revealed preference
RPL	random parameter logit
RR	relative risk
RUT	random utility theory
$	dollar

SAR	spatial autoregressive model
SCLC	small cell lung cancer
SEM	spatial error model
SP	stated preference
T-M-M	tobacco-maize-maize (rotation)
TCPS	technical commission of the permanent sample (IMA)
US	United States (of America)
VLM	Flemish Land Agency
VSL	value of a statistical life
WTP	willingness to pay
y	year
Zn	zinc

1 Introduction

1.1 Environmental and soil pollution

Natural resource systems yield valuable goods and services that contribute to human well-being in numerous ways. The wide range of environmental service flows is generally broken down into four services to society (Freeman, 2003). In the first place, the natural environment is a source of raw materials, fuels and fibers that can be used as input factors in various production processes. Moreover, it offers a range of regulating services as well. For example, forests control surface runoff due to erosion and improve air quality, thereby reducing the number of illnesses in the population. The amenity services provided by the environment are the third category of services. Although an increased level of biodiversity and scenic views related to environmental landscapes are not consumed directly, these services can provide utility and welfare to society and future generations. Lastly, the environment has the capacity to absorb shocks presented by human activities and serve as a sink for the dispersion, transformation and storage of waste products (Freeman et al., 1973).

However, in some cases human activities have overextended the latter service and discharged waste products beyond the absorbing capacity of the environment. Consciousness with regard to the importance of these environmental goods and services among governments and policy makers has led to an increased attention for protecting the environment from disturbances in recent decades. The United Nations Conference on Environment and Development, held in 1992 in Rio de Janeiro (Brasil), is often considered to be the first conference in which the global community explicitly acknowledged the need to integrate the environment into economic development objectives. More than 20 years later there still is a long way to go towards a better understanding of the true value resource-environmental systems provide to society. In this setting, an important role is played by environmental research for achieving this objective. Environmental research needs to draw from a multi- and interdisciplinary approach in order to captivate all aspects from the environmental system and obtain an integrated conception of the situation. An environmental economic perspective is one of the basic elements in this context to determine the optimal allocation of environmental resources in view of achieving a maximum amount of welfare.

Although the focus in environmental economics research has somewhat shifted towards the issue of climate change in recent years (Dietz and Maddison, 2009), environmental pollution continues to be a substantial problem in many countries. The principal economic effect generally relates to the impact of pollution on public health (Boyd and Genuis, 2008; Valent et al., 2004). Developing countries, often on the verge of or in the process of transforming to a more industrialized economy, are faced with higher health risks due to elevated pollution levels (Briggs, 2003). Although governments mostly aim to protect their population from exposure to harmful substances, the legislative framework and hence the implementation of stricter environmental regulations in these countries is mostly lagging behind economic evolutions. In this setting, it might be useful for these countries to take advantage of the experience acquired in industrialized countries. Since most western countries have already passed a period of rapid industrial development, often without firmly established environmental regulations, some knowledge has been built up with regard to the (long-term) consequences of environmental pollution. Besides reaping the benefits from the lessons learned in other countries, this will allow developing countries to increase their competitiveness with industrialized countries in a sustainable way.

In this dissertation, we will especially focus on one particular form of environmental pollution, i.e. soil pollution. Soil pollution is much more static in comparison with other forms of environmental pollution such as air and water pollution, which are continuously being spread by environmental flow systems. This implies that even long after the pollutants have been introduced in the soil, they can still be present in above average amounts and cause harm to public health and the environment. As a consequence, soil remediation strategies require a specific approach, taking into account the different levels of stakeholders that have been affected as well as the temporal aspect of soil contamination. Obviously, the landowner will be one of the most important stakeholders in this setting, since the loss of use value might have an impact on the economic returns of this land. However, it is important to incorporate the adverse economic effects to individuals close to the pollution and society as a whole as well. This research will therefore focus on both aspects of soil pollution, i.e. the economic damage to the landowner and to society as a result of the pollution. In the next section, the economic theory underlying environmental pollution is discussed in more detail, while section 1.3 provides an overview of nonmarket valuation techniques. In section 1.4, we will elaborate on the case study that is analyzed throughout this dissertation. The research design and objectives will be considered in the final section of this chapter.

1.2 Market failure and externalities

One of the main objectives in welfare economics is to maximize social welfare by allocating resources as efficiently as possible. Resource allocation refers to (1) the types of goods and services produced, (2) the combination of inputs used to produce these goods and (3) the way in which the goods are distributed between persons. Although an efficient allocation of resources is a necessary condition for reaching a social welfare maximum, (perfectly competitive) market economies are themselves incapable of achieving this goal. An additional condition requires that it ought to be impossible to increase social welfare by transferring consumption goods between persons. This relates to the distributive nature of resource allocation. Therefore, government intervention can be needed to redistribute welfare among its citizens in order to pursue a maximum amount of social welfare (Perman et al., 1996).

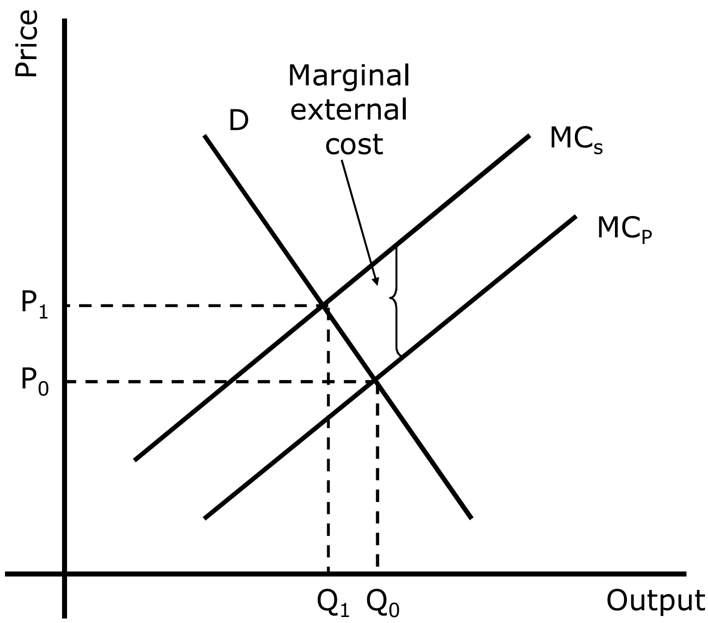
However, in case markets fail to allocate resources efficiently in the first place, government intervention can also be required. There are a number of ways in which the conditions for a perfectly competitive market can be violated[1]. Information asymmetry is an example of a market failure that is rather common in contamination cases. In most cases, not all parties involved are equally informed on the severity and the consequences of the contamination, leading to distorted market outcomes. We will return to this type of market failure in the last two chapters of this dissertation. However, the presence of contamination itself is often a result of a market failure. More specifically, this kind of market failure is referred as an externality. An externality (or an external effect) occurs when the production or consumption decisions of one agent affect the utility of another agent in an unintended way and when no compensation is made by the producer of the external effect to the affected party (Perman et al., 1996). From this definition it is clear that externalities are inherent to economic behavior. Although there are examples of positive externalities, negative externalities are far more abundant. Since this dissertation focuses on environmental pollution as an external cost, the subject will be discussed further from this perspective.

Mostly, the occurrence of pollution is a result of a lack of incentives on the polluter's side to take into account the effects of private decisions on third parties. If it is assumed for simplicity that the pollution is brought about by a production facility, the company's decisions will be aimed at maximizing its profits by producing the amount of output at which the demand curve D intersects the company's private marginal cost curve (MC_P). In Figure 1.1, this corresponds to an output of Q_0. However, in this situation the polluter fails to take into account external effects if no rules are established to

[1] For an elaborate discussion on this subject, we refer to Kolstad (2000) and Perman et al. (1996).

compensate for the damage caused. Hence, the external costs will be borne by a third party or society as a whole. The actual, social cost produced by the polluter is represented by the marginal social cost curve (MC_S). Consequently, if the external effects are internalized, the production will decrease to a level Q_1 at a higher price P_1. The marginal external cost at a certain output level equals the difference between the marginal social cost and the marginal private cost.

Figure 1.1: Marginal cost curves in the presence of external costs

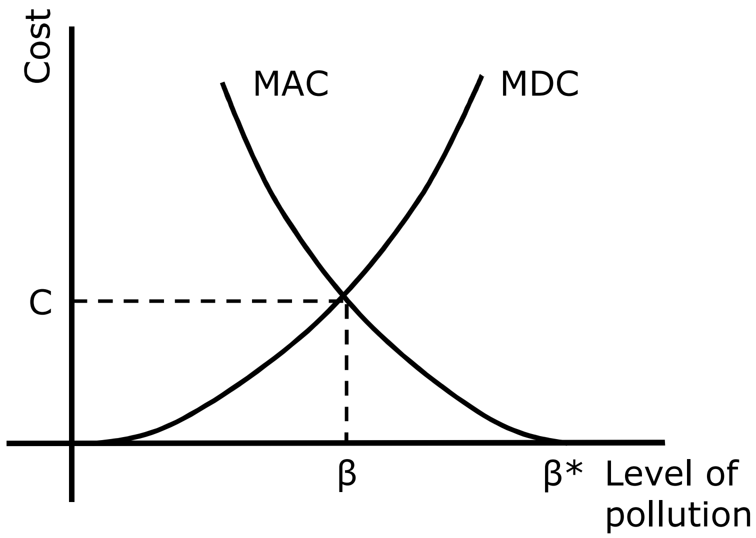


When defining the concept of pollution, various disciplines focus on different aspects of the pollution. For a natural or environmental scientist, pollution can be described by the physical effects of some pollutants on the environment. This can be exhibited in numerous ways, such as biological effects in the form of ecosystem stress or chemical reactions such as damages to buildings caused by acid rain. However, from an economic point of view, the extent of the pollution is commonly measured by means of the population's reaction to these physical effects. Humans might experience feelings of inconvenience, distaste, unpleasantness, stress or anxiety in the presence of pollution and change their behavior accordingly in order to adjust to the situation. The term economic pollution therefore refers to the human reactions and the resulting loss of welfare associated with pollution (Pearce and Turner, 1990).

This definition of economic pollution entails that the presence of polluting elements in the environment does not necessarily imply that there is economic pollution. The economist is particularly interested in pollution flows that exceed the environment's absorbing capacity and that cause direct or indirect damage to the population. Moreover, even if there is economic pollution, it does not necessarily mean that it needs to be eliminated (completely) from an economic perspective. An optimal level of economic pollution can be obtained by intersecting the marginal abatement cost curve (MAC) and the marginal damage cost curve (MDC) (Figure 1.2). The former curve represents the cost that is required to reduce pollution levels by one extra unit, while the MDC curve entails the extra damage cost produced by one unit of pollution. An efficient level of pollution (abatement) is reached at their intersection, implying an optimal pollution level β at cost C in Figure 1.2. The efficient level of pollution abatement is the distance β* - β, in which β* equals the pollution level in the situation without abatement[2].

[2] For a formal derivation of optimal pollution levels, we refer to Pearce and Turner (1990).

Figure 1.2: Optimal pollution level



1.3 Nonmarket valuation

Calculating the costs for abating pollution levels, for example by implementing advanced environmental regulations, is generally a rather straightforward task. However, revealing the economic damage caused by environmental pollution is more difficult to quantify because there is no actual market for this 'good', making it is impossible to estimate market demand (or supply) curve directly (Boardman et al., 2011). Hence, the economic damage costs for environmental pollution have to be derived indirectly by means of nonmarket valuation techniques. These methods are based on the fact that people are continuously making tradeoffs and expressing their preferences by choosing certain products and services to purchase over other products and services. Economic theory assumes that individuals are motivated by a maximization of personal utility in their purchase decisions. Since these transactions mostly occur in terms of monetary units, exchanging money for goods indicates the people's willingness to pay for these goods. However, environmental goods and services are generally not traded in the market and often have public good characteristics. The latter implies that people cannot be withheld from consuming them by other people's consumption and that the consumption is non-excludable[3]. Since monetary values are the yardstick for indicating gains and losses in utility in welfare economics, an important objective in environmental economics is to put a monetary value on environmental services in order to reflect people's preferences. As Pearce and Turner (1990) put it: "Environmental economists simply have to bear the burden of trying to explain what the use of money measures means, and what it does not mean."

Following Mitchell and Carson (1989), methods for estimating economic values can be classified on the basis of two characteristics. The first characteristic concerns the origin of the data, either from real-life market behavior in which people are faced with the consequences of choosing from multiple alternatives or from people's responses to hypothetical questions, i.e. without actual (financial) consequences. The second characteristic involves the way in which monetary values are inferred from these data, either directly or indirectly. A general overview of principal economic valuation techniques can be found in Table 1.1.

Table 1.1: Nonmarket valuation methods

	Observed behavior	**Hypothetical questions**
Direct	Competitive market price	Bidding games
	Simulated markets	Willingness-to-pay questions
Indirect	Travel cost	Discrete choice experiments
	Hedonic property values	Contingent ranking
	Averting behavior	Contingent referendum

Source: Adapted from Freeman (1999)

Mostly, these methods are classified according to the first characteristic. Methods based on observed behavior are referred to as revealed preference

[3] The latter condition is not required by all authors to be labeled a public good (Perman et al., 1996).

(RP) techniques. Direct RP methods are rarely used in case of environmental service flows because this would imply that the flow is effectively traded in the market and people can maximize their utility by choosing their preferred quantity at a given price. On the other hand, indirect RP techniques aim to value nonmarket goods or services implicitly by considering a complementary or substitution good that is traded in the market. For example, the hedonic property value method assumes that certain environmental goods – e.g. scenic vistas or soil pollution – are incorporated in the prices people are willing to pay for real estate. This way, an economic value for the environmental good can be inferred from property values.

Methods based on hypothetical questions are labeled as expressed or stated preference (SP) techniques, since they allow individuals to express their preferences with regard to goods or services. Generally, direct and indirect SP methods are referred to as contingent valuation (CV) and choice modeling approaches, respectively. CV is a survey-based methodology for eliciting values people place on nonmarket goods and services (Boyle, 2003). Unlike CV approaches, individuals are not asked to directly assign a monetary value to the product of interest in choice modeling. Rather, individuals are stimulated to order preferences with regard to (attributes of) the product in a number of questions. If price is included in the order preference task, a monetary valuation for the nonmarket good can be derived indirectly. Choice modeling methods include discrete choice experiments (DCE), contingent ranking and referendum approaches (Freeman, 1999).

There are a number of issues that have to be taken into account when working with nonmarket valuation methods (Freeman, 2003). First, individuals ought to be well informed on the nature of the environmental service flows under consideration and how these flows will affect their well-being. Otherwise, the observed behavior or responses to questions will reflect their ignorance rather than the true value they attach to the service. Secondly, individuals face income constraints which will also be reflected in the economic values they assign to environmental service flows. The results of empirical research must therefore be considered in its context and benefit transfers should be applied cautiously. Lastly, even though environmental goods are not traded in the market, it is assumed that individuals have well defined preferences for bundles of goods that will contribute most to their well-being. This implies that, also with regard to environmental service flows, people are aware which attribute will render them the most utility and are able to compare them to each other.

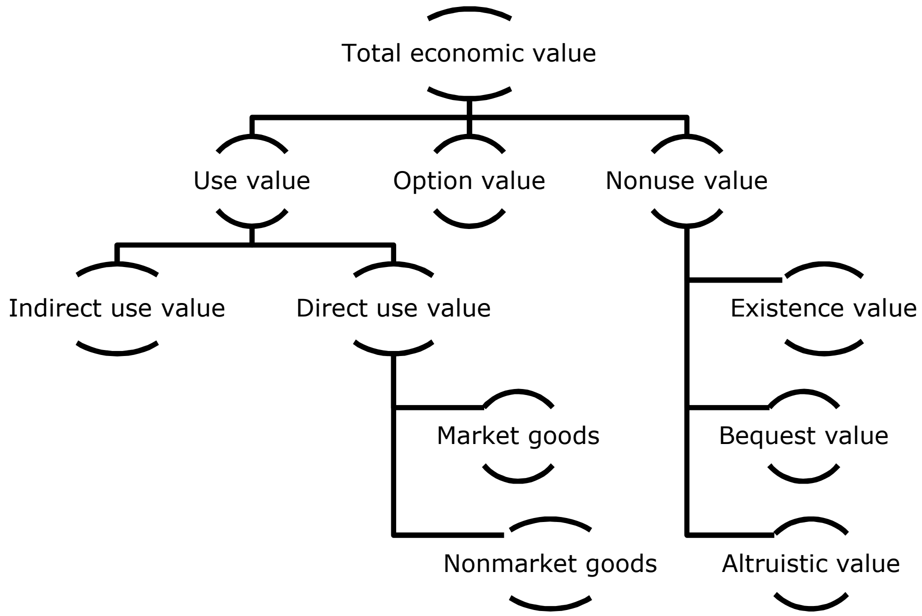
Specifically concentrating on SP methods, there are some additional difficulties that come into play. Boardman et al. (2011) mention four broad issues that have to be considered in environmental valuation studies using SP techniques, which we will discuss briefly. Each SP method has to deal with some of them to a certain degree. The first issue involves the respondents' capability of understanding the questions being asked and placing the good under consideration in an economic context. This problem seems to be more severe in goods that respondents are not expected to consume or that they are not familiar with, as will be discussed further in section 6.4.3. Secondly, since SP studies often handle controversial and complex themes, it is important to uphold a condition of neutrality in phrasing the questions being asked. The best way to manage this issue is by pre-testing the survey in independent focus groups. The third problem involves judgment bias, which covers a number of biases SP studies have been confronted with in the empirical literature. These biases include noncommitment bias, order effects, embedding effects and starting point bias. The last issue in SP methods concerns the possibility of strategic responses to the hypothetical questions. If respondents have an incentive to behave strategically, for example if they believe answering questions in a certain way will result in beneficial outcomes, the SP analysis will not reflect the respondents' true preferences.

The problematic aspects related to SP techniques are the most compelling argument why many economists tend to favor RP over SP methods.

However, also RP methods deal with restrictions in some situations. The issue of asymmetric information mentioned previously is particularly a problem in RP methods, since it is assumed here that individuals buying a product are fully aware of all product characteristics and the way in which it might affect utility. Obviously, this may not be the case in all circumstances. Well-designed SP studies are capable of (partly) circumventing this issue, because it allows to explicitly inform the respondents on the environmental service flows and their impact on utility. Moreover, if there is no clearly identifiable link between the product of interest and a market good – for example, in case the research involves an environmental condition that is currently not applicable yet – market data are unavailable and using RP methods is simply not possible. In this case, SP methods can reveal the public's preferences and willingness to pay with regard to the environmental service.

The last argument concerns the inability of RP methods to estimate nonuse values, also referred to as passive use values (Flores, 2003). Krutilla (1967) was one of the first to suggest that RP techniques might not accurately measure the complete societal value of a certain product, because it ignores nonuse values. Although there is a lot of controversy in the literature with regard to this concept and which values it actually includes (Boardman et al., 2011), generally three nonuse values are distinguished, i.e. bequest values, altruistic values and existence values (Kolstad, 2000). The former two values hinge on people's motivations for maintaining the product in a certain state for other people. More specifically, the bequest value refers to the utility obtained from conserving the product for future generations, while the altruistic value attaches value to the utility other people derive from the product. On the other hand, some individuals might value environmental service flows merely for its existence, which is referred to as the existence value. The separate category of option values, for which the literature disagrees in which place it belongs in the total economic value scheme (Figure 1.3), refers to an individual's utility from preserving a product in order to maintain the possibility to use it in the future.

Figure 1.3: Total economic value scheme



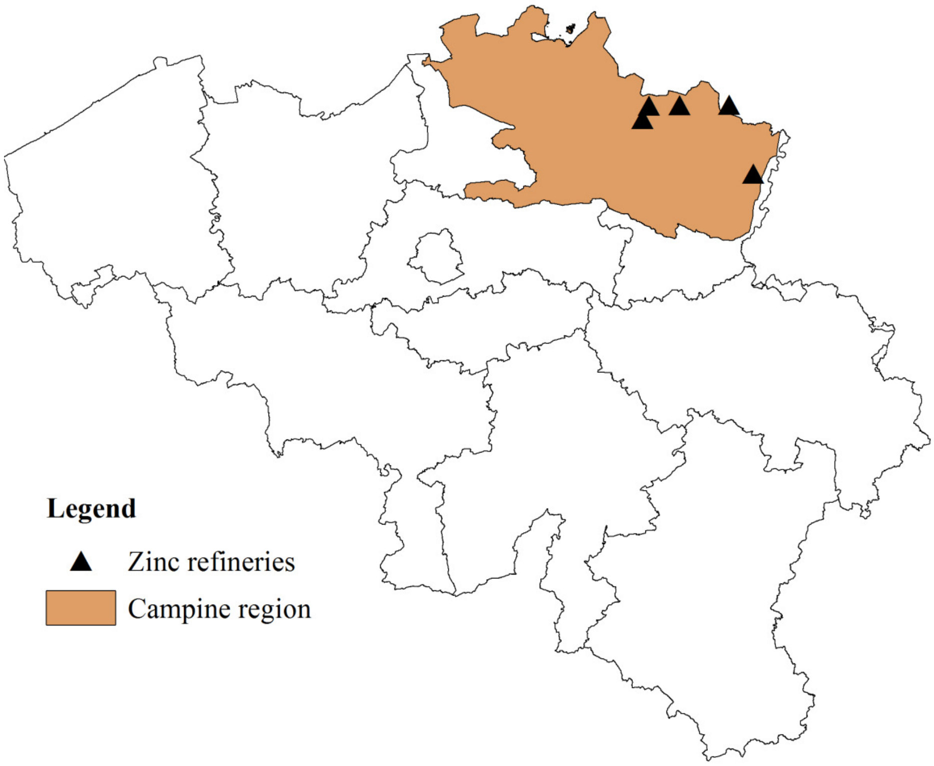
Source: Adapted from Munasinghe (1992)

In its most fundamental form, society is faced with the problem of choosing the amount of environmental resource flows that generates the highest possible level of well-being for humans (Freeman, 2003). Therefore, nonmarket valuation methods are useful in creating a framework that allows for an adequate comparison between market and nonmarket goods using the same yardstick, i.e. monetary values. On the subject which method will return the most accurate value estimates, it should not be necessary to prefer one category of valuation methods over another, but rather consider them as complementary to each other. In case it is difficult to estimate parameter coefficients using RP models alone or it is simply not clear or possible to correctly take into account the variable of interest, SP methods can provide a useful complementary tool to check whether the expressed preferences correspond to the behavioral responses. Nevertheless, each method deals with its own strengths and weaknesses that have to be taken into account when interpreting value estimates.

1.4 Campine region

This dissertation focuses on one particular case study, i.e. the environmental pollution problem in the Campine region in Belgium. In the end of the nineteenth century the metallurgic industry was attracted to the area by a combination of high quality transport infrastructure, low population densities, high unemployment rates and existing coal and ore stocks nearby. Some nonferrous smelters were introduced in Hoboken and Olen, but the main activity of zinc (Zn) refinery was established in Balen, Lommel, Overpelt, Dilsen-Stokkem and Budel (The Netherlands) (Figure 1.4). Until the 1970s, the process used to extract this metal involved heating up Zn ores up to a temperature of 1400°C in order to condense and capture pure zinc. In this production process, a lot of other heavy metals such as cadmium (Cd), lead (Pb) and copper (Cu) were volatilized as well. In case these metals were not captured by condensers, they were expelled into the air through smoke stacks (Clauw, 2007). This way, Zn smelters have diffusely contaminated an extensive area of 700 km² in both Belgium and the Netherlands with heavy metals (Hogervorst et al., 2007).

Figure 1.4: Situating the Campine region and the Zn refineries in Belgium



Except for parts of chapter 2, we will focus solely on the contamination in the Belgian part of the Campine region in this study. The contamination in this area was produced by the cluster of Zn smelters – three in Belgium and one in the Netherlands – located close to each other on the north side of the Campine region. The introduction of strict emission guidelines and especially the transfer from a pyro-metallurgic to an electrolytic refining process initiated a steady decrease in heavy metal emissions since the middle of the 1970s. The dispersion of heavy metals was brought down to insignificant levels in the mid-1980s (Staessen et al., 1995). Nevertheless, the area has been severely affected by the emissions and particularly soils still have to deal with elevated levels of heavy metals.

Currently, only soil Cd concentration levels have been found to exceed threshold values set by the Flemish government in some residential and agricultural areas (Table 1.2). In Flanders, all subjects relating to soil contamination are managed by the Soil Decree (Flemish Government, 2007). This decree discerns three soil concentrations for pollutants, i.e. threshold values, guide values and target values. The former two are further specified according to zoning type (residential, agricultural, industrial,…). The threshold value and the guide value for Cd contamination in agricultural soils are 2 ppm and 1.2 ppm, respectively. When a remediation project is set up, the Soil Decree determines that soil pollutant concentrations should at least decrease to guide values. The target value (0.7 ppm) corresponds to the supposed background contamination level in unharmed soils.

Table 1.2: Soil standards for Cd in Flanders (in ppm) (Flemish Government, 2007)

Land use	**Nature**	**Agriculture**	**Residential**	**Recreation**	**Industry**
Target value	0.7	0.7	0.7	0.7	0.7
Guide value	1.2	1.2	1.2	1.2	1.2
Threshold value	2	2	6	9.5	30

A substantial amount of research efforts have been put into the environmental pollution issue in the Campine region since the end of the 1970s. Numerous studies have tried to register the situation from different perspectives. First of all, research focused on the health effects due to the contamination. This epidemiologic research comprised three study periods, i.e. Cadmibel (1985-1989), Pheecad I (1991-1995) and Pheecad II (1985-2004). More details on these study periods and on the follow up BONK[4] research ordered by the Flemish Government in 2006 will be provided in chapter 3. From 2004 until 2008, the Interreg-project BeNeKempen aimed to assemble all the information with regard to the case study that has been

[4] BONK: Blootstellingsonderzoek Noorderkempen (Exposure Research Northern Campine)

gathered so far. This has led, for example, to the construction of a very extensive database of soil sample measurements that will be used in chapter 2. Furthermore, the doctoral thesis of Frederik Clauw focused on the impact of subjective and objective measures of risk on residential property values (Clauw, 2007).

Since Cd is the trace element with the highest bioaccumulation index in green plants (Kabata-Pendias and Pendias, 1992), another research topic were the consequences of the soil contamination for the agricultural sector in the Campine region. Farmers should be cautious when cultivating crops on Cd contaminated land, because food from plants generally holds higher amounts of Cd than food from animals such as meat, eggs and dairy products (Jarup and Akesson, 2009). The highest Cd concentrations in food from plants are found in leafy green vegetables, root vegetables, potatoes, cereals, seeds and nuts (Egan et al., 2007). Given that the Cd levels in some agricultural parcels in the Campine region can pose problems to farmers, a working group was established in the BeNeKempen-project in order to clarify under which circumstances it is safe to cultivate a number of crops which are frequently produced in Belgium (Smolders et al., 2007). Datasets from different Belgian and Dutch institutions including soil and plant Cd samples were combined to verify at which soil Cd concentrations crops were likely to exceed food standards. This resulted in the cultivation advice presented in Table 1.3.

Table 1.3: Soil Cd concentrations (in ppm) at which crops exceed food thresholds (Smolders et al., 2007)

	5% exceedance of threshold					**10% exceedance of threshold**				
pH	**4,5**	**5**	**5,5**	**6**	**6,5**	**4,5**	**5**	**5,5**	**6**	**6,5**
Potato	0.8	1	1.3	1.7	2.2	1.5	2	2.6	3.4	4.3
Endive	0.2	0.3	0.6	1.2	2.3	0.3	0.5	0.9	1.7	3.2
Celeriac	0.2	0.2	0.2	0.3	0.3	0.2	0.2	0.3	0.3	0.4
Cucumber			7.2					9.5		
Cabbage			1.7					1.7		
Onion			1.8					1.8		
Peas			0.8					0.8		
Asparagus			3.4					3.4		
Beans	1.5	2.1	3	4.2	5.8	1.9	2.7	3.8	5.3	7.4
Scorzonera	<0.1	<0.1	<0.1	0.2	0.5	<0.1	<0.1	0.1	0.3	0.7
Lettuce	0.3	0.4	0.5	0.7	0.9	0.5	0.6	0.8	1	1.3
Spinach	0.2	0.3	0.3	0.3	0.4	0.3	0.4	0.4	0.5	0.5
Tomato	0.9	1.1	1.4	1.7	2.2	1.6	2.1	2.6	3.3	3.9
Carrot	0.2	0.2	0.3	0.4	0.5	0.3	0.3	0.5	0.6	0.9
Celery	<0.1	<0.1	0.1	0.2	0.2	<0.1	0.1	0.2	0.2	0.3
Leek	0.1	0.2	0.3	0.5	0.9	0.2	0.3	0.5	0.8	1.3
Grass	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Corn (cob)	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Corn (plant)	0.6	0.7	0.9	1.1	1.3	1	1.2	1.4	1.8	1.9

The findings from Egan et al (2007) regarding the crops that are susceptible to Cd admission are confirmed by this research. Most green vegetables (lettuce, spinach, celery, celeriac and endive) and root vegetables (carrot, leek and scorzonera) can only tolerate very low soil Cd concentrations. Potatoes were also rather susceptible to exceeding Cd thresholds. Notably, even at Cd concentrations below the 2 ppm threshold some crops are able to take up contaminants to such an extent that food threshold values will be exceeded. The remaining vegetables such as cabbage, onions, beans, asparagus and tomatoes can stand moderate amounts of Cd, while a few crops (cucumbers and corncob) are quite tolerant to Cd in soils. Considering that the sandy soils in the Campine region generally have pH values around 5.5, it is clear that the agricultural potential of soils affected by contamination are substantially reduced with regard to crop cultivation.

The Centre for Environmental Sciences at Hasselt University was one of the pioneers in phytoremediation research, a plant-based technology to remove or dissolve pollutants from natural mediums (Vangronsveld et al., 1995). In this dissertation, phytoremediation is used as an umbrella term for all forms of sustainable marginal land management using plants. Since phytoremediation is particularly suitable for application to agricultural land because these soils are used for growing crops, the impact of phytoremediation on farmer income has been economically analyzed for the Campine region in the doctoral thesis of Nele Witters (Witters, 2011). It was concluded that phytoremediation by means of energy maize takes (much) more time but is also more profitable to farmers than cultivating willow for removing the contaminants. On the other hand, Tom Kuppens modeled the viability of pyrolysis of short rotation coppice from phytoremediation in his doctoral thesis (Kuppens, 2012).

1.5 Research design

The previous section has demonstrated that the pollution in the Campine region has been extensively researched from different perspectives in former years. This raises the question whether the contamination still presents a problem from an economic point of view. And if so, is it still worthwhile to remove the contaminants from (agricultural) soils? In order to provide an appropriate answer to these questions, the topic has to be handled from a private as well as a societal perspective. With regard to the former perspective, we will focus on the asset value depreciation faced by agricultural landowners with contaminated soils in the last two chapters of this research. For the latter perspective, the health economic damage to society as a result of the pollution will be handled in the two preceding chapters. Nonmarket valuation techniques will be an indispensable tool for measuring and quantifying the external costs associated with the contamination. If it turns out that the pollution appears to bring about significant external costs, these costs should be internalized in the cost-benefit analysis of remediation strategies, which will allow for a more accurate analysis. Therefore, the research question to be answered in this dissertation is as follows:

To which extent are there still external costs associated with the environmental pollution in the Campine region?

In order to operationalize this research question, the following objectives are handled in the chapters of this dissertation.

- Chapter 2: Get a better view on the extent of the pollution and determine the amount of contaminated farmland suitable for phytoremediation using geostatistical prediction techniques

Since we will focus on merely one case study in this dissertation, it is important to obtain an excellent insight into the dissemination of the pollution in the area. Therefore, we will aim to reproduce the pollution contours using a Geographic Information System (GIS) in this chapter. Besides providing a better awareness of the contamination levels and its spatial distribution, these contours will also serve as a basis for the research in chapters 3 and 5. In section 2.2 we will elaborate on spatial interpolation techniques, and kriging methods more particularly, that can be used to predict pollution levels at unsampled locations. Data for generating these predictions maps are recovered from former research projects in the area. In section 2.3 an exploratory spatial data analysis will allow us to obtain a better understanding of the data and test whether the data comply to assumptions of normality, stationarity and spatial dependence needed for spatial interpolation. After selecting the correct model specification, section 2.4 will provide prediction maps in which the pollution levels are situated throughout the research area. As a sensitivity analysis, the prediction maps are accompanied by standard error maps in which the uncertainty associated with the estimated values is exposed.

Since the contamination of agricultural soils is a significant topic of interest in this dissertation, a second objective in this chapter is to extract the farmland that has been excessively affected by heavy metals. We will focus more specifically on farmland which can potentially be allocated to phytoremediation in order to reduce contamination levels to below threshold values. Therefore, a methodology is developed in section 2.2 for determining ranges of pollution levels for which phytoremediation could represent an appropriate remediation strategy. This approach builds on the fact that only mildly contaminated parcels can be completely remediated by means of phytoremediation. Moreover, farmers will only be interested in its application in case remediation time frames are restricted. In section 2.3 we will discuss the data that are used to determine these ranges. A phytoremediation field experiment that has been recently performed in the Campine region was very helpful in this setting to provide more details on biological parameters. Combining the results of the spatial interpolation of contamination levels with the ranges determined in this setting will allow us to find out how much farmland can be restored by means of phytoremediation in section 2.4.

- Chapter 3: Establish the extra incidence of illnesses attributable to environmental pollution in a health risk assessment

In section 3.1 a concise literature review provides more insight into environmental exposure pathways and toxic effects on public health caused by cadmium (Cd), the most prevalent heavy metal in the Campine region. Numerous epidemiologic studies have shown relationships between Cd exposure and kidney damage, decreased bone density and cancer mortality. Specifically for the Campine region, research has provided evidence for the existence of an increased incidence of osteoporotic fractures and lung cancer among highly Cd exposed individuals. In this chapter we will extend on this framework in order to determine how many illnesses in the area are attributable to environmental pollution. In section 3.2 we will discuss two approaches that can be used to identify the economic burden produced by an adverse environmental effect and explain why the damage function approach is most suitable in this context. The data used in this chapter (section 3.3) emerge from different sources: 1) the relative risks for incurring these illnesses in the area are provided by the peer-reviewed papers from Staessen and Nawrot, 2) the exposure assessment is based on the prediction maps created in the previous chapter and 3) the incidence rates come from the Belgian Cancer Registry for lung cancer and from research at the University of Liège for hip fractures. These data sources will then be combined to determine the illnesses attributable to environmental pollution in section 3.4. After the results of our analysis are described, we will provide some information on a follow-up epidemiologic study in order to put our findings in the proper context.

- Chapter 4: Determine the health care costs attributable to environmental pollution in a health economic assessment

In chapter 4, we will elaborate on the results of chapter 3. More specifically, the health care costs of the illnesses discussed in chapter 3 will be analyzed so that the economic damages due to pollution can be estimated. Since the costs for hip fractures in Belgium have recently been assessed by researchers at University of Liège, we will focus solely on determining the costs for an average case of lung cancer in Belgium. Therefore, a restricted framework for cost assessment is adopted, seeing that a broader perspective would require a much more extensive research design which is beyond the scope of this dissertation. In section 4.2, the concept of a case-control design is introduced and explained. This research design will serve to establish the direct medical costs faced by lung cancer patients. Cost data will be provided by the Intermutualistic Agency, a governmental institution that combines the databases of all sickness funds in Belgium (section 4.3). This entails micro-data on the resource use of a sample of lung cancer patients as well as two matching control samples. One control sample is matched only on the basis of socio-economic characteristics, while the other control sample is matched additionally for smoking comorbidities, seeing that smoking is the most important risk factor for lung cancer. The results in section 4.4 will distinguish three kinds of cost estimates, i.e. average direct medical costs for lung cancer and incremental direct medical costs in comparison with the two control samples. Combining these results with the findings of chapter 3 will allow us to estimate the pollution-attributable health care costs for hip fractures and lung cancer in the Campine region.

- Chapter 5: Analyze the impact of soil contamination on farmland values in a hedonic pricing analysis

Coming back to the effect of contamination on the agricultural sector, chapter 5 aims to determine to what extent the pollution levels are capitalized in farmland prices. Hedonic pricing analysis is selected as an appropriate methodology for estimating the welfare effects of the soil contaminants on land prices. Section 5.2 provides an overview of the most important literature on hedonic studies aimed at establishing factors which seem to affect farmland values. The extensiveness of the literature review shows that the hedonic methodology is widely accepted for empirical applications of farmland models. Section 5.3 will first provide some background on the theory underlying hedonic pricing analysis. Secondly, this section will focus on the spatial effects that arise when the regression of an inherently spatial dependent variable such as farmland values is concerned. We will elaborate on spatial econometric techniques that have been developed to take these effects into account. Data, which are obtained from the National Registry Office in Belgium, will be described in section 5.4. The prediction maps for pollution levels created in chapter 2 will be linked to the data in GIS as a measure for the environmental risk associated with each parcel. Section 5.5 will provide the most important results for a classic linear regression model, as well as a quantile regression analysis. As a complementary paragraph, the policy analysis in section 5.6 will aim to clarify the procedures used for land sales transactions in Belgium and how land buyers are informed on adverse characteristics such as soil contamination.

- Chapter 6: Investigate the farmer's preferences in land purchase decisions in case of soil contamination in a discrete choice experiment

In chapter 6 we will advance on the topic of farmer's land purchase decisions by complementing the revealed preference hedonic pricing analysis with a stated preference analysis. This way, we aspire to further unravel the economic and motivational framework underlying farmland price determination with a particular focus on the impact of soil contamination. A stated preference analysis allows us to circumvent any information asymmetry with regard to the soil contamination that might have been at play in the hedonic pricing analysis by providing complete information on the contamination and its consequences to the farmers. In section 6.2 we will elaborate on the technique of discrete choice experiments (DCE), the methodology that is applied in this chapter. DCEs aim to identify an individual's indirect utility function by analyzing the tradeoffs they make between the attributes and attribute levels of different product alternatives. The steps that are required to set up a DCE study will be investigated in section 6.3. One particular aspect in this design is the introduction of a soil contamination attribute level in which a typical pollution situation in the Campine region is hypothesized. In section 6.4, first some descriptive statistics of the respondent sample will be provided. Then, a basic multinomial logit model will be estimated which will indicate the general farmer population's preferences with regard to soil contamination. In order to control for possible heterogeneity in the sample, interaction effects will be added to the analysis afterwards to examine whether farm-level or socio-economic characteristics might have a confounding impact on the soil contamination attribute level.

2 Geostatistical analysis of soil contamination in agricultural land[5][6]

2.1 Introduction

In order to understand and determine the external effects produced by the environmental contamination in the Campine region, it is important to first obtain a decent insight into the extent of the pollution. Over the years, soil scientists gathered an abundance of data regarding heavy metal and particularly cadmium (Cd) concentrations in the area. However, a classic issue with samples of environmental variables is that they are usually measured at some point in space, while the variable is actually spatially continuous. To resolve this issue, spatial interpolation methods have been used to predict values for the variable of interest in the entire research area by means of these sample measurements. This way, values can also be estimated for unsampled locations. Concerning the problem at hand, this chapter aims to create soil Cd prediction maps for the Campine region using spatial interpolation techniques. These predictions will also serve as a basis for other chapters, more particularly chapter 3 and 5.

It has been discussed in the previous chapter that particularly professional farmers have to deal with some adverse effects due to the soil contamination

---

[5] The geostatistical analysis in this chapter is partly based on the work of Frederik Clauw (Clauw, 2007) and Thomas Voets, a former colleague at Hasselt University. We would like to thank both of them for their efforts and the helpful comments in finalizing this chapter.

[6] Parts of this chapter have been published in: Schreurs, E., Voets, T. and Thewys, T. (2011). GIS-based assessment of the biomass potential from phytoremediation of contaminated agricultural land in the Campine region in Belgium. Biomass & Bioenergy 35(10): 4469-4480.
Parts of this research have also been presented at the 8th International Phytotechnologies Society (IPS) conference: Schreurs, E., Voets, T. and Thewys, T. (2011). GIS-based assessment of the agricultural land suitable for heavy metal phytoextraction: A case study of the Campine region (Belgium). IPS conference, Portland (OR), 13-16 September 2011.

because the agricultural potential of their land is affected by the contaminants. In this light, the impact of soil contamination on farmland values will be researched in the last two chapters of this thesis. However, it might be useful to first investigate the extent of the problem specifically related to farmland. Moreover, for agricultural land a solution for the pollution can be provided by phytoremediation[7], whereas this technique is less appropriate for other zoning types such as residential or industrial areas. The second goal of this chapter will therefore focus on determining the amount of farmland that has been affected by soil contamination and that can possibly be remediated by phytoremediation. This approach has already been used to determine the biomass potential that can be realized from phytoremediation (Schreurs et al., 2011).

The remainder of this chapter is structured as follows. In section 2.2, the spatial interpolation techniques used for predicting soil Cd concentrations in the area will be discussed in detail, while section 2.3 will focus on the data needed to (1) make these predictions and (2) determine the farmland suitable for phytoremediation. In section 2.4, the results from these spatial analyses will be provided. Some concluding remarks will be given in section 2.5.

2.2 Methodology

2.2.1 Spatial interpolation

2.2.1.1 Geostatistical analysis

The interpolation of a spatially continuous variable from point samples is a common problem in spatial analysis. Generally, two main spatial interpolation techniques can be distinguished, i.e. deterministic and geostatistical techniques. Although both methods originate from Tobler's

[7] Note that we have defined the term 'phytoremediation' in section 1.4 as an umbrella term for all forms of sustainable marginal land management using plants.

First Law of Geography, which states that everything is related to each other but near things are more related than distant things, they differ in the functions used for interpolation. Deterministic techniques such as inverse distance weighting and spline rely on mathematical functions, while geostatistical techniques such as kriging are based on mathematical *and* statistical functions (Johnston et al., 2001). In other words, deterministic interpolation techniques use mathematical formulas based on the distance between the sample points to predict values for unsampled locations, whereas geostatistical interpolation also uses the statistical properties of the sample points to account for the spatial configuration close to the prediction location (Clauw, 2007).

Geostatistics is a subset of statistics that specializes in the analysis and interpretation of geographically referenced data (Goovaerts, 1997). It describes the spatial continuity of natural phenomena while providing adaptations of classical regression techniques to take advantage of this continuity (Isaaks and Srivastava, 1989). Geostatistical interpolation offers distinct advantages over deterministic interpolation techniques. For example, by including the spatial structure of data, geostatistical interpolation will predict the most precise values from the available sample values (Wackernagel, 2003). Moreover, since geostatistical interpolation is based on random function theory, it is capable of modeling the uncertainty associated with the predicted values. So besides generating predicted values, this technique can also be used to analyze the reliability and variability of predicted values (Houlding, 2000). Deterministic interpolation techniques are unable to provide these kind of measures. Given the advantages of geostatistical interpolation techniques over deterministic methods, only the geostatistical techniques will be used to predict soil Cd values in this chapter.

In order to create surfaces for spatially continuous variables geostatistical techniques rely on three assumptions, i.e. the normality, stationarity and spatial dependence assumption. Similar to its statistical variant, geostatistical normality assumes that the distribution of point values in the sample population resembles a normal distribution. Stationarity is a property of the random function model for inference, which assumes that the characteristics of a random function remain the same when shifting a set of *n* points from one part of the region to another. The assumption implies that the model does not change fundamentally when moving in the universe (Wackernagel, 2003). Spatial dependence can be described as the existence of a functional relationship between what happens at one point in space and what happens elsewhere (Anselin, 1988). While the previous two assumptions were also applicable in general statistics, the assumptions of spatial dependence is undesirable in classic regression analyses because it violates the assumption of independent observations[8]. However, spatial dependence is a prerequisite in geostatistics because it allows to predict soil properties at unsampled locations. For an elaborate discussion of geostatistical assumptions, we refer to Clauw (2007).

2.2.1.2 Kriging

Kriging is the most well-known stochastic method in the category of geostatistical interpolation techniques (Eberly et al., 2004). Kriging methods belong to the family of generalized least squares regression algorithms (Goovaerts, 1997). As explained in the previous section, it derives predicted values based on the distance between points in space and the variation between measurements as a function of distance. Therefore, kriging first quantifies the spatial structure by fitting a spatial dependence model to the dataset. In combination with samples values close to the prediction location, the fitted model is subsequently used to predict values for all unsampled locations (Johnston et al., 2001).

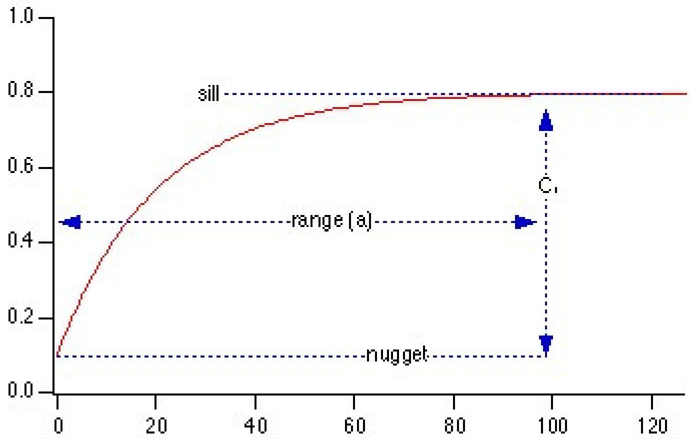
The spatial structure in a dataset can be quantified by means of semivariograms, which capture the spatial dependence between samples by plotting the semivariance of paired sample measurements against the distance between samples in a semivariogram cloud. The semivariance

[8] The issue of spatial dependence in regression analysis will be explained in more detail in chapter 5.

equals half the squared difference between point values separated by a distance d – assuming no direction preference – and is expected to increase as the distance between samples increases, because near samples are assumed to be more similar than distant samples. Since plotting all pairs of locations would quickly lead to a congestion of the plot, the number of points in the empirical semivariogram is reduced by grouping the pairs of locations based on their distance from each other and directions. This process is referred to as 'binning' (Johnston et al., 2001). For each bin, an empirical semiovariogram is created.

The next step involves fitting an appropriate (non)parametric model that captures the structure of the empirical semivariogram. The model is used for developing kriging predictors and standard errors by calculating distance weights for interpolation. A graphical representation of the semivariogram enables us to obtain a better insight into the spatial correlation between data points and their neighbors. There are some basic concepts that are used to describe a semivariogram. The separation distance at which there is considered to be no more spatial autocorrelation between observations – i.e. the point at which the model starts to flatten out – is known as the range. The corresponding semivariance coefficient is referred to as the sill. Since theoretically the semivariance should be zero at zero separation distance, the nugget is the semivariance at the separation distance that approximates zero. The (random) example of a semivariogram in Figure 2.1 might clarify these concepts.

Figure 2.1: Example semivariogram



There are two main parametric kriging models, i.e. simple kriging and ordinary kriging[9]. Both models deal with similar assumptions. More specifically, it is assumed that the spatial variation is homogenous throughout the study area and only depends on the distance between samples points, that there is no underlying trend in the data and that all variation is statistical (Eberly et al., 2004). The difference between simple kriging and ordinary kriging is situated in its assumption about the mean. The former model supposes there is a known, constant mean, while the latter method assumes the constant mean is unknown and needs to be estimated based on the data. More theoretical background on both models can be found in Clauw (2007) and Eberly et al. (2004).

2.2.2 Farmland for phytoremediation

In order to determine the farmland that can potentially be allocated to phytoremediation, it needs to be a realistic proposal from the point of view of the farmers who are confronted with soil pollution. This relates in first instance to the level of contamination, since phytoextraction should only be applied to soils containing rather moderate pollutant concentrations (Zhao et

[9] Other parametric models and nonparametric kriging models such as indicator kriging are not discussed here.

al., 2003). Otherwise, the time needed to achieve target concentrations might surpass practical time frames and farmers might be reluctant to adopt phytoremediation strategies. In this light, the time span is treated as an independent variable in this study. Moreover, if levels of contamination are too high, crop growth of remediating plants might be inhibited as well (Evangelou et al., 2004; Wang et al., 2007). Another important factor is the limited depth that can be remediated, since phytoremediation is restricted by the roots of the plants (Van Nevel et al., 2007). Deep contamination can only be removed by plants with an extended root system distribution, which strongly reduces the number of crops to be considered. For this reason, phytoextraction should focus on shallow contaminated soils.

Given that the time span is restricted to reasonable time frames, phytoremediation is only applicable within a limited range of soil pollutant concentrations. Therefore, the outer values of this range have to be determined. The lower limit is established by legal threshold values that have to be obtained when remediation projects are set up. On the other hand, the upper limit has to be assessed by calculating the decrease in soil pollutant concentrations that can be achieved by means of remediation crops. Hence, the maximal soil concentration that can be handled by means of phytoremediation is found by adding the target soil concentration to the amount of contamination that is removed within a particular time frame. This is illustrated in the following formulas:

$$c_{max} - c_t = \frac{u * t}{d * b * (10{,}000\ \mathrm{m^2/ha})} \quad [1]$$

$$c_{max} = c_t + \frac{u * t}{d * b * (10{,}000\ \mathrm{m^2/ha})} \quad [2]$$

With:
c_{max}: maximal soil Cd concentration (mg kg^{-1})
c_t: target soil Cd concentration (mg kg^{-1})
u: plant uptake (mg ha^{-1} y^{-1})

t: time span (y)

d: depth (m)

b: soil density (kg m^{-3})

It is clear that a number of parameters are needed for this assessment. Firstly, plants that may be used for phytoremediation are selected. Crops used for phytoremediation should combine a number of characteristics (Keller, 2006). They need to be suited to local circumstances such as climate, type of soil, geo-morphology,… The phytoextraction potential for the metal(s) to be removed and the tolerance to other contaminants present in the soil must have been shown in experiments. Furthermore, crops should preferably coincide with current agricultural practices and yield marketable products, which will enhance the acceptability to farmers.

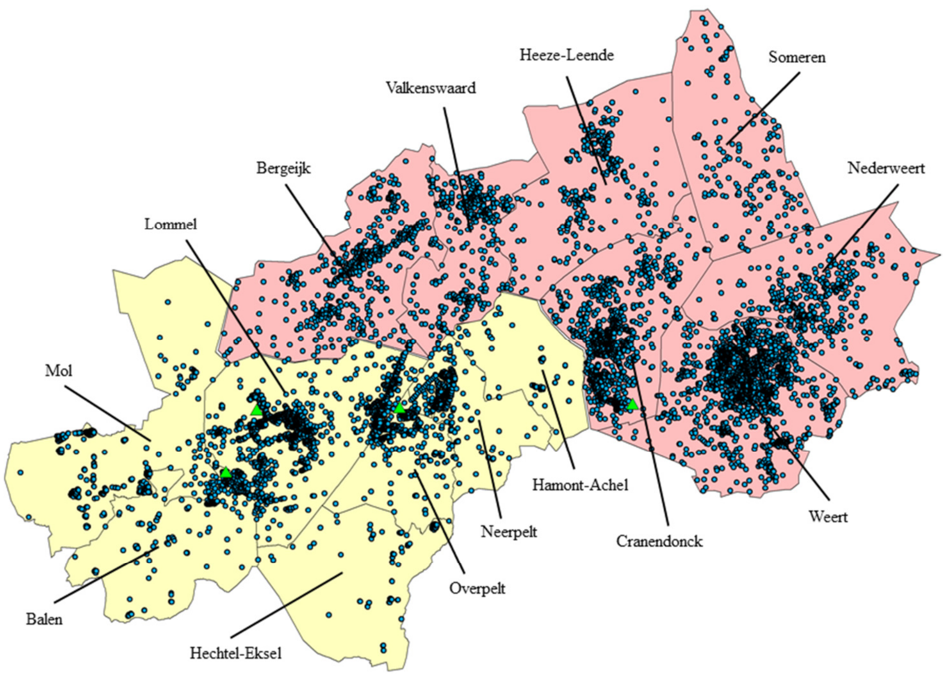
Subsequently, the uptake capabilities of the chosen crops are analyzed. For a realistic assessment of uptake values, it is best to search for field studies using these crops, conducted in soil profiles comparable to the one under investigation. When the other parameters (i.e. target soil concentration, time span, depth, and soil density) are determined, the maximal soil pollutant concentrations can be calculated using equation [2]. Hereafter, the range of soil concentrations is used to determine the agricultural area that can be allocated to phytoremediation. This can be achieved by classifying the region into geographical subareas using the lower and the upper limit of the soil pollutant concentrations as boundaries. The agricultural soils are then extracted from these subareas to obtain the amount of farmland suitable for phytoremediation.

2.3 Data

2.3.1 Geostatistical analysis

As already mentioned in the previous chapter, numerous individual studies have tried to register the soil situation in the Campine region from different perspectives. In the Interreg-project BeNeKempen (2004-2008) all the information with regard to heavy metal concentrations gathered so far was assembled. This way, a very extensive database of soil sample measurements was constructed, containing 11,885 soil Cd samples. The study area for this research consisted of 7 Belgian municipalities, i.e. Mol, Balen, Lommel, Hechtel-Eksel, Hamont-Achel, Neerpelt and Overpelt, and 7 Dutch municipalities, i.e. Bergeijk, Cranendonck, Heeze-Leende, Nederweert, Someren, Valkenswaard and Weert. After georeferencing all observations using the Belgian Lambert72 coordinate system, the map in Figure 2.2 shows that the measurements (represented by the blue dots) were predominantly concentrated in the direct vicinity of the major pollution sources (represented by the green triangles). Additionally, more soil samples have been taken in the Dutch part of the study area.

Figure 2.2: Soil Cd sample measurements in the study area



Before the data can be analyzed geostatistically, an exploratory spatial data analysis is performed to find out whether the data comply to the assumptions of normality, stationarity and spatial dependence. In order to test the normality of the sample, histograms of soil Cd samples as well as its logarithmic transformation are created (Figure 2.2). Additionally, some summary statistics are provided to obtain a better understanding of the data (Table 2.1).

Figure 2.3: Histogram soil Cd samples

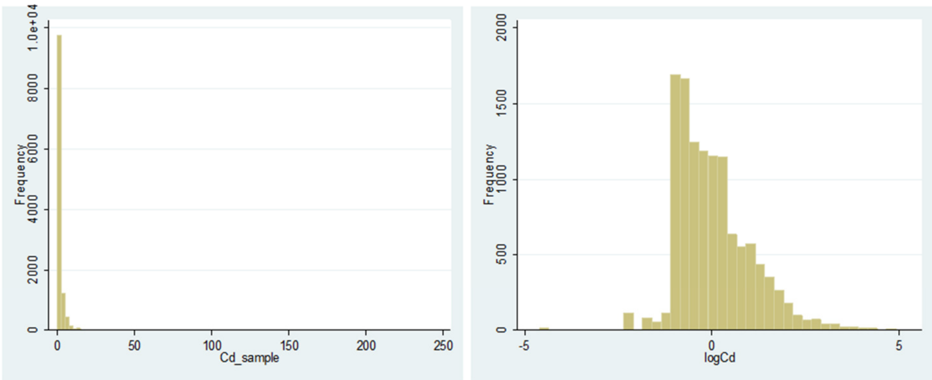


Table 2.1: Summary statistics soil Cd samples

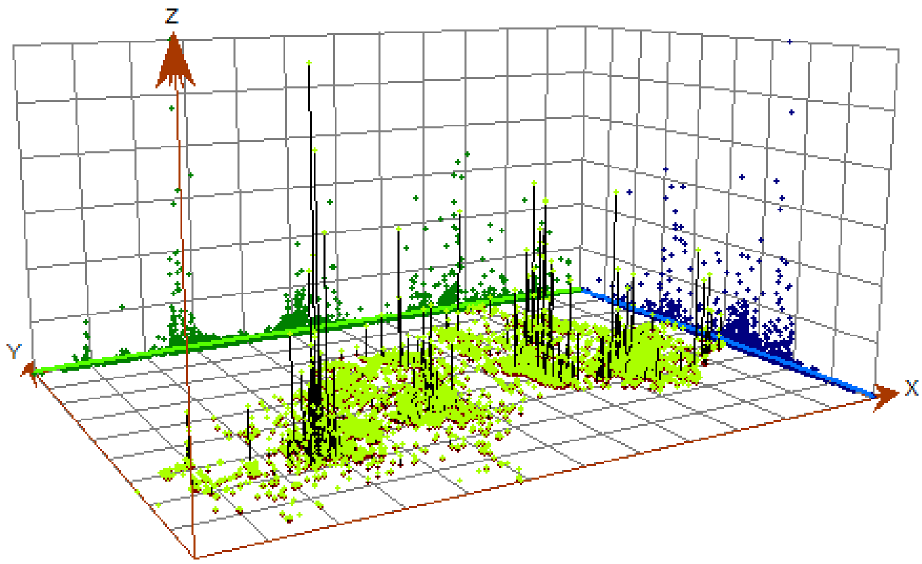
	Cd	**logCd**
Observations	11,885	11,885
Mean	2.22	0.05
Median	0.9	-0.11
Minimum value	0.01	-4.61
Maximum value	228	5.43
Standard deviation	6.56	1.01
Skewness	13.56	0.89
Kurtosis	279.61	5.03

The histogram in Figure 2.3 reveals that the dataset particularly includes low values for Cd. There are only a few (very) high values up to 228 ppm, which are located in close proximity to the Zn smelters. Since there is no reason to assume that these values are unrealistic or measured incorrectly, they are not excluded from the dataset. If the Cd concentrations are transformed into its logarithm, the distribution resembles more closely to a normal distribution. This is confirmed by the summary statistics in Table 2.1. The mean value is closer to the median value, the skewness coefficient is closer to zero and the kurtosis coefficient is much closer to 3, all indicating that the logarithmic Cd values are more normally distributed than the untransformed Cd values. Therefore, it seems to be reasonable to incorporate logarithmic Cd values in the geostatistical analysis.

Since kriging interpolation techniques will produce suboptimal maps in case the data fail to comply with the stationarity assumption, a trend analysis was performed in order to verify if the dataset is stationary (Figure 2.4). If a trend can be identified in the data, it should be removed by transforming the data to stabilize the variance (Cressie, 1993). The tailored ArcGIS tool tries to detect a trend by creating a 3D graph from the dataset. All sample points are plotted on the X,Y plane by means of the longitude and latitude coordinates. The Cd value of each sample is used to determine the height in the Z dimension. All samples are then projected onto the X,Z and Y,Z plane as scatter plots. Subsequently, polynomial functions are fitted through the scatter plots on the projected planes (Johnston et al., 2001). Since there is no particular trend noticeable in the functions of the Cd sample points in Figure 2.4 – represented by the blue line in the X,Z plane and the green line in the Y,Z plane – no global trend appears to exist and the Cd data are considered to be stationary[10].

[10] Besides the global trend, a directional influence – referred to as anisotropy – can affect the points in the semivariogram, even after the trend is removed. Anisotropy differs from a trend, since it usually cannot be described by a single mathematical formula and it does not originate from a single source.

Figure 2.4: Trend analysis



The presence of spatial dependence in a dataset can be measured by means of the nugget to sill ratio (Cambardella et al., 1994). If this ratio is smaller than 25%, the variable of interest is considered to be strongly spatial dependent. A ratio between 25% and 75% indicates moderate spatial dependence, while ratios higher than 75% show only weak spatial dependence (Ahmadi and Sedghamiz, 2007; Liu et al., 2006). For the entire Cd dataset, the nugget to sill ratio equals 36% which denotes a rather moderate spatial dependence. Since we are particularly interested in the higher levels of contamination in this research, a subsample was created in order to determine the degree of spatial dependence close to the pollution sources. This subsample contained all Cd sample values within a range of 3 km from the three active zinc smelters (n = 4535)[11]. Here, the nugget to sill ratio was 24 %. This indicates that spatial dependence is more dominant in close proximity to the pollution sources.

[11] In Figure 2.5 to 2.7, it can be observed that the range of the semivariogram converges to approximately 3 km.

2.3.2 Phytoremediation potential

In accordance with the requirements stated by Keller (2006), four crops are chosen that seem to be acceptable for phytoremediation in the Campine region. These crops are short rotation coppice willow (*Salix*), silage maize (*Zea mays*), tobacco (*Nicotiana tabacum*) and rapeseed (*Brassica napus*). The latter two crops have to be rotated in a phytoremediation scheme, considering they are very susceptible to funghi like sclerotinia. It is recommended that tobacco[12] and rapeseed will only be cultivated once in three or four years. In the case of maize, there is no need for a strict rotation scheme. Although it is not advisable, in practice maize is often cultivated in monoculture by farmers. However, in the long run this will lead to lower biomass yields. In this study, silage maize will only be used in rotation with rapeseed and tobacco.

Vangronsveld et al. (2009) recently conducted a series of field experiments with numerous phytoremediation crops in the Campine region (Table 2.2). The results for the four crops under consideration will be applied in our model. For simplicity, it is assumed here that the uptake rates and biomass production will remain constant throughout the entire time frame[13].

Table 2.2: Phytoremediation field experiment

Crop	**Plant Cd concentration** (ppm)	**Biomass production** (Mg ha^{-1} y^{-1})	**Total Cd removal** (g ha^{-1} y^{-1})
Maize	3	20	60
Rapeseed	6	8	48
Tobacco	24	8	192
Willow - leaves	60	2.4	144
Willow - twigs	24	8	192
Willow - total		10.4	336

[12] Tobacco for phytoremediation cannot be used for human consumption, seeing that it contains even more Cd than in normal circumstances. Alternatively, it might be used as biomass flow for bioenergy (anaerobic digestion).
[13] For a scenario analysis that incorporates decreasing uptake rates, we refer to Schreurs et al. (2011).

Table 2.3: Other parameters			
Soil density (kg m^{-3})	**Depth** (m)	**Target Cd concentration** (ppm)	**Time frame** (y)
1500	0.4	1.2	21 42

The other parameters are summarized in Table 2.3. Since the soil structure in the Campine region is fairly homogeneous and predominantly consist of sandy particles, soil density is assumed to be around 1500 kg m^{-3} in the entire study area. The Cd contamination is mainly present in the upper 40 cm of agricultural soils (Thewys et al., 2010; Witters et al., 2009). This is probably due to continuing agricultural cultivation on most soils after contamination had occurred. In this way, Cd contamination in the plough layer (assumed to be 40 cm) was homogenized. For all crops under consideration, the majority of the roots are found in this range (Kamh et al., 2005; Keller et al., 2003), although some roots may penetrate deeper into soils. However, since little or no contamination can be found in underlying layers, it is unnecessary to consider larger depths.

With respect to phytoremediation duration, time frames of 21 and 42 years are modeled. In phytoremediation literature, durations in the range of a couple of decades are deemed acceptable to most authors (Dickinson and Pulford, 2005; Maxted et al., 2006; Zhao et al., 2003). Another reason for applying these time frames is for practical reasons. In case of willow cultivation in the Campine region, the most cost-effective option for farmers is a rotation of 7 x 3 years (Ruttens et al., 2008). So when farmers choose to cultivate willow as the phytoremediation crop, it is assumed that the willow rotation period is completely finished. Keeping the time frame to a limited extent, a maximum of two complete rotations is presumed. Although the annual rotations including tobacco and maize are not fixed to a certain amount of time, the same time frames are considered for the sake of comparability. It has not been empirically tested whether these time frames are actually acceptable to farmers.

In determining the amount of farmland suitable for phytoremediation, the research area will be constrained to the Belgian part of the Campine region for two reasons. In first instance, there are differences between legal soil threshold values in Flanders and the Netherlands. In order to premise a single target concentration for the phytoremediation project, the Flemish soil thresholds will be used. As mentioned in Table 1.2, the guide value that should be reached when remediation projects are set up in agricultural soils in Flanders corresponds to 1.2 ppm. Therefore, this value will be adopted as the target concentration and as the lower limit for the range of soil Cd concentrations.

The second argument relates to the maps that are used to extract the farmland eligible for phytoremediation, once the prediction maps are classified using the target soil Cd concentration and the maximal soil Cd concentrations as boundaries. Each year, the Flemish Agency for Agriculture and Fishery (ALV) creates a GIS-layer of farmland that is currently in agricultural cultivation in Flanders. This layer is referred to as 'Landbouwgebruikerspercelen' and is made publicly available to government organizations and universities. Since no comparable GIS-layers are published on a yearly basis in the Netherlands, a joint derivation would include different kinds of farmland between both countries. The ALV-layer for the year 2010 is used, which consists of 10,343 parcels for the study area.

2.4 Results

2.4.1 Geostatistical analysis

2.4.1.1 Empirical semivariogram

By means of the empirical semivariogram the most appropriate model is fitted to the points in the semivariogram in order to reveal the function that relates the semivariance to the distance between observations. According to Eberly et al. (2004), there are three parametric models commonly employed in the semivariogram modeling processs, i.e. the spherical, exponential and Gaussian model. In Figure 2.5 to 2.7, these three models are fitted to the binned values (represented by red dots) in the empirical semivariogram. The blue crosses are averaged points, which have been created by binning empirical semivariogram points. Since it cannot be determined by plain observation which model will best fit our dataset, some additional statistics are required.

Figure 2.5: Spherical model

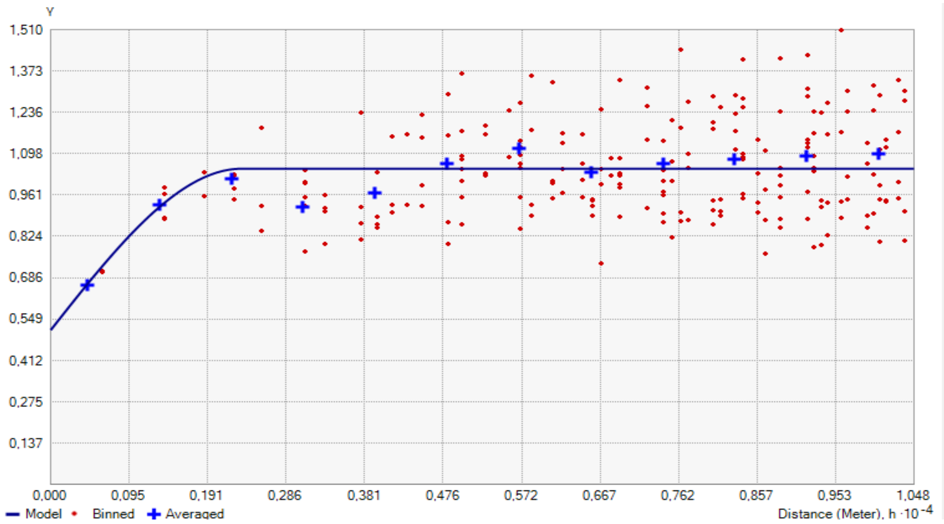


Figure 2.6: Exponential model

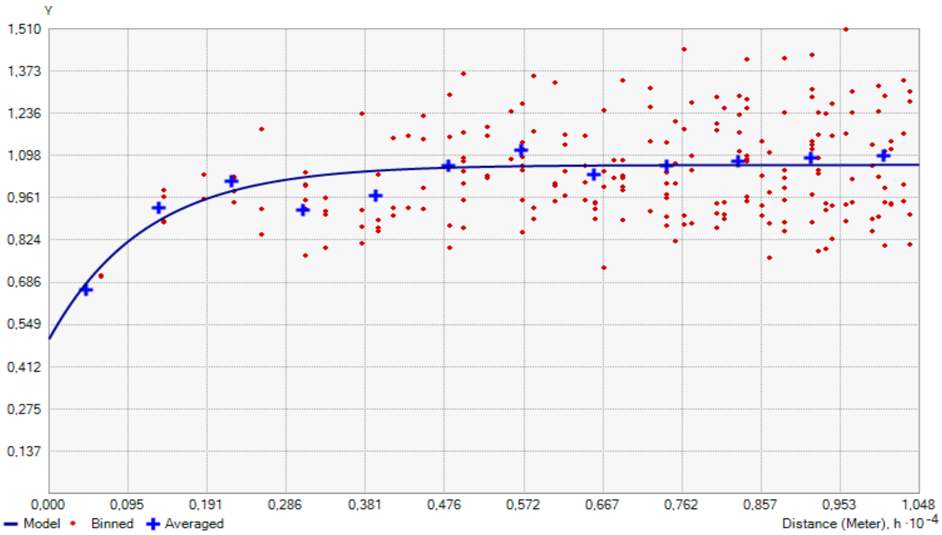
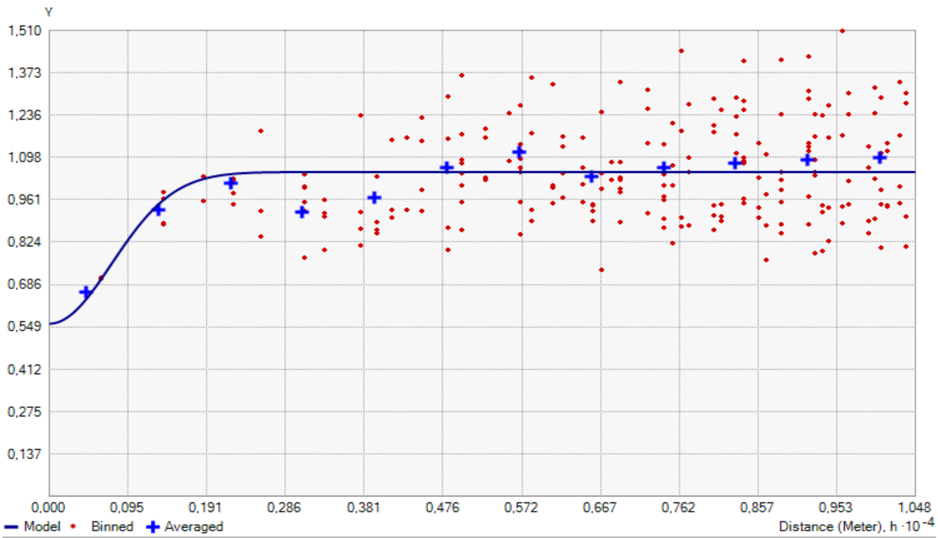


Figure 2.7: Gaussian model



Crossvalidation techniques are able to provide a better understanding of the degree to which the model accurately predicts values at unknown locations. Crossvalidation relies on a comparison between model predictions and observed values by withholding one (or more) observation(s), after which a prediction of the removed data location is made using the other sample points. Since removing one (or few) sample point(s) will have a negligible effect on the prediction model – particularly in large datasets such as this one – predicted values can then be compared to observed values. Deviations from actual observations are described by some summary statistics (Table 2.4).

Table 2.4: Crossvalidation statistics

Model	**Spherical**	**Exponential**	**Gaussian**
ME	-0.258	-0.244	-0.291
RMSE	5.582	5.532	5.744
MSE	-0.042	-0.028	-0.093
RMSSE	2.629	2.488	2.915
ASE	2.861	2.913	2.727

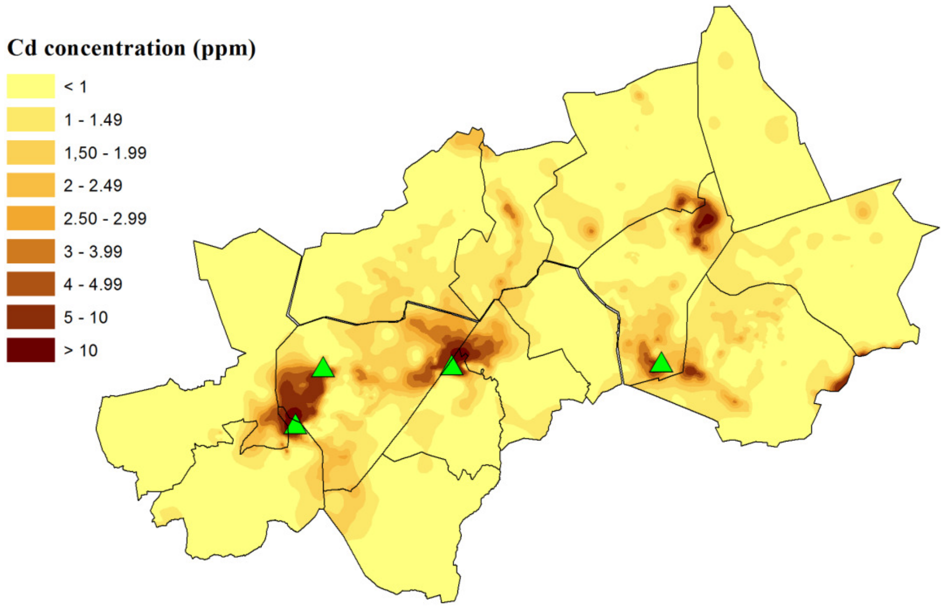
The statistics in Table 2.4 can serve as diagnostic tools to indicate which model is suitable for map production. In order to check whether the model provides accurate predictions, the mean error (ME) and the root-mean-square prediction error (RMSE) should converge to zero. The variability in the predictions is assessed by means of average standard errors (ASE) and root-mean-square standardized errors (RMSSE). If ASE is smaller than RMSE, the variability in the predictions is underestimated, as is also indicated by a RMSSE that is greater than one. The mean standardized prediction error (MSE) should be close to zero for unbiased prediction errors (Clauw, 2007; Johnston et al., 2001). In this setting, it is clear that the exponential model outperforms the other two models in our dataset. Therefore, this model will be selected to create the prediction maps.

2.4.1.2 Prediction maps

Using the empirical semivariogram the ordinary kriging method can be applied to obtain a prediction map of the Cd contamination in the Campine region (Figure 2.8). The prediction map indicates that the highest levels of

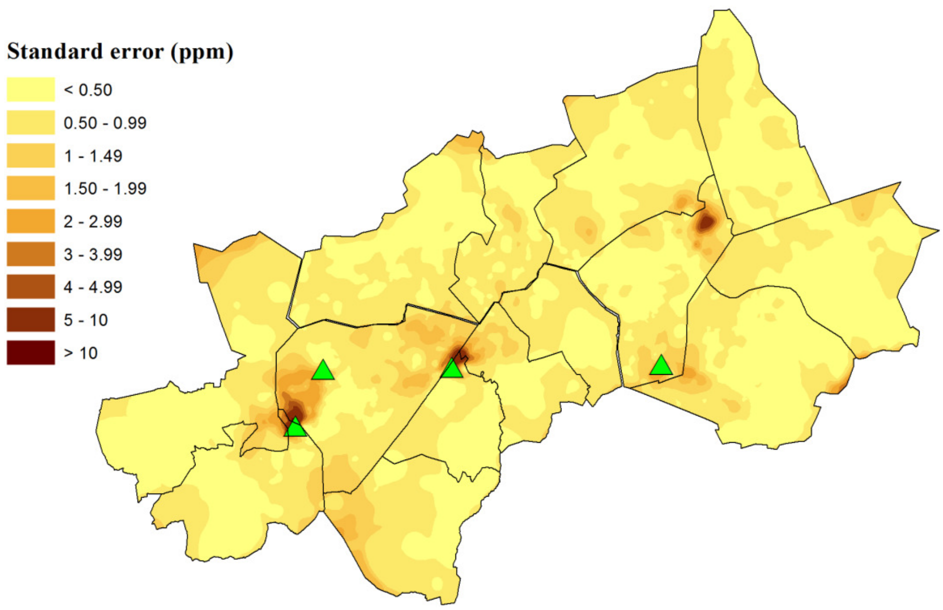
Cd can be found in the vicinity of the three active Zn smelters. The factory in Lommel Maatheide (the second Zn smelter to the left), which has been dismantled in the 1970's, deals with significiantly lower contamination levels as a result of the numerous soil sample measurements that have been taken after the site remediation in 2003 in order to confirm its clean status. Apart from the rather high concentrations in the proximity of the pollution sources – particularly the Balen smelter – the soil contamination in the study area is mainly diffuse and lower than 5 ppm Cd. The elevated Cd levels in the northern part of the Cranendonck municipality are due to the industrial Philips-Maarheeze site where commodities were processed to fluorescent powder.

Figure 2.8: Prediction map of soil Cd concentrations



The prediction map in Figure 2.8 represents the most likely values at all locations in the study area. One of the strengths of geostatistical interpolation methods is that they provide the opportunity to create a statistical measure of prediction uncertainty. Standard error maps can be used as a measure for possible variation and uncertainty associated with an estimated value. Therefore, prediction maps should preferably be accompanied by uncertainty surfaces to get an indication of the probability that the estimated value will equal the true value (Clauw, 2007). The standard error map in Figure 2.9 displays the uncertainty associated with the predictions in the Campine region. It can be observed that the standard errors tend to be the highest in locations with high predicted Cd values. This was expected since there is a larger variability in soil Cd samples close to the pollution sources. As a result, the standard error map resembles quite closely to the prediction map. At the borders of the research area, standard error values are somewhat higher as well due to the limited amount of sample measurements in these locations.

Figure 2.9: Standard error map



2.4.2 Farmland for phytoremediation

Results from the phytoremediation field experiment (Table 2.2) show that two crops – i.e. willow and tobacco – are capable of removing significant amounts of Cd from soils in the Campine region. Therefore, this analysis will solely focus on these two cultivations. In common willow cultivations for bioenergy production, the crop is typically harvested in winter time when leaves have fallen. However, in case of phytoremediation harvest should preferably be completed before leaf senescence, as leaves contain a considerable amount of Cd (Hammer et al., 2003; Keller et al., 2005). Since it is assumed that willow is harvested every three years, leaves can only be removed once in three years. For tobacco (T), a three-year rotation in combination with maize (M) is hypothesized, since this has been found to achieve the highest Cd removal (Schreurs et al., 2011).

Assuming uptake rates will remain constant throughout the entire time frame, phytoremediation using willow leads to meaningful decreases in soil Cd concentration. Within a time frame of 42 years, soils containing up to 2.88 ppm Cd can be remediated to the guide value of 1.2 ppm (Table 2.5). Although the rotation including tobacco and maize might be more acceptable to farmers, it is far less effective from a remediation point of view than continuous short rotation cropping. In case the T-M-M rotation is maintained for 42 years, soils with a maximal Cd concentration of 1.93 ppm can be brought back to guide values, while a willow cultivation can remediate soils up to 2.04 ppm in 21 years.

Table 2.5: Farmland suitable for phytoremediation

Cultivation	**Time span** (y)	**Maximal soil Cd concentration** (ppm)	**Farmland for phytoremediation** (ha)
Willow	21	2.04	1813.37
	42	2.88	2396.75
T-M-M	21	1.56	974.46
	42	1.93	1699.79

Combining the prediction maps with the ALV layer in GIS allows us to determine the agricultural land that can be dedicated to phytoremediation. It is assumed that parcels dealing with contamination levels lower than 1.2

ppm do not have to be remediated, so this will be the target contamination level. Then, the study area is divided into geographical subareas according to the maximal soil Cd concentrations (Table 2.5). For example, in case of willow cultivation for 21 years, the area that deals with (predicted) soil Cd concentration between the range of 1.2 ppm and 2.04 ppm is extracted. Subsequently, this subarea is intersected with the ALV layer in order to calculate the amount of farmland in this area. This way, it is found that more than 1800 ha of farmland can be remediated by means of willow within 21 years. Another 583 ha can be remediated by means of willow within 42 years. Since the level of Cd that can be remediated using tobacco and maize is considerably lower in comparison with willow, the potential farmland suitable for phytoremediation is smaller as well.

Overall – thus regardless of phytoremediation restrictions – 3027 ha of agricultural land exceeded the guide value of 1.2 ppm Cd in the Belgian part of the Campine region. More than 40% of the contaminated farmland, and approximately 9% of the complete farmland area in the study area, surpassed the threshold value of 2 ppm Cd. This result indicates that a considerable amount of farmland in the area deals with elevated Cd concentrations, which is a rationale for tackling the subject further from an economic point of view in later chapters.

2.5 Conclusion and discussion

This chapter has aimed to fulfill two objectives. First of all, it was necessary to obtain a better understanding of the extent and severity of the soil contamination in the Campine region. Therefore, spatial interpolation techniques are applied to a dataset of 11,885 Cd sample measurements to predict soil Cd levels throughout the entire study area. Geostatistical interpolation methods were found to be preferable over deterministic methods because they are based on mathematical as well as statistical functions and they allow for mapping the uncertainty associated with the predictions. More specifically, the ordinary kriging method was applied in this research. Since the exploratory spatial data analysis (ESDA) pointed out that the dataset was characterized by a positively skewed distribution with a small number of very high values, Cd values were transformed to its logarithm. Furthermore, the ESDA indicated that the data were likely to be stationary and there was considerable spatial dependence present in the data, particularly in samples close to the pollution sources.

Geostatistical methods incorporate the spatial structure of the dataset as a function of distances between observations. This function is fitted in an empirical semivariogram. Although there are a number of models capable of fitting the semivariogram cloud, the exponential model appeared to be most appropriate for this dataset based on multiple crossvalidation statistics. This model was then used to create prediction surfaces for the study area. As expected, the resulting map showed that the highest Cd concentrations were located in close proximity to the pollution sources and particularly in northern and eastern direction. This corresponds to what was expected based on wind current flows in Belgium, which predominantly originate from southwestern directions. The contamination levels might be somewhat underestimated in the vicinity of the Lommel Maatheide site, because a lot of sample measurements have been taken on the piece of land that has been remediated in 2003. However, since the time of sample measurement was unknown to us, it was impossible to exclude observations based on whether they were taken before or after the remediation. Moreover, the area to the north of the Maatheide site is labeled as nature zoning, which is not particularly relevant in this dissertation. Therefore, it is expected that this (potential) bias will not affect the results of the analyses in the following chapters.

The second goal of this chapter was to determine the amount of agricultural land that has been affected by soil contamination. By means of the prediction maps, it was found that more than 3000 ha of farmland in the Campine region exceeded the Cd guide value set by the Flemish

Government. More than 40% of this area even surpassed the threshold value of 2 ppm Cd. Since soil contamination in agricultural land can be handled by means of phytoremediation, we aimed to extract the amount of farmland which was eligible for remediation. Therefore, the range of soil Cd concentrations that can be remediated by phytoextraction was determined using multiple parameters. The target Cd concentration was set at the Flemish guide value of 1.2 ppm for agricultural land. Biological parameters such as uptake rates, biomass production, soil density and root depth are obtained from a phytoremediation field experiment recently conducted in the Campine region (Vangronsveld et al., 2009). Since it was hypothesized that farmers were only willing to dedicate land to phytoremediation if a clean status can be achieved within a practical time frame, periods of 21 and 42 years were modeled. The results showed that phytoremediation is capable of restoring 60% and 79% of contaminated farmland in respectively 21 and 42 years by means of willow cultivation. In case a rotation of tobacco and maize is used, the amount of farmland suitable for phytoremediation decreases considerably.

3 Economic burden of illnesses attributable to environmental pollution. Part I: Health risk assessment[14]

3.1 Introduction

Human exposure to Cd can have adverse effects on public health. The population's most important exposure pathways to Cd are pulmonary and gastrointestinal absorption (Jarup and Akesson, 2009)[15]. Besides occupationally exposed persons and people living close by to industries emitting cadmium, the primary source of inhalatory cadmium exposure is tobacco smoke (Adams et al., 2011). Tobacco leaves contain a significant amount of Cd, which is inhaled for approximately 10% when smoking cigarettes (Nordberg et al., 2007). About half of the cadmium content in the cigarette smoke is resorbed by the lungs (Godt et al., 2006). Although Cd contaminated dust can be another inhalatory exposure pathway, the effects on urinary Cd concentrations are rather small (Hogervorst et al., 2007).

For non-smokers taking in food containing cadmium is the most important source of exposure (Adams et al., 2011). There are numerous ways in which foodstuff can be contaminated with Cd. Frequently, this occurs when crops and vegetables are grown in soils holding elevated Cd concentrations. Some foodstuffs have the capacity to easily transfer metals from the roots to its edible parts (Satarug et al., 2010; Vromman et al., 2008). Especially

[14] Parts of this chapter have been presented at the 8th Conference on Sustainable Development of Energy, Water and Environment Systems (SDEWES): Schreurs, E., Cleemput, I., Nawrot, T. and Thewys, T. (2013). Economic impact assessment of environmental contamination: A damage function approach [Poster presentation]. SDEWES, Dubrovnik (Croatia), 22-27 September 2013.

[15] In theory, another exposure pathway is dermal contact with contaminated media like soil or water. This implies that cadmium is absorbed by the skin and enters the blood vessels. However, there has been little research on this topic and the amount of Cd which is absorbed by the skin is generally very small.

homegrown food sources can be a risk factor since people may be unaware of the presence or the risks of contamination in their gardens (Hough et al., 2003a). The highest Cd concentrations are found in leafy green vegetables, root vegetables, potatoes, cereals, seeds and nuts (Egan et al., 2007). Therefore, it has been argued that people with a vegetarian or high-fiber diet have an increased risk of dietary cadmium intake (Jarup and Akesson, 2009; Nawrot et al., 2010). Contaminated water only has a limited contribution to dietary cadmium intake (Olsson et al., 2002).

Since cadmium does not have any physiological or biochemical function within the human body (Van Cauwenbergh et al., 2000), exposure to Cd can be harmful to humans. In most cases the amount of metals present in the body is used as a biomarker for exposure to metals. For Cd, blood cadmium levels are a marker for recent exposure, given that the half-life of Cd in blood is 2 to 3 months (Welinder et al., 1977). Urinary Cd concentrations with a half-life of 10 to 30 years represent the lifetime exposure to Cd (Jarup and Akesson, 2009; Staessen et al., 1992). In numerous dose-effect and dose-response studies it has been examined whether there are functional relationships between these markers of exposure and the incidence of certain diseases.

These epidemiologic studies have generated evidence for a number of adverse health effects as a result of exposure to Cd. Earlier research particularly focused on kidney effects, seeing that this is the part of the body in which Cd accumulates. In extreme cases, this has been shown to result in itai itai disease, a combination of renal tubular dysfunctions and bone disorders (Inaba et al., 2005). This severe illness has been diagnosed in some areas in Japan where rice paddies have been contaminated with Cd (Kobayashi et al., 2009). However, also in case of more moderate exposure there have been some reports of an increased risk for tubular impairment (Akesson et al., 2005; Thomas et al., 2009).

More importantly, Cd exposure has also been associated with osteoporosis (James and Meliker, 2013), often in combination with kidney tubular damage reports (Jarup and Alfven, 2004). Osteoporosis is an age-related illness which is characterized by a decreasing bone density and a deterioration of bone tissue. This leads to a higher bone fragility and consequently a higher risk of fractures (Jarup and Akesson, 2009). One of the proposed mechanisms for this effect is that Cd-induced renal tubular damage leads to decreasing calcium reabsorption in the nephron and hence a lower bone mineral density (Staessen et al., 1994). Many epidemiological studies have shown that an elevated Cd body burden can result in a higher relative risk of fractures (Akesson et al., 2006; Alfven et al., 2000; Engstrom et al., 2011; Schutte et al., 2008; Staessen et al., 1999; Thomas et al., 2011).

Moreover, cadmium has been classified as a carcinogenic metal as well (Verougstraete et al., 2003). It has been shown that cadmium exposure increases all-cause mortality, particularly related to cancer (Menke et al., 2009). Exposure to Cd can cause multiple varieties of cancer, such as lung cancer (Nawrot et al., 2006), breast cancer (McElroy et al., 2006) and endometrial cancer (Akesson et al., 2008). Although there have been reports of effects on diabetes (Schwartz et al., 2003), cardiovascular diseases (Tellez-Plaza et al., 2013) and other illnesses (Nawrot et al., 2010; Satarug et al., 2010) as well, damages to the bone and skeletal system and cancer mortality seem to be the most important health effects due to Cd exposure.

In order to assess the health related impact on a societal level, it might be useful to economically evaluate the (health) damages environmental Cd pollution has generated in the Campine region. Epidemiologic research in the area (Nawrot et al., 2006; Staessen et al., 1999) has shown that the presence of increased environmental Cd levels has a detrimental effect on public health. In the next two chapters the health economic burden attributable to environmental pollution will be assessed. In chapter 3, a health risk assessment will estimate the number of illnesses which is attributable to Cd exposure, while the economic consequences are handled in chapter 4. This chapter is structured as follows. In section 3.2, the methodology that is adopted will be described. Data (sources) will be discussed in section 3.3, while the results of the risk assessment will be provided in section 3.4. In the last section, some concluding remarks and limitations of the analysis will be given.

3.2 Methodology

In identifying the monetary values of changes in human health due to an environmental effect, two links must be established. First, there is the link between the environmental change and the impact on public health. The second relationship has to be established between the change in health status and its monetary equivalent, usually expressed in terms of willingness to pay or willingness to accept compensation. Environmental economics methods provide two kinds of approaches that can be implemented for estimation of these effects. In the first approach a comprehensive model is developed in which individual behavior and choice is a function of the environmental change under consideration. This allows to directly elicit WTP estimates from this functional relationship. The second approach tries to set up these linkages separately. Hence, the economic values for a change in health status are determined, while the risk of health effects as a result of the environmental change is derived independently (Freeman, 1999). The most prevalent applications of both approaches are respectively the averting behavior method and the damage function approach, and more specifically the damage cost method for the 'health – economic value' relationship. Although they are often considered jointly because both of them try to value the underlying changes in pollution that cause changes in outcomes, both methods originate from a different background.

The averting behavior method refers to the actions that people are taking to reduce their exposure to environmental pollutants as well as to the actions to mitigate adverse effects of exposure. The method assumes that a rational person will pose defensive behavior and make defensive expenditures as long as his/her personal assessment of damage costs avoided is larger than the cost of the defensive action. Hereby, individuals implicitly reveal that the benefits derived from avoiding the environmental damage exceed that expenditure or at least equal it at the margin (Garrod and Willis, 1999). Besides defensive expenditures people can also change their behavior or time allocation in order to adapt to the environmental effect. In this case, the opportunity costs associated with changed behavior or time use have to be quantified first before inferences can be made. From the expenditure patterns of people exposed to environmental damage, their willingness to pay for avoiding it can be inferred (Dickie, 2003; Young, 2005). Since economic values are derived from actual observable behavior, the method is classified as a revealed preference technique.

Unlike methods focusing on behavioral changes, the damage cost method establishes a physical linkage between the environmental effect and the willingness to pay for avoiding the effect (Callan and Thomas, 2010). It measures all direct and indirect resource costs that are effectively brought about by the environmental effect. This way, these damage costs can be used to estimate the benefits of reducing environmental pollution (Young, 2005). As mentioned before, the damage cost method is merely the second stage of a complete damage function approach (DFA). In the first stages of the DFA epidemiologic research has to determine the functional relationship between the environmental effect and the associated (health) damages while controlling for socioeconomic influences and other confounding variables such as diet, life style and occupational exposures to harmful substances (Freeman, 1999).

Dickie (2003) asserts that there are two main differences between averting behavior and damage cost methods. First, while the former method focuses on the way human behavior responds to changes in environmental effects, the latter method implicitly excludes the effect of behavioral responses to these changes. Secondly, the averting behavior method is capable of estimating – or at least bounding – theoretically consistent measures of economic value such as WTP. In the theoretical model that is put forward, he demonstrates that damage costs will not correspond to WTP because (1) damage costs only focus on the resource costs and do not include the utility value of changes in service flows (for example, the WTP to avoid pain and suffering from health effects caused by pollution) and (2) damage costs do not take into account the changes in defensive behavior, while empirical evidence indicates that defensive behavior is not a static matter but increases with pollution. This implies that the true economic values of changes in environmental quality depend on the degree of pollution. Freeman (1999) argues that from a conceptual perspective the damage cost method will return only partial measures of the total WTP to avoid adverse health effects. However, when we dig deeper into the assumptions and data requirements for determining theoretically WTP estimates using revealed preference methods, it can be observed that they are quite burdensome to say the least. When damages in a health context are concerned, they are actually nearly impossible to comply with. Let me elaborate.

For averting behavior methods to determine theoretically consistent WTP estimates, three conditions need to be fulfilled (Holland et al., 2010): (1) defensive expenditures must be voluntary, (2) costs must be actually paid by the affected persons and (3) there must be no joint production of additional (dis)utility associated with the defensive behavior. The first condition assumes that people have not been urged or enforced to take measures to protect themselves from harmful exposure. When our case study is put forward as an illustration, the population should have taken the initiative themselves for not cultivating any vegetables in their gardens anymore, not because the government urged them to. Secondly, the defensive expenditures must come solely at the expense of the affected individuals. Preventive actions and costs made by other parties which are not affected by the environmental degradation need to be left out of the equation. For our case study, this implies that the costs for implementing strict environmental regulations, which are borne by the government and the polluting companies, cannot be taken into account.

Especially the last requirement is a major obstacle in using the averting behavior method for estimating the effect of changes in environmental quality (Dickie, 2003). It is particularly hard to achieve since practically all defensive actions will provide some additional utility or disutility to people. On the one hand, defensive behavior might have an impact on multiple health outcomes. If the number of symptoms exceeds the number of mitigating and defensive actions, the marginal WTP will include some unobservable marginal utility terms and cannot be correctly estimated using the averting behavior approach. Suppose, for instance, that individuals in the Campine used to stay indoors due to the elevated Cd levels in the air. As we will see later on in this chapter, this would have reduced their probability of incurring more than one illness. On the other hand, the defensive actions might enter the utility function directly as well. If the population in the affected area was urged to refrain from eating homegrown vegetables, these people needed to purchase vegetables elsewhere. Besides the extra expenditures that might be required for purchasing these substitution goods, people might experience some disutility from the fact that they will now need transportation in order to obtain their vegetables.

Averting behavior methods also assume that the damage avoided would otherwise be dealt with in a competitive market situation. However, there are a number of market imperfections that are very likely to distort the market in a situation of environmental degradation. First of all, the exposed population is assumed to have perfect information on all aspects in which the environmental effect will affect their utility. This entails, for example, that people are aware of the additional disease risk they are running due to the environmental exposure and adjust their behavior accordingly. However, the literature on people's risk assessment suggests that subjective risks might deviate considerably from objective risks (McClelland et al., 1990). Furthermore, this assumption also implies that exposed individuals are properly informed on the costs associated with the damage, which brings us to the next market imperfection. In the health care market there are no competitive market mechanisms between demand and supply at play because the market is completely dominated by government intervention[16]. Current prices are the result of a complex negotiation between the government, doctors, hospitals and other stakeholders in the health care system. Governments have developed the health care system in such a way that social mechanisms have been integrated which redistribute the economic burden of illnesses across the entire population. Health care insurance shifts (some) direct and indirect medical costs away from the sick individual, while the burden of sick leave from work is (partly) compensated by the employer and the government. This way, people only have to pay a small fraction of the effective resource costs brought about by the illness directly[17].

This will have its repercussions in the application of averting behavior methods as well. The basic assumption of the method is that a rational person will pose defensive behavior and make defensive expenditures as long as his personal assessment of damage costs avoided is larger than the cost of the defensive action. However, an individual's assessment of damage costs avoided will only entail the costs that are at their own expense. When health damages are concerned, people will only take into account the medical costs that have to be paid out-of-pocket and not the economic burden that is suffered by society as a whole when making rational choices concerning the defensive expenditures. The total, societal value of preventing an illness is thus likely to exceed the aggregation of all individual WTP estimates. The damage cost approach, on the other hand, is capable of taking these public expenditures into account in order to put the societal value of preventing an illness into perspective. However, in using this approach it must be considered that the health care market is distorted by government intervention and asymmetric information.

All in all, the inability to determine theoretically consistent measures of WTP does not give sufficient weight to deny the damage function approach in

[16] See section 4.3 for a brief description of the health care system in Belgium.
[17] Of course, people contribute indirectly to the health care system and the social security system in general by paying taxes.

favor of the averting behavior method. The latter approach deals with such substantial restrictions and assumptions that deriving correct WTP estimates will be very hard as well, especially in case health damages are concerned. Therefore, the damage function approach will be adopted in this analysis to estimate the relationship between environmental pollution levels and the economic consequences of the physical damages it causes. Although these physical damages can involve damages to crops or structures as well, we will focus on the effects on public health. Research using DFA is generally divided into three stages (Table 3.1). In the first stage the functional relationship between pollution levels and the physical damage is estimated. When health issues are concerned, an epidemiologic research design will to try to unravel this relationship by establishing risk or odds ratios for detrimental health effects. In the second stage the damage coefficient – estimated in the first step – will be used to estimate the impact of the contamination on a population level. A health risk assessment will determine the amount of people that have been exposed to elevated pollution levels in order to estimate the number of people that have experienced adverse health effects. In the last stage, the economic consequences of these health effects are estimated. The methodology for producing cost estimates will be extensively discussed in the next chapter (see section 4.2).

Table 3.1: Stages of the damage function approach

Stage	**Objective**	**Method**
1	Relative risk	Epidemiologic research
2	Attributable cases	Health risk assessment
3	Economic evaluation	Damage cost/Cost of Illness

The decision to adopt a DFA in this study is partly based on the fact that a significant amount of epidemiologic research in the area has already been done in the past. Seeing that this stage is the most time consuming and data intensive part of a DFA, the amount of research to be done in order to estimate the damage costs is relatively small in comparison with earlier work. Moreover, as far as the remaining research – i.e. the second and third stage of DFA – goes, environmental economists are able to handle this research if assisted by an interdisciplinary team of epidemiologists, risk analysts and health economists. In the first stage of the DFA, however, the contribution of economists would have been quite limited.

The epidemiologic research performed in the past has investigated the correlation between environmental pollution and illnesses by establishing relative risks (RR), the ratio between an event (illness) occurring in exposed and non-exposed situations. These RR will be used as a starting point for the health risk assessment in this chapter. The population-attributable fraction (AF) can be derived from the RR using the following formula (Baker and Nieuwenhuijsen, 2008):

$$AF = \frac{RR-1}{RR} \qquad [3]$$

Subsequently, the number of people exposed to environmental pollution as well as the degree of pollution they have been exposed to, will be determined by means of an exposure assessment. The research area will be separated into geographic clusters *i* according to the exposure levels, ranging from no exposure to high Cd exposure. The RR and the size of the population will be determined for each cluster separately. In last instance, the standard incidence of the illnesses in Belgium needs to be multiplied with the RR in each cluster in order to estimate the number of pollution-attributable (PA) illnesses:

$$PA\ illnesses = \sum_i AF_i \times Population_i \times (Standard\ incidence \times RR_i) \qquad [4]$$

3.3 Data

3.3.1 Epidemiology

The epidemiologic research performed in the past comprised three study periods, i.e. Cadmibel (1985-1989), Pheecad I (1991-1994) and Pheecad II

(1985-2004). The Cadmibel study was a cross-sectional population study that collected data on the people's environmental exposure to Cd, Cd body burden, blood pressure and renal tubular function. Two population samples were defined, i.e. a sample living in six districts close to the three pollution sources (high exposure area: HEA) and a sample living in four districts that were located more than 10 km from the pollution sources and were not affected by Cd (low exposure area: LEA). The two samples had similar baseline characteristics, except for internal and external exposure to Cd, and were approximately the same size (n = +/-550).

Following up on the Cadmibel findings, Pheecad I was a longitudinal study focusing on bone mineral density and risk of fractures. In this study the Cadmibel participants that still lived in the area were contacted once again and invited for measurement of their forearm bone density. Although 60% of the Cadmibel sample responded positively, more than 100 men were excluded because they reported to be exposed to heavy metals at work. In the end, the Pheecad I analysis included 506 participants (Staessen et al., 1999). In the second Pheecad study, the incidence of cancer in Cadmibel participants was followed up for almost 20 years. Participants who had a history of illnesses (cancer or pneumoconiosis) that increased the risk of lung cancer or from whom no urine samples were obtained, were excluded from analysis. The study population consisted of 521 people for the HEA and 473 people for the LEA. Since smoking is a major source of exposure to Cd in the general population, every person that has ever used smoking materials was classified as a smoker. The proportion of smokers in the HEA and the LEA were quite similar (Nawrot et al., 2006)[18].

Positive and significant relationships were found between different indicators for environmental exposure to Cd and the incidence of fractures (Table 3.2). The study controlled for several potentially confounding factors such as smoking, alcohol intake, oral contraceptives, hormone replacement therapy

[18] For more details regarding the research design and the results, we refer to Lauwerys et al. (1990) for Cadmibel, Staessen et al. (1999) for Pheecad I and Nawrot et al. (2006) for Pheecad II.

and supplements of calcium or vitamin D, but none of them entered any of the models. The relative risks for fractures in women have only been adjusted for age, while no significant covariates were found for men. The results of the lung cancer study (Table 3.3) have been corrected for confounding factors age, sex and smoking. Since occupationally exposed persons were faced with a higher Cd body burden and an increased risk of cancer, these 42 individuals were excluded for determining the relative risks in Table 3.3. Also when other ways of controlling for smoking were used (e.g. by means of the number of pack-years), the results remained consistent. The population-attributable risk of lung cancer was 67% in the HEA, compared to 73% for smokers.

Table 3.2: Relative risk (95% CI) of fractures (Staessen et al., 1999)

	Men		**Women**	
	Fracture	**p**	**Fracture**	**p**
Baseline Cd excretion (nmol per day)*	1.2 (0.75-1.93)	0.44	1.73 (1.16-2.57)	0.007
Residence in polluted area (0-1)	2.76 (1.07-7.13)	0.036	4.3 (1.77-10.4)	0.001
Cd in soil (mg/kg)*	1.39 (1.04-1.86)	0.024	1.54 (1.19-2.00)	<0.001
Cd in leek (mg/kg dry weight)*	1.93 (1.05-3.93)	0.034	2.27 (1.31-3.94)	0.004
Cd in celery (mg/kg dry weight)*	1.69 (1.02-2.79)	0.039	2.07 (1.31-3.27)	0.002

* Relative risk associated with a doubling of the Cd concentration

Table 3.3: Relative risk (95% CI) of lung cancer (Nawrot et al., 2006)

	Lung cancer	**p**
24-h urinary cadmium excretion (nmol/day)*	1.73 (1.09-2.72)	0.019
Residence in polluted area (0-1)	3.58 (1.00-12.7)	0.049
Cadmium in soil (mg/kg)*	1.49 (1.04-2.14)	0.032

* Relative risk associated with a doubling of the Cd concentration

3.3.2 *Exposure assessment*

Internal biomarkers of Cd – i.e. urinary and blood Cd concentrations – are generally the best indicators for exposure to Cd. Since these biomarkers have only been measured in the HEA and the LEA and not in the entire contaminated area, it is not possible to use these RR in the health risk assessment. However, one of the external biomarkers of environmental exposure – i.e. soil Cd concentrations – *can* be extrapolated to the entire area, allowing us to extend the health risk analyses to the population level. It has been acknowledged before that the degree of cadmium in soil is an appropriate predictor for the amount of environmental pollution people have been exposed to in the Campine region (Wildemeersch, 2008). Using the prediction maps for soil Cd concentrations created in chapter 2, the contaminated area can be partitioned into risk areas according to the degree of soil contamination. Since the relative risks associated with soil Cd concentrations (Table 3.2 & 3.3) specify the extra risk of incurring an illness in case the amount of cadmium in soil is doubled, each pollution level can be assigned its own risk index.

The population living in each risk area is determined by means of the Central Reference Address File (CRAB), a geographic layer managed by the Flemish government and its municipalities. The layer contains all residential address points located in Flanders. If it is assumed that each address point represents one household, intersecting the layer with the risk areas results in the number of housing units per risk area. Subsequently, the average number of people per household will be calculated by dividing the population in the research area by the total amount of address points. The exposed population is estimated by multiplying the addresses in the risk areas with the average number of people per household.

3.3.3 *Incidence rates*

Incidence rates of lung cancer and fractures in Belgium are needed in order to estimate the number of cases attributable to environmental pollution. For each illness the most recently available incidence rates are provided. Lung cancer data are obtained from the Belgian Cancer Registry (BCR), which registers all cases of cancer that have been diagnosed in Belgium. Since 1999 the BCR publishes cancer incidence rates for Flanders on a yearly basis (Table 3.4) and since 2004 for all regions in Belgium. Data for lung cancer incidence in Flanders in 2010 will be adjusted to fit the local population structure with regard to gender.

Table 3.4: Lung cancer incidence* (per 100,000 inhabitants) in Flanders

Crude rate	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Men	110.0	107.0	108.3	110.5	106.5	105.8	108.5
Women	27.2	28.2	28.8	32.1	30.7	34.2	37.7

* Total number of newly registered diagnoses of lung cancer across the general population, no (age-)standardization has been applied.

With respect to fractures there are three major arguments to focus solely on hip fractures. Firstly, Staessen et al. (1999) has shown that bone mineral density is negatively correlated with urinary Cd excretion, linking elevated Cd exposure to osteoporosis. The most common osteoporotic fractures are hip, wrist or vertebral fractures. In case of hip fractures, osteoporosis accounts for nearly all fractures, particularly when individuals older than 50 are concerned (Reginster et al., 2001). Wrist and vertebral fractures, on the other hand, often result from traumatic events as well. Therefore, it would have been necessary to differentiate incidence data based on the cases which are attributable to osteoporosis. As a second argumentation, hip fractures have been shown to produce the largest economic burden in comparison with wrist and vertebral fractures (Borgstrom et al., 2006). Since we aim to determine the health care costs attributable to the environmental pollution, resource costs produced by hip fractures are likely to be the most important contributor to this objective. Lastly, hip fractures are the only common osteoporotic fractures for which data are collected on a regular basis in Belgium (e.g. in the National Health Inquiry). Moreover, hip fracture incidence in Belgium has been published in international peer-reviewed journals by researchers at University of Liège (Gillet et al., 1995;

Hiligsmann et al., 2008a; Reginster et al., 2001). The most recent data are available for the year 2007 (Hiligsmann et al., 2011a), so these data will be used in our analysis (Table 3.5).

Table 3.5: Annual absolute number of hip fractures and age-standardized* incidence of hip fracture per 100,000 for Belgian population (≥ 50 years), over a 8 year period (2000-2007) (Hiligsmann et al., 2011a)

Year	**Absolute number of hip fractures**			**Standardized incidence (per 100,000 persons older than 50)**	
	Women	**Men**	**Total**	**Women**	**Men**
2000	10,349	3164	13,512	560	215
2001	10,517	3326	13,843	564	220
2002	10,671	3404	14,074	563	220
2003	10,787	3423	14,210	562	216
2004	10,560	3443	14,003	543	213
2005	10,682	3602	14,285	538	216
2006	10,754	3675	14,429	527	212
2007	10,937	3808	14,744	522	213
2000-2007	85,256	27,845	113,101	548	216

* The year 2005 was used as reference for standardization.

3.4 Results

In order to determine the population exposed to elevated Cd levels, it is important to first establish the exposure level for which there are supposed to be no extra illnesses. Actually, this implies that the area of low exposure – for which the RR is expected to equal 1 – has to be redefined in terms of soil Cd levels. For simplicity, it is assumed here that the background (low level) exposure to cadmium corresponds to a soil Cd concentration of 1 ppm. In the LEA defined in the epidemiologic studies, soil Cd levels were on average 0.81 ppm (Nawrot et al., 2006). However, this LEA was located in Hechtel-Eksel, more than 10 km from the pollution source. In our study area the average distance to the pollution sources is smaller and thus the average soil

Cd level is expected to be somewhat higher. Moreover, according to the soil thresholds set by the Flemish Government, soil Cd levels of 1 ppm are still well within the acceptable range of soil concentrations without health effects (see Table 1.2).

Assuming that the health effects are linear on a population level, the relative risks can be extrapolated to more highly exposed areas, i.e. areas with higher soil Cd levels. Therefore, the contaminated area is categorized into risk areas according to predicted soil Cd concentrations[19]. For example, the risk area with a soil Cd level of 2 ppm – a doubling in comparison with the background exposure – coincides with a relative risk of 1.49 for lung cancer. This procedure can be followed to determine the RR for all superjacent soil Cd levels and for each illness. However, an upper threshold for environmental exposure is imposed at a soil Cd level of 8 ppm. Soil Cd concentrations higher than 8 ppm are exceptional in the area and can only be found in the direct vicinity of the pollution sources (see chapter 2). Further categorization would create rather marginal risk areas. Moreover, the average soil Cd concentration in the HEA was 7.97 ppm (Nawrot et al., 2006), so the 8 ppm level approximately corresponds to the soil Cd level in the HEA. This seems to be a plausible assumption, since the (calculated) RRs at the 8 ppm level (Table 3.6, 3.7 and 3.8) correspond quite closely to the RRs in the HEA (Table 3.2 and 3.3). For lung cancer, the calculated RR was 3.31, compared to 3.58 in the HEA. For hip fractures, the calculated RR was 2.69 in men and 3.65 in women, while the RR in the HEA were 2.76 and 4.30, respectively. The calculated RR appear to somewhat underestimate the actual RR in the HEA. The AFs are then derived from the RRs at each contamination level using equation [3].

The number of housing units per risk area is determined by means of the CRAB layer. Multiplying with the average number of people per household results in population estimates per risk area. For lung cancer, population data for 2010 are used because this corresponds to the most recently

[19] Intermittent soil Cd levels are rounded to integer values.

available year for cancer incidence rates. Incidence rates from Table 3.4 are weighted by means of the local population's structure with regard to gender. Hip fracture incidence for the population aging older than 50 is reported in Hiligsmann et al. (2011a) with data from 2007. Hence, population data from 2007 are employed. Moreover, the cost estimates for lung cancer are based on data from 2007-2011, while the hospitalization costs for hip fractures originate from 2007 (see chapter 4).

Table 3.6 shows that the yearly number of lung cancer cases attributable to environmental pollution amounts to approximately 22 in the Campine region. For hip fractures, there are expected to be 8.5 cases in men (Table 3.7) and 31 cases in women (Table 3.8) attributable to environmental pollution on a yearly basis. There are substantially more cases of hip fractures in women in comparison with men, which can be partly explained by the higher relative risks and partly by the higher hip fracture incidence in women.

Table 3.6: Lung cancer cases attributable to environmental pollution

Soil Cd level	RR lung cancer (95% CI)	AF lung cancer (95% CI)	Population* (in 2010)	Incidence** (per 100,000 persons)	Pollution-attributable cases per year***
1	1.00	/	105,086	73.3	/
2	1.49 (1.04-2.14)	33% (4-53)	22,339	109.2	8.02 (0.94-12.99)
3	1.88 (1.06-3.34)	47% (6-70)	10,597	137.9	6.85 (0.83-10.24)
4	2.22 (1.08-4.58)	55% (8-78)	3273	162.7	2.93 (0.39-4.16)
5	2.52 (1.10-5.85)	60% (9-83)	1059	185.0	1.18 (0.18-1.62)
6	2.80 (1.11-7.15)	64% (10-86)	818	205.4	1.08 (0.17-1.45)
7	3.06 (1.12-8.46)	67% (10-88)	433	224.5	0.65 (0.10-0.86)
8	3.31 (1.12-9.80)	70% (11-90)	882	242.4	1.49 (0.23-1.92)
				Total amount:	**22.20 (2.84-33.24)**

* The population per risk area is calculating by multiplying the number of housing units per risk area (found using the CRAB layer) with the average number of people per household in 2010.
** Standard lung cancer incidence (73.3 cases per 100,000 persons) has been weighted by the local population's structure with regard to gender (Table 3.4).
*** See equation [4].

Table 3.7: Hip fracture cases in men attributable to environmental pollution

Soil Cd level	RR hip fracture (95% CI)	AF hip fracture (95% CI)	Population* (male, >50, 2007)	Incidence** (per 100,000 persons older than 50)	Pollution-attributable cases per year***
1	1.00	/	17,860	213.0	/
2	1.39 (1.04-1.86)	28% (4-46)	3797	296.1	3.15 (0.43-5.20)
3	1.69 (1.06-2.67)	41% (6-63)	1801	359.0	2.63 (0.37-4.04)
4	1.93 (1.08-3.46)	48% (8-71)	556	411.5	1.10 (0.17-1.63)
5	2.15 (1.10-4.22)	53% (9-76)	180	457.6	0.44 (0.07-0.63)
6	2.34 (1.11-4.97)	57% (10-80)	139	499.0	0.40 (0.07-0.55)
7	2.52 (1.12-5.71)	60% (10-82)	74	536.9	0.24 (0.04-0.33)
8	2.69 (1.12-6.43)	63% (11-84)	150	572.0	0.54 (0.09-0.72)
				Total amount:	**8.50 (1.25-13.10)**

* The population per risk area is calculating by multiplying the number of housing units per risk area (found using the CRAB layer) with the average number of men per household in 2007.
** Age-standardized hip fracture incidence for men in 2007 (Table 3.5) is used as the starting point.
*** See equation [4].

Table 3.8: Hip fracture cases in women attributable to environmental pollution

Soil Cd level	RR hip fracture (95% CI)	AF hip fracture (95% CI)	Population* (female, >50, 2007)	Incidence** (per 100,000 persons older than 50)	Pollution-attributable cases per year***
1	1.00	/	18,692	522.0	/
2	1.54 (1.19-2.00)	35% (16-50)	3973	803.9	11.20 (5.10-15.97)
3	1.98 (1.32-3.00)	50% (24-67)	1885	1034.9	9.67 (4.73-13.00)
4	2.37 (1.42-4.00)	58% (29-75)	582	1238.0	4.17 (2.13-5.40)
5	2.73 (1.50-5.00)	63% (33-80)	188	1422.6	1.70 (0.89-2.14)
6	3.05 (1.57-6.00)	67% (36-83)	146	1593.7	1.56 (0.84-1.93)
7	3.36 (1.63-7.00)	70% (39-86)	77	1754.3	0.95 (0.52-1.16)
8	3.65 (1.69-8.00)	73% (41-88)	157	1906.5	2.17 (1.22-2.62)
				Total amount:	**31.41 (15.44-42.23)**

* The population per risk area is calculating by multiplying the number of housing units per risk area (found using the CRAB layer) with the average number of women per household in 2007.
** Age-standardized hip fracture incidence for women in 2007 (Table 3.5) is used as the starting point.
*** See equation [4].

However, it can be observed that there is a significant amount of variation associated with these estimations. One of the reasons for this is that the estimations are based on relative risks which resulted from epidemiologic studies that were completed a number of years ago. The Pheecad study by Staessen et al. (1999) ended 18 years ago, while the Pheecad II study by Nawrot et al. (2006) ended 8 years ago. The question arises whether the relative risks have remained constant until the time frame 2007-2010, since the population's exposure levels to Cd might have changed in the meanwhile. As mentioned before, the aerial Cd levels in the area were brought down to insignificant levels in the middle of the 1980s. Subsequently, prevention campaigns to inform the population about the risks of cultivating vegetables in Cd contaminated soils at home have been carried out as well. Therefore, it seems to be rather likely that the environmental exposure to Cd has decreased in recent years.

In 2006 the Flemish Government ordered a follow-up study with regard to the pollution issues in the Campine region. The findings in its final report (Wildemeersch, 2008) are quite helpful in this setting[20]. In this study the population's exposure to heavy metals in the area was measured once again and a exposure model was developed in order to find out in which way the population's body burden has evolved over time. The results with respect to cadmium show that internal biomarkers for exposure (i.e. blood and urinary Cd concentrations) in the HEA are still somewhat higher than in the LEA. The population's urinary Cd concentrations exceed the critical value of 1 µg/g creatinine more often in the HEA (17%) in comparison with the LEA (11%). Aerial Cd concentrations, however, are well below critical thresholds in the entire area.

The exposure pathways can be linked to the population's internal Cd levels by means of an exposure model. Model estimates show that in 2007 daily Cd intake is mainly due to Cd in food, particularly non-homegrown foodstuff in this case (Figure 3.1). The contribution of smoking actively is rather

[20] Unfortunately, the study has not been published in the international scientific literature and is therefore only available in Dutch.

moderate because it concerns the average contribution over the entire sample. The sample includes 53% non-smokers, 29% past smokers and 18% current smokers, distributed similarly across the study areas. With respect to cumulative Cd levels (Figure 3.2), the same factors that were important contributors to blood Cd concentrations (i.e. food and smoking) come into play. Besides these factors, indoor Cd concentrations and ingestion are substantial contributing factors as well. The indoor Cd concentrations are related to dust particles containing Cd elements, possibly originating from blown up soil particles with elevated Cd concentrations.

Figure 3.1: Exposure model recent Cd intake (Wildemeersch, 2008)

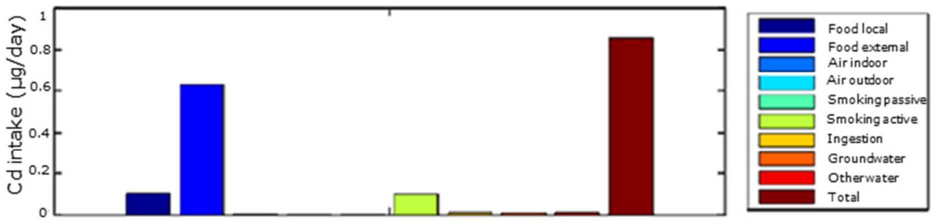
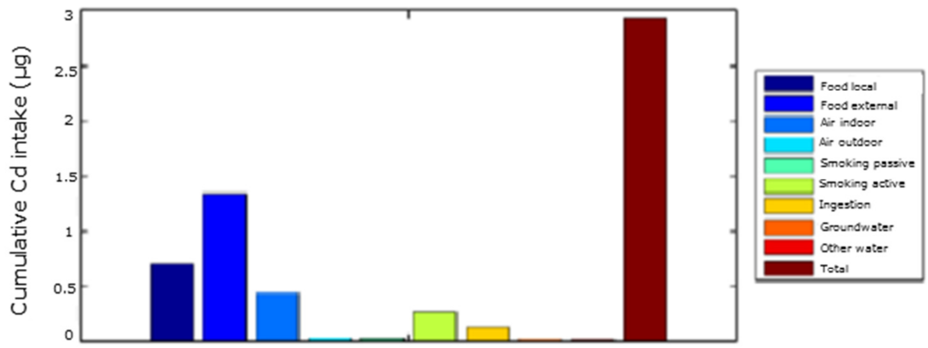


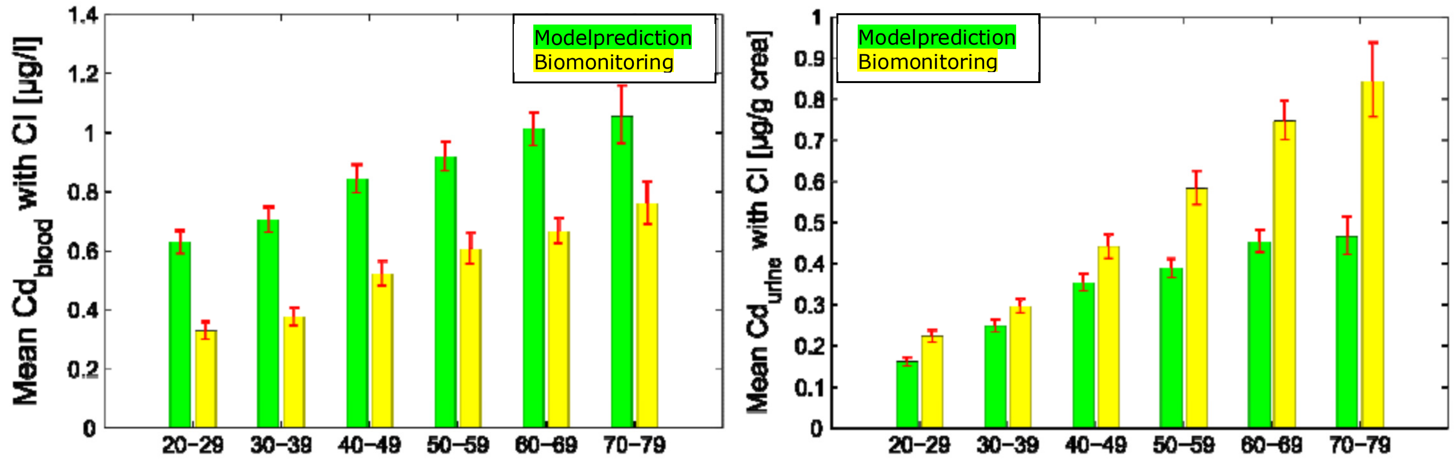
Figure 3.2: Exposure model cumulative Cd intake (Wildemeersch, 2008)



Although the exposure model provides results that are consistent with biomonitoring data, model prediction levels for Cd in blood are on average overestimated by 95%, while urinary Cd levels are underestimated by 17%. When the results are shown in more detail (Figure 3.3), it can be seen that the model consistently overestimates blood Cd concentrations for all age groups. However, urinary Cd concentrations are particularly underestimated in the older population segments, indicating there are still significant amounts of Cd accumulated in this population segment. As mentioned before, cadmium accumulates in the body over the years and the trace element has an exceptionally long half-life of 10 to 30 years in urine. Probably, environmental Cd concentrations that the population has been exposed to in the past were higher than assumed so far in the model. Nevertheless, the exposure model demonstrates that health effects are likely to decrease in the future since the percentage of the population that has been exposed to the historically elevated Cd concentrations will steadily decrease.

Since particularly the older population segment still seems to deal with a significant Cd body burden, the main results of this chapter – i.e. the pollution-attributable illnesses – should focus on this population segment. However, this is incorporated implicitly in the analysis by means of the incidence rates, because the considered illnesses are generally occurring in later stages of life. It can be observed in the Belgian and Flemish incidence rates that age-standardized incidence reaches its peak in the age interval of 75-80 years for lung cancer and 80-85 years for hip fractures and that both illnesses are only occurring exceptionally in people younger than 45 years old (data not shown). Moreover, for hip fractures the health risk assessments only includes the population over 50 years old.

Figure 3.3: Exposure model Cd in blood and Cd in urine (Wildemeersch, 2008)



In light of the findings in Wildemeersch (2008), the assumption of constant RR seems to provide an upper threshold for the amount of pollution-attributable illnesses in the Campine region. The population's exposure to cadmium has decreased considerably over time and is very likely to decrease further in the future, since the most important Cd exposure pathways have been broken. The fact that blood and urinary Cd concentrations are still somewhat higher in the HEA is for the largest part due to higher exposure levels in older population segments, which have been exposed to high aerial Cd concentrations in the past. At present, the actual amount of pollution-attributable illnesses will probably be situated within the lower range of the estimations given in Tables 3.6, 3.7 and 3.8. In general, it can be hypothesized that the constant RR estimations were particularly applicable in former years and decades. In recent years and in the future, as the average Cd body burden in the population is declining, the attributable cases will converge to the lower limit of these ranges.

3.5 Conclusion and discussion

Epidemiologic research has shown that human exposure to critical doses of Cd can cause a number of adverse health effects. For example, Cd exposure has been associated with a decreasing bone density, even when low-level exposure is concerned. Consequently, many epidemiologic studies have shown that an elevated Cd body burden results in higher relative risks of fractures. Since cadmium has been identified as a carcinogenic metal, Cd exposure increases cancer incidence and all-cause mortality as well. More specifically for the Campine region, research by Staessen and Nawrot indicated that the presence of elevated environmental Cd levels in the area has resulted in an increased risk of osteoporotic fractures and lung cancer.

The objective of chapter 3 and 4 is to estimate the damage costs attributable to the environmental Cd pollution in the Campine region by means of a damage function approach (DFA). The first stage of the DFA was already carried out by the epidemiologic research performed in the past. The second stage, which entails a health risk assessment with the aim of determining the amount of pollution-attributable illnesses on the population level, is the subject of this chapter. In the third stage (chapter 4), the damage cost method will be used to estimate the costs related to those illnesses. Unlike revealed preference techniques, the damage cost method establishes a physical linkage between the environmental effect and the economic consequences. The primary weakness of the damage cost method is the lack of incorporating human responses to adverse environmental effects, which prevents the method from revealing theoretically consistent WTP estimates. However, it has been shown in this chapter that this is nearly impossible to determine using revealed preference methods such as the averting behavior method as well, seeing that the required assumptions are very hard to fulfill. Moreover, the DFA allows for an integration of societal burden brought about by health damages, while the averting behavior method is incapable of dealing with imperfect markets such as the health care market.

For the health risk assessment, we relied on the relative risks provided by the first stage of the DFA. In an exposure assessment the amount of people that have been exposed to elevated pollution levels has been determined. In order to estimate to which extent the population has been affected by Cd, predicted soil Cd values from chapter 2 are used, because these are available for the entire research area while internal biomarkers of Cd exposure were only administered in the HEA and the LEA. The number of people that have experienced adverse health effects attributable to the pollution was estimated by means of the most recently available incidence data from Belgium or Flanders.

More than 20 years after the principal Cd exposure pathway has been interrupted, the population in the Campine region – particularly the older segment – still suffers from relatively high urinary Cd concentrations, which is a biomarker for long term (cumulative) Cd exposure. Since the epidemiologic studies have demonstrated that elevated urinary Cd concentrations are associated with the incidence of osteoporotic fractures and lung cancer, it is reasonable to expect that there are still illnesses attributable to the pollution at this moment. If it is assumed that the relative risk factors have remained constant up to the time frame 2007-2010, there are approximately 22 cases of lung cancer, 8.5 hip fractures in men and 31 in women attributable to the pollution in the Campine region.

The main objective of this chapter was to extrapolate the results of the cohort studies by Staessen and Nawrot to the population level in order to be able to economically evaluate the health damages generated by the pollution in the next chapter. However, this damage function approach requires making some assumptions, which we have tried to deal with as correctly as possible. One of the principal assumptions with regard to the outcomes is that the RR have remained constant so far. Given the results from the follow up study by Wildemeersch (2008), it seems to be unlikely that this is effectively the case. Therefore, it is assumed that the exact amount of pollution-attributable illnesses at present is actually situated within the broad range that is obtained when using the RR's 95% confidence interval's lower limit. This way, the uncertainty with regard to the persistence of the relative risks is (to some extent) taken into account. Moreover, the resulting intervals show there still is expected to be a non-negligible amount of illnesses attributable to pollution, especially hip fractures in women. Other assumptions that have been made in this research include the linear distribution of relative risks across the entire contaminated area and the lack of incorporating population migration over the years.

At this moment there are fairly little actions that can be adopted in order to reduce the amount of pollution-attributable illnesses, besides continuing the efforts that were put in place before. The aerial Cd concentrations, for instance, are currently well below critical thresholds in the Campine region. The only environmental medium that still contains elevated Cd levels is the soil. However, according to the exposure model put forward in Wildemeersch (2008), the contribution of soil Cd to recent Cd intake is quite limited. If it is assumed that indoor aerial Cd concentrations partly emanate from soil particles, the contribution to cumulative Cd intake is somewhat larger. The largest part of Cd exposure originates from food – homegrown and external – and smoking. Therefore, removing the remaining Cd pollution from (agricultural) soils will have an impact on the environmental exposure to Cd and hence on the pollution-attributable illnesses, but the effect will be considerably less significant than breaking the former exposure pathways.

Nevertheless, this research has its merits in showing that a lack of strict environmental policies can have some long-lasting public health effects. Polluting elements can accumulate in the body or change the genetic composition of individuals. Moreover, external effects of ineffective regulations seem to primarily affect the lower social classes in the population. This research can therefore serve as an example and even more so as an incentive for governments to implement rigorous environmental regulations as an instrument to increase and redistribute social welfare for its population. In this setting, it is important to enhance environmental epidemiologic research in order to get a better insight into the way in which pollutants are affecting public health. In case an integrated and multidisciplinary research framework is adopted, the damaging health effects of contamination can be quantified as to raise public support for environmental regulations.

4 Economic burden of illnesses attributable to environmental pollution. Part II: Health economic assessment[21]

4.1 Introduction

This chapter advances on the results obtained in the previous chapter and handles the third stage of the damage function approach. This entails that the economic consequences of the health damages caused by the environmental pollution in the Campine region will be determined. The health care costs of the two considered illnesses, i.e. lung cancer and osteoporotic hip fractures, will be estimated and multiplied with the number of pollution-attributable cases found in the previous chapter. This way, the damage costs associated with environmental pollution in the Campine region can be analyzed.

Although there is a substantial amount of literature researching costs of illnesses, transferring health care costs from internationally published studies is only considered acceptable in rare cases. The most important reasons for this are differences in national health care funding systems and clinical practices and varying study designs (Molinier et al., 2006). Therefore, the results from studies performed in other countries can merely serve as a reference point in most cases. In addition, the continuous improvement in medication and treatments also shifts the cost structures of diseases, making older cost studies obsolete (Cipriano et al., 2011). These health care improvements will frequently result in higher medical costs for treating illnesses. When cancer treatments are concerned, increased costs

[21] Parts of this chapter have been presented at the 8th Conference on Sustainable Development of Energy, Water and Environment Systems (SDEWES): Schreurs, E., Cleemput, I., Nawrot, T. and Thewys, T. (2013). Economic impact assessment of environmental contamination: A damage function approach [Poster presentation]. SDEWES, Dubrovnik (Croatia), 22-27 September 2013.

for cancer therapies such as radiotherapy and particularly chemotherapy account for the largest part of rising costs over time (Warren et al., 2008). According to Yabroff et al. (2011), the economic burden of treating cancer is unlikely to decrease in the future. In general, it can be concluded that data used in health economic research should comply with two conditions: (1) emerge from the country it is applied to and (2) as recent as possible to obtain up to date cost assessments.

With respect to the considered illnesses, hip fractures have been extensively investigated by the research group of Jean-Yves Reginster (Université de Liège) in recent decades. Besides reporting the incidence of hip fractures in Belgium (Gillet and Reginster, 1999; Hiligsmann et al., 2010; Hiligsmann et al., 2008a) and providing clinical evidence for treating osteoporosis (Reginster and Brandi, 2010; Reginster et al., 2000), the research group also focused on the economic evaluation of osteoporosis treatments and the direct medical costs of osteoporosic fractures in Belgium (Hiligsmann et al., 2009; Reginster et al., 1999). Recently, they presented mean hospitalization costs for hip fracture patients using data from a Belgian sickness funds and originating from 2007 (Hiligsmann et al., 2011b). In this perspective, the two most important conditions described above for transferring cost estimates in a health setting are fulfilled, so there is no need to redo the cost assessment for hip fractures. The cost figures from Hiligsmann et al. (2011b) will be transferred in order to estimate health care costs from environmental pollution related to osteoporosis.

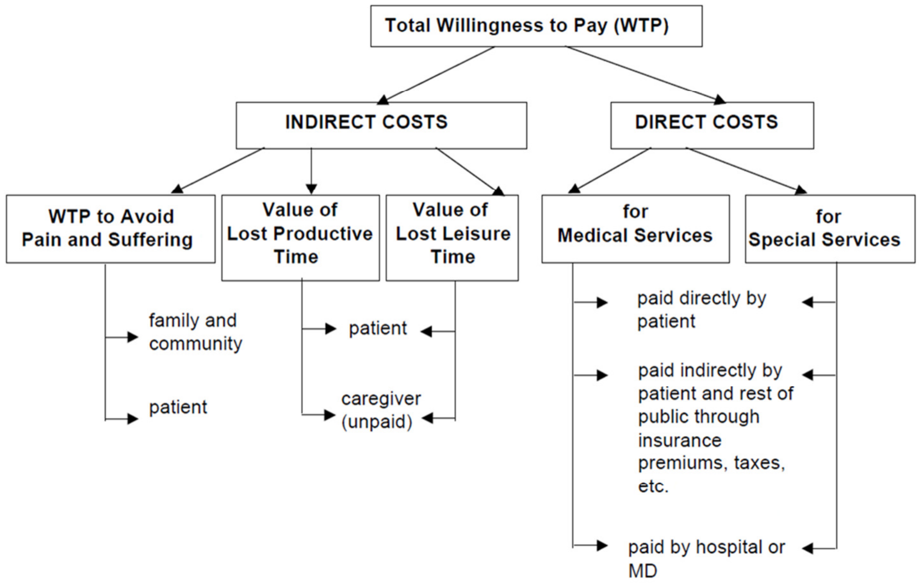
For lung cancer, however, no Belgian cost assessments have been performed and published in the international scientific literature in recent years. Therefore, the main objective of this chapter is to analyze the health care costs associated with lung cancer in Belgium in order to estimate the pollution-attributable health care costs in the Campine region. This chapter continues as follows. In section 4.2, the methodology on how these cost factors are estimated will be put forward. Section 4.3 will elaborate on the data that were used and the health care system in Belgium, while the results of the analysis will be provided in section 4.4. In section 4.5 we will discuss a number of studies on the willingness to pay for avoiding mortality and morbidity and some cost studies on lung cancer and hip fractures. In the last section, some concluding remarks will be given.

4.2 Methodology

As discussed in section 3.2, the third stage of the damage function approach considers the economic evaluation of the damages caused by the pollution. If the damage cost methodology is applied in a health context – such as here – it is also referred to as the Cost Of Illness (COI) approach (Dickie, 2003). COI appraises the value of the resources that are expended or foregone as a result of a health problem. Therefore, the direct and indirect resource costs that are effectively brought about by the environmental effect are comprised in the COI estimate. Figure 4.1 provides an overview of all societal cost elements of illnesses. Since damage costs are focused on resource costs, the method will exclude changes in the individual's utility value, which is displayed in Figure 4.1 by the WTP to avoid pain and suffering.

The direct costs contain all costs made by the patient, social security and insurance companies which can be directly related to the illness and its treatment. Besides costs for medical services such as hospital admissions, doctor visits and medication, this also includes non-medical costs for tasks that were usually performed by the patient, but which he/she is unable to do because of the illness (e.g. housekeeping). With respect to the indirect costs, there are opportunity costs associated with the loss of productive and non-productive time due to the illness. The opportunity costs of family members and other informal caregivers to provide assistance in helping the patient are also included in the indirect costs. The final cost component consists of the willingness to pay for avoiding the disutility associated with the illness and is thus excluded in COI estimates. This component is discussed more extensively in section 4.5.1.

Figure 4.1: Elements of societal cost of illnesses (EPA, 2006)



In this study, the scope of COI is further reduced to merely include the direct medical costs associated with lung cancer. Consequently, the cost estimate that is provided will be rather conservative for the actual damage costs caused by the environmental pollution. This restricted design is primarily due to the objective of our research to provide a rough estimate of the economic damage produced by the presence of environmental pollution in the Campine region. However, we do not intend to perform an extensive health economic study on lung cancer. From a research design perspective, data collection with regard to the direct medical costs is relatively straightforward in Belgium (see section 4.3). For other direct costs and indirect costs data are needed on non-medical expenditures made by lung cancer patients, time allocation of patients and caregivers, productivity losses and so on. Therefore, a much more elaborate research design would have been necessary. However, this falls outside the scope of our analysis[22].

[22] In the discussion section of this chapter (section 4.5.1), we will elaborate on the methodologies and results from WTP studies for avoiding mortality and morbidity. This

A case-control design was used to compare the direct medical costs of a sample of lung cancer patients in Belgium to the direct medical costs that have been incurred by a control sample of Belgian citizens with similar socio-demographic and socio-economic characteristics as the patient sample but without lung cancer. Patients were selected within the time frame from 2007 until 2010, given that 2010 was the last completely available year in the database used for data extraction (see section 4.3). This time frame allows us to define a patient group which is large enough to be statistically significant, while at the same time working with relatively recent data. Selected patients are followed up longitudinally. Health care costs are administered from 3 months prior to diagnosis until 18 months after diagnosis or until death. The period after diagnosis is broken down into an acute phase – from diagnosis until 3 months later – and a chronic phase – from 3 months until 18 months after diagnosis. Self-employed individuals are excluded from selection because parts of their medical costs were not reimbursed by the national health insurance scheme up until 2007. Therefore, their inclusion would lead to inconsistent results.

In order to select a control sample, a matching procedure was established that was based on the criteria age, sex, employment situation and preference classification (people receiving increased reimbursements) in the patient sample. First, a sample was created that included all individuals that complied to the conditions established by the matching procedure. Consequently, a number of individuals were randomly selected from this sample to create a control sample that was comparable in size to the patient sample. People in the control sample are followed up longitudinally as well. This allows for a calculation of the incremental medical costs which are attributable to lung cancer by taking the difference between the costs made by the two samples.

However, since this analysis aims to estimate the damage costs attributable to environmental pollution in the Campine region, some specific

aims to serve as a reference point for putting the resource costs found in our analysis into perspective.

characteristics of lung cancer patients have to be taken into account. It has been extensively demonstrated that the principal risk factor for lung cancer is cigarette smoke (Khuder, 2001). Ezzati and Lopez (2003) found that 71% of lung cancer mortality was attributable to smoking, while Souhami and Tobias (2003) claim that tobacco is responsible for as much as 90% of lung cancer incidence. Therefore, in sampling a group of lung cancer patients, there will be a considerable overrepresentation of smokers in comparison with a random population sample. Given that tobacco abuse is also an important risk factor for morbidity and mortality due to – among others – cardiovascular (Aboyans et al., 2010) and pulmonary (Taylor, 2010) diseases, the additional medical costs caused by these illnesses will also be included in the costs of the patient sample. This might distort the comparison between costs for the patient sample and a random control sample. If the effect of smoking comorbidities can be taken into account in the calculations (by controlling for comorbidities), the incremental costs which are effectively a result of the lung cancer diagnosis – thus excluding the costs due to smoking – will be discovered. Consequently, a second control sample was created which adjusted for the principal comorbidities of smoking.

4.3 Data

In Belgium the social security system requires every citizen to have a health care insurance, covering the costs for all basic health care interventions. The mandatory federal health care insurance is free of charge and is managed by seven sickness funds that are active in Belgium. These sickness funds are under the supervision of the National Institute for Health and Disability Insurance (RIZIV-INAMI), which is – among other things – responsible for the organization of reimbursements of medical costs and the acknowledgment of all reimbursable medical interventions in Belgium. Each of these interventions receives a unique 6-digit nomenclature number, so that every time a caregiver performs a medical intervention, the nomenclature number, the place and date of the intervention are added to the patient's record in his/her sickness fund database. Consequently, all sickness funds gather an abundance of data concerning the socio-economic status and the medical costs made by their members.

The Intermutualistic Agency (IMA) is an organization that has been founded to analyze the data gathered by sickness funds with the aim of supporting research for improving the economic efficiency of the Belgian health care system. All organizations in the health care system, but also other government institutions and universities, can file a request for an analysis of IMA's Permanent Sample (EPS), which randomly selects 1 out of 40 persons entitled and 1 out of 20 persons aging older than 65 from the complete sickness fund database. If the request is approved by the Technical Commission of the Permanent Sample (TCPS), holding all stakeholders involved in IMA, data extractions from EPS can be started.

It is in this setting that the cost analysis of lung cancer is performed. Although the IMA database delivers high quality and very detailed data, there are a number of constraints that have to be considered when working with administrative-financial databases before the results can be analyzed. First of all, diagnosis codes and other medical information are *not* registered in the sickness fund databases, only the place, date and health care costs associated with each nomenclature number are documented. Therefore, the definition of a patient group is based on a number of proxies, being interventions which patients with the illness of interest are very likely submitted to. It is important to define the patient group in a balanced way, restricting the patient group to those patients of which it can be expected that the probability that they have been diagnosed with the required illness is as high as possible, while at the same time avoiding to define the illness in such a strict way that a selection bias is created and the sample is not representative for the patient population of interest for the study. Moreover, the conditions have to be set up in such a way that a sufficiently large sample can be extracted from IMA's Permanent Sample.

Another constraint concerns the completeness of the cost factors that can be extracted from the database. The IMA database only registers the medical costs which are acknowledged by RIZIV-INAMI and hence reimbursed by the sickness funds. This implies that all medical costs that are not reimbursed and all non-medical costs are not included in the databases. Non-reimbursed medical costs include all supplementary medical costs that are not covered by the mandatory health care insurance, such as experimental cancer treatments, non-reimbursed medication and supplements for private hospital bedrooms. Non-medical costs involve costs of non-medical home care – which is the responsibility of regional governments in Belgium – and costs made by informal caregivers.

Furthermore, hospitalization costs are only partially included in the IMA database. Hospitals in Belgium are financed by a complicated and rather opaque system which is largely a joint effort of RIZIV-INAMI and the Federal Public Service (FPS) of Public Health (FOD Volksgezondheid). The FPS Public Health provides an annual budget for hospital activities – also referred to as the 'budget of financial means' (BFM) – which is composed out of capital costs (section A), operational costs (section B) and corrective measures (section C) (Table 4.1). The operational costs make up more than 90% of the entire budget and can be, in its turn, divided into multiple subclasses. The common operational costs (category B1) and clinical costs (category B2) are the most important cost factors, covering the majority (80-90%) of all operational costs. Approximately 20% of B1 and B2 is financed via an invoice that hospitals can submit to the sickness funds of the admitted patients, which then allocates funds according to the number of admissions (50%) and the number of nursing days (50%)[23]. The rest of this budget is fixed and paid to hospitals on a monthly basis (Cleemput et al., 2012).

[23] This is the situation for general acute hospitals. In case it concerns another kind of hospital, the variable part depends completely on the number of nursing days.

Table 4.1: Components of the hospital budget of financial means (Cleemput et al., 2012)

Section	Cost factor	Category	Description
A	Capital costs	A1	Investment charges
		A2	Short-term credit burdens
		A3	Investment charges for some medico-technical services which are exclusively financed via the hospital budget
B	Operational costs	B1	Administration, maintenance, laundry
		B2	Personnel and medical equipment
		B3	Medico-technical departments
		B4	Specific lump sum costs (as a result of legal obligations)
		B5	Pharmacy costs
		B6	Social agreements for personnel (not included in B2)
		B7	Teaching hospital or university function
		B8	Social function of the hospital
		B9	Extra-legal financial benefits
C	Corrective measures		

The other major part of hospital financing in Belgium comes from the contribution of honorarium fees to the hospital's operational costs. The rationale for this arrangement is that hospitals provide doctors with the necessary equipment in order to ensure an appropriate amount of care for patients. That is why doctors working in hospitals are required to transfer a certain percentage – different between hospitals and between specialisms – of their honorarium fees to the hospital. This part of hospital financing is referred to as 'fee for service' (FFS). Since the RIZIV-INAMI is responsible for hospital financing based on FFS and the variable part of BFM requires nomenclature numbers to be submitted to sickness funds, these hospitalization costs are included in the IMA database. However, this entails only a part of the complete hospitalization costs made by patients.

The partial inclusion of hospitalization costs in IMA estimates can be resolved by indirectly estimating the total hospitalization cost using the duration of inpatient stay. If the number of inpatient days are multiplied with an average per diem hospitalization price for Belgium, the complete hospitalization cost can be estimated. Since the hospital per diem price varies between hospitals, the Belgian guidelines for economic evaluations of health care interventions (Cleemput et al., 2012) suggest using a weighted average per diem price per type of hospital stay, in which the weights are determined by the size of the hospitals. The resulting per diem hospital prices can be found in Table 4.2. Depending on the time frame in which the hospital stays have been made, a cost factor can be attributed to the inpatient days to obtain the total cost of hospitalization.

Table 4.2: Weighted average per diem hospitalization prices per hospital type (in €) (Cleemput et al., 2012)

Hospital type	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Acute	289	288	308	322	352	373	388
Burns	1061	1075	1135	1155	1208	1253	1294
Geriatrics	174	179	179	196	212	216	215
Palliative	402	402	418	426	450	463	472
Psychiatric	178	177	186	196	210	215	222
Rehabilitation	194	194	208	220	229	235	246

Furthermore, as a result of privacy considerations access to the IMA database is restricted to IMA data analysts only. Data are delivered 'as is' in aggregated form, which causes some other constraints. Since this precludes working with individual level data, we are unable to perform extra analyses on the data. For example, it prevented us from checking whether there are significant within-sample cost differences and if so, to what they are attributable. Moreover, the aggregation of the results also implies that it is unknown in which year the patient's medical costs have been made in study designs that span multiple years. Consequently, it is not possible to discount the economic data to a common time frame.

4.4 Results

4.4.1 Sample selection

4.4.1.1 Patient sample

Selecting lung cancer patients from the IMA database requires thorough knowledge of the procedures that are in place when dealing with a case of lung cancer in Belgium. Therefore, three pneumonologists – of which two were specialized in respiratory oncology – were contacted for guidance in choosing the conditions that have a high probability of leading to the selection of lung cancer patients. They all agreed that a Multidisciplinary Oncologic Consult (MOC), a meeting between doctors with different specialisms such as an oncologist, radiologist, surgeon,… which is standard procedure for diagnosis of a cancer case in Belgium, should definitely be one of the conditions for selecting any cancer patient. The nomenclature number for a MOC charged by the doctor-coordinator is 350372-350383[24]. If this intervention is organized and charged by a pneumonologist, chances are very likely it concerns a case of lung cancer.

Furthermore, a number of other conditions were suggested by the pneumonologists. For example, a consultation of a pneumonologist in the 3 months prior to diagnosis is likely to have occurred in lung cancer patients. Also, the diagnosis of lung cancer is in most cases confirmed by a bronchoscopy or by anatomopathologic research of lung tissue. However, the introduction of these supplementary conditions might bias the patient sample that is selected from the database, since not *all* lung cancer patients are submitted to these interventions. Therefore, the two former conditions, a MOC that has been charged by a pneumonologist, are retained for selecting lung cancer patients. The socio-economic characteristics of the patient sample (n = 359) can be found in Table 4.3.

[24] Since 1st November 2010 two supplementary nomenclature numbers for MOCs have been introduced: 350291-350302 for an additional MOC performed in a different hospital than the first MOC and 350276-350280 for a follow up MOC. However, seeing that these nomenclature numbers have come into effect at the very end of our selection time frame 2007-2010, they are not taken into account.

Socio-economic characteristics		Frequency	Percentage	Cumulative percentage
Sex	Male	260	72.42%	72.42%
	Female	99	27.58%	100%
Age	<46	5	1.39%	1.39%
	46-50	5	1.39%	2.79%
	51-55	15	4.18%	6.96%
	56-60	23	6.41%	13.37%
	61-65	24	6.69%	20.06%
	66-70	54	15.04%	35.10%
	71-75	79	22.01%	57.10%
	76-80	85	23.68%	80.78%
	81-85	49	13.65%	94.43%
	>85	20	5.57%	100%
Employment situation	Resident	6	1.67%	1.67%
	Active	42	11.70%	13.37%
	Disabled	20	5.57%	18.94%
	Retired	268	74.65%	93.59%
	Deceased partner	21	5.85%	99.44%
	Unknown	2	0.56%	100%
Preference classification	Yes	100	27.86%	27.86%
	No	259	72.14%	100%

Table 4.3: Socio-economic characteristics of the patient sample

The patient sample predominantly consists of males and almost 80% of the patients are more than 65 years old. Consequently, the majority (74,65%) of the patient sample is retired. When comparing the patient sample with the lung cancer population in Belgium in the time frame 2007-2010 (Table 4.4a), there are slightly more female patients in the sample. However, with regard to age there is a clear overrepresentation of people aged between 71 and 80 as well as an underrepresentation of 51-65 year old persons in the lung cancer sample. For a large part this can be explained by the structure of IMA's EPS, which includes 1 out of 40 persons entitled and 1 out of 20 persons aging older than 65 (as mentioned in section 4.3). If the patient group is (hypothetically) adjusted so that each person entitled has a 1 out of 40 chance of being included – the probability of persons older than 65 being admitted is halved – there is a rather close match between the (age-adjusted) patient sample and the lung cancer population with respect to age (Table 4.4).

Table 4.4: Sex and age distribution of (a) population of lung cancer patients in Belgium from 2007 until 2010 and (b) age-adjusted patient sample

		a. Lung cancer population 2007-2010			**b. Age-adjusted patient sample**	
		Frequency	**%**	**Cumulative %**	**%**	**Cumulative %**
Sex	Male	22,010	73.47	73.47	72.42	72.42
	Female	7946	26.53	100	27.58	100
Age	<46	571	1.91	1.91	2.32	2.32
	46-50	1036	3.46	5.36	2.32	4.64
	51-55	2223	7.42	12.79	6.96	11.60
	56-60	3373	11.26	24.05	10.67	22.27
	61-65	4344	14.50	38.55	11.14	33.41
	66-70	4393	14.66	53.21	12.53	45.94
	71-75	5017	16.75	69.96	18.33	64.27
	76-80	4781	15.96	85.92	19.72	83.99
	81-85	2951	9.85	95.77	11.37	95.36
	>85	1267	4.23	100	4.64	100

There are still slightly too little 61-70 year old persons and too many 71-80 year old persons in the age-adjusted patient sample in comparison with the lung cancer population, but the differences are quite small. Given the relatively good representation of lung cancer patients in the age-adjusted patient sample, this might be an indication that the selection procedure for lung cancer patients has worked in a correct way. Nevertheless, the patient sample is biased towards people older than 65, which generally have higher health care costs than younger people. From this point of view, the results will overestimate the true direct medical costs of lung cancer.

4.4.1.2 Control sample

A control sample (n = 357) was selected from IMA's EPS on the basis of reference points to the patient sample with regard to sex, age, employment situation and preference classification (Table 4.5). It can be seen that all these characteristics are very similar in the control sample in comparison with the patient sample.

Table 4.5: Socio-economic characteristics of the control sample

Socio-economic characteristics		**Frequency**	**Percentage**	**Cumulative percentage**
Sex	Male	259	72.55%	72.55%
	Female	98	27.45%	100%
Age	<46	5	1.40%	1.40%
	46-50	5	1.40%	2.80%
	51-55	15	4.20%	7.00%
	56-60	23	6.44%	13.45%
	61-65	23	6.44%	19.89%
	66-70	54	15.13%	35.01%
	71-75	79	22.13%	57.14%
	76-80	84	23.53%	80.67%
	81-85	49	13.73%	94.40%
	>85	20	5.60%	100%
Employment situation	Resident	6	1.68%	1.68%
	Active	42	11.76%	13.45%
	Disabled	20	5.60%	19.05%
	Retired	268	75.07%	94.12%
	Deceased partner	21	5.88%	100%
	Unknown	0	0%	100%
Preference classification	Yes	100	28.01%	28.01%
	No	257	72.55%	100%

The objective of the patient-control comparison is to generate incremental direct medical costs for an average lung cancer patient in Belgium. However, since it is hypothesized that the majority of lung cancer cases in the patient sample is caused by tobacco smoke while the focus of this research is on lung cancer attributable to environmental pollution, the results should be adjusted to account for comorbidities related to smoking (see section 4.2). In order to reveal the sole health care costs of lung cancer without the medical costs related to smoking, a second control sample (n = 356) was created with similar socio-economic characteristics to the patient sample, but consisting of people which are expected to be smokers and do not have lung cancer.

Analogous to the patient sample selection procedure tobacco usage had to be derived indirectly, as smokers cannot be defined in a direct way from the EPS database. Therefore, a number of additional conditions had to be established that indicated which patients were very likely to be smokers or had a history of smoking, besides the socio-economic characteristics that were used to determine the first control sample. The pneumonologists were contacted again to suggest in which way a sample of smokers could be selected from the database. It was suggested that the most probable consequences of smoking were an increased risk of chronic obstructive pulmonary disease (COPD) and high levels of cholesterol. These assumptions regarding comorbidities in lung cancer patients due to smoking are supported by multiple studies (Lopez-Encuentra and Bronchogenic Carcinoma, 2002; Perin et al., 2012; Young et al., 2009). However, since these illnesses can lead to a broad spectrum of medical interventions, the specialists recommended to focus on the medication that is used to treat tobacco dependence and these specified illnesses.

In order to reduce tobacco dependence in patients, Champix (ATC N07BA03) and Zyban (ATC N06AX12) are the only medicines that are reimbursed by RIZIV-INAMI. Although some sickness funds choose to reimburse nicotine replacement therapy for smokers as well as part of their complementary insurance package, it is not reimbursed by RIZIV-INAMI. For handling high levels of cholesterol the most frequently prescribed medicines in Belgium are simvastatine (ATC C10AA01), pravastatine (ATC C10AA03) en atorvastatine (ATC C10AA05). The most important medication for treating COPD patients are ATC codes R03BB04, R03AK06, R03AK07 and R03AK07. However, these medicines can also be used for treating asthma, which is not a comorbidity caused by tobacco smoke. Therefore, an additional condition has to be imposed for distinguishing the COPD patients and the asthma patients. Seeing that asthma patients are frequently prescribed medication from ATC codes R03DC03 or R03DX05 as well, this was an *excluding* condition for selecting COPD patients.

If considered separately, all medication mentioned in the previous paragraph can be prescribed to patients as a result of numerous causes outside of tobacco abuse. Particularly medication for treating high levels of cholesterol is prescribed very frequently, for example in overweight or older patients. Therefore, focusing on particularly selecting the patients who are on medication because of their smoking habits, the rule was established that a patient in the second control sample should take medication from *at least two categories* treating different symptoms. The combination of two potentially tobacco-related medicines increases the probability that the selected patient effectively is a smoker (Table 4.6).

Table 4.6: Socio-economic characteristics of the control sample which is adjusted for comorbidities related to smoking

Socio-economic characteristics		Frequency	Percentage	Cumulative percentage
Sex	Male	258	72.47%	72.47%
	Female	98	27.53%	100%
Age	<46	5	1.40%	1.40%
	46-50	5	1.40%	2.81%
	51-55	15	4.21%	7.02%
	56-60	23	6.46%	13.48%
	61-65	22	6.18%	19.66%
	66-70	54	15.17%	34.83%
	71-75	79	22.19%	57.02%
	76-80	84	23.60%	80.61%
	81-85	49	13.76%	94.38%
	>85	20	5.62%	100%
Employment situation	Resident	6	1.69%	1.69%
	Active	42	11.80%	13.48%
	Disabled	20	5.62%	19.10%
	Retired	268	75.28%	94.38%
	Deceased partner	20	5.62%	100%
	Unknown	0	0%	100%
Preference classification	Yes	99	27.81%	27.81%
	No	257	72.19%	100%

Obviously, the selection procedure for smokers is not without flaws. Although we aimed to minimize the probability of including non-smokers, there is no way of telling with 100% certainty that they have not been selected. Moreover, using this selection procedure forces us to assume that all smokers take some kind of tobacco related medication, while this is of course not the case. Consequently, the smoking-adjusted control sample is more likely to include heavy smokers and overestimate the health economic burden for smokers. That is why comparing the medical costs from the patient sample and the smoking-adjusted control sample has to be done with care so that the results are not misinterpreted. Nevertheless, considering the limits of working with a financial-administrative database, this procedure seemed to be the best option for selecting a smoker control sample.

4.4.2 Lung cancer costs

4.4.2.1 IMA costs

In order to obtain a detailed insight into the cost data provided by IMA, first an overview of the original IMA data extraction results are provided for all samples. The results in Table 4.7 have not been manipulated and include all medical costs acknowledged by RIZIV-INAMI and the variable part of hospitalization costs. As mentioned before, IMA only supplied aggregated cost figures on a patient level such as mean costs, median costs and standard deviations. Distinction is made between specified phases and to whom the costs are accounted.

Table 4.7: IMA data extraction results

		Prediagnostic phase		**Acute phase**		**Chronic phase**	
		RIZIV-INAMI	Patient costs	RIZIV-INAMI	Patient costs	RIZIV-INAMI	Patient costs
Patient sample	Mean	4056.02	640.10	5730.19	623.66	11,155.88	951.93
	Median	3088.91	330.79	4853.84	370.44	7122.18	553.64
	Standard deviation	4348.17	924.62	4890.74	881.96	13,539.52	1644.76
Control sample	Mean	694.23	85.87	799.54	104.94	4627.36	599.26
	Median	165.85	23.53	175.84	26.29	1262.01	184.69
	Standard deviation	1598.87	181.84	2011.08	287.51	9524.25	1828.91
Smoking-adjusted control sample	Mean	1099.15	135.32	1071.56	161.28	5427.76	596.88
	Median	211.91	28.93	237.23	27.27	2252.25	240.95
	Standard deviation	2379.13	426.79	2112.26	732.55	8517.68	974.63

Wald tests point out that mean costs for the patient sample are significantly higher in comparison with both control samples for all disease stages. The smoking-adjusted control sample deals with significantly higher mean costs than the control sample in the prediagnostic phase (at the 1%-level for reimbursed costs and at the 5% level for patient costs). Both mean and median costs are higher in the other disease stages as well, but not significantly. These results illustrate that the sample of smokers is dealing with higher medical costs in comparison with a standard control sample, which rationalizes the creation of a smoking-adjusted control sample in this setting. With respect to the distribution of the cost data, it can be observed that mean costs are overall higher than median costs, particularly in the control samples. This indicates that the distribution is skewed to the right and the samples include a minority of individuals incurring high(er) medical costs. Coefficients of variation confirm that variation is generally more eminent in the control samples (Table 4.8), although this can be related to the lower absolute values in these samples as well.

Table 4.8: Coefficients of variation in IMA costs

	Prediagnostic phase		**Acute phase**		**Chronic phase**	
	RIZIV-INAMI	Patient costs	RIZIV-INAMI	Patient costs	RIZIV-INAMI	Patient costs
Patient sample	1.07	1.44	0.85	1.41	1.21	1.73
Control sample	2.30	2.12	2.52	2.74	2.06	3.05
Smoking-adjusted controle sample	2.16	3.15	1.97	4.54	1.57	1.63

4.4.2.2 Hospitalization costs

Since the IMA assessments only include the variable part of hospitalization costs, complete hospitalization costs are calculated separately in order to add them to the IMA costs in a later stage. These costs are estimated by multiplying the number of inpatient days with an average per diem hospitalization price for Belgium, as recommended by Cleemput et al.

(2012). In this setting, the amount of inpatient days is established per phase and per sample in Table 4.9.

Table 4.9: Number of inpatient days per phase and per sample

Sample	Prediagnostic phase	Acute phase	Chronic phase
Patient sample	3413	3246	3995
Control sample	264	250	1108
Smoking-adjusted control sample	408	330	1567

In assigning a cost factor for an average per diem hospitalization price, it appears to be rather simplistic to allocate one single per diem price to the entire sample, considering that the study's time frame spans four years (2007-2010) and average per diem prices have increased steadily in recent years (Table 4.2). Moreover, since costs in the chronic phase are incurred up to 21 months after prediagnostic costs have been made, it is plausible to assume that the former phase accounts for higher hospitalization costs. Therefore, it seems to be reasonable to create differentiated per diem prices per disease stage, so that assigned prices will correspond more closely to the effective prices. On the one hand, this differentiation will be based on the weighted average hospitalization costs for acute hospitals (Table 4.2). Lung cancer cases in Belgium are most likely to be treated in this type of hospital. On the other hand, the probability that a per diem price accounts for the hospitalization cost in a certain disease stage will be determined in order to calculate an expected per diem price for each phase. In determining these probabilities it is assumed that diagnosis dates are linearly distributed across the year of diagnosis.

Although the same logic applies to all phases, the probability distribution and the expected per diem hospitalization prices will be determined for the prediagnostic phase as an example. For patients being diagnosed with lung cancer in 2007, the prediagnostic phase will fall completely within 2007 since this phase covers the three months before diagnosis. Per diem hospitalization costs for the prediagnostic phase of these patients will equal the weighted average per diem price for 2007, more specifically €322. However, for patients being diagnosed in 2008 there is a one in four chance that parts of the prediagnostic phase will be incurred in 2007. Since it is assumed that the diagnosis dates are linearly distributed over the year of diagnosis, there is a 1 to 8 probability that an inpatient day has occurred in 2007. Therefore, the expected per diem hospitalization price in the prediagnostic phase for incidences in year t will amount to:

$$E(P_{t,\ prediagnostic}) = 1/8\ P_{t-1} + 7/8\ P_t \qquad [5]$$

If the probability distributions are combined with the weighted average per diem prices given in Table 4.2, this results in the expected per diem prices per phase and per year of diagnosis (Table 4.10). Since only cases are included that could be followed up during the entire time frame (21 months) or until mortality, the number of lung cancer cases differs per year of diagnosis. The probability of a 2008 lung cancer case being admitted are greatest, seeing that all cases from 2008 can be admitted while all other years deal with some restrictions. For example, patients diagnosed in the first three months of 2007 could not be included since we were unable to calculate complete costs for their prediagnostic phase. The lowest incidence is found in 2010 due to the inability to cover the costs for their chronic phase. Consequently, all 2010 lung cancer cases must have suffered from early mortality.

Table 4.10: Expected per diem hospitalization price per phase (in €) and lung cancer cases per year of diagnosis

Year of diagnosis	**Prediagnostic phase**	**Acute phase**	**Chronic phase**	**Lung cancer (n)**
2007	322.00	325.75	347.35	97
2008	348.25	354.63	369.78	143
2009	370.38	374.88	384.25	93
2010	386.13	388.00	388.00	26
Weighted average per diem price	349.63	354.49	368.79	/

Given the difference in inclusion probability, the number of lung cancer cases per year of diagnosis has served as a weight factor for calculating the weighted average expected per diem hospitalization price per disease stage (depicted in the last line of Table 4.10). In combination with the duration of inpatient stays (Table 4.9), these prices are used to assess the total hospitalization costs.

4.4.2.3 Total costs

After the hospitalization costs are added to and the variable part of hospitalization costs is removed from the IMA results, the total costs for lung cancer cases can be analyzed. The entire patient sample (n = 359) consumed health care costs amounting to nearly €11 million over the 21 months they were followed up (Table 4.11). A great majority of these costs (93,4%) were covered by RIZIV-INAMI, which was slightly higher than the control sample and the smoking-adjusted control sample (90,7% and 91,6%, respectively). Table 4.11 also includes the sample sizes for all groups. This illustrates that mortality had a great impact on the size of the patient sample. Thirty percent of the sample (n = 105) had passed away 3 months after they had been diagnosed with lung cancer. In the control groups the sample sizes represent the amount of people that have incurred health care costs in the period they were followed up. For example, in the control sample only 311 individuals have incurred health care costs, which implies that 46 (= 357 – 311) individuals have not made any costs in these three months. For both control samples it is assumed that no individual has passed away up to 6 months after they have been included in the sample – i.e. up to the chronic phase – because there was no information with respect to time of death for these samples.

Average medical costs per lung cancer case run up to more than €36,000 over the complete study time frame (Table 4.12). Patient’s out-of-pocket costs over the 21 month time frame constitute €2314.64. When the costs made by non-lung cancer patients with similar socio-economic characteristics – i.e. the control sample – are taken into account, the incremental costs of lung cancer are obtained. It is observed that lung cancer patients experience a substantial amount of extra costs in comparison with the average Belgian population. Almost €28,000 in medical costs, consisting of 94.5% reimbursed costs and 5.5% out-of-pocket costs, are attributable to lung cancer. The largest part of these costs (55%) are made in the first two phases. When the comorbidities of smoking are left out of the equation, it can be noticed that the incremental costs of lung cancer are somewhat lower in all phases. However, these costs still amount to €25,385.60.

In Table 4.13 the monthly incremental costs per lung cancer case are calculated. These calculations enable us to compare the costs between different stages in lung cancer progression on the same basis. The highest medical costs are incurred in the first 3 months after diagnosis, €2776.44 and €2652.66 incremental costs per month in comparison to the control sample and the smoking-adjusted control sample, respectively. In the prediagnostic phase monthly incremental costs are also rather high and the patient costs are even higher than in the acute phase. Monthly costs are substantially lower in the chronic phase of the illness. The average monthly incremental cost over the entire 21 month time frame amount to €1320.06 compared to the control group and €1208.84 when smoking comorbidities are taken into account. Approximately 6% of these costs are not covered by RIZIV-INAMI.

Table 4.11: Total costs per sample and per phase (in 1000€)

a)	**Prediagnostic phase**			**Acute phase**		
	RIZIV-INAMI	Patient costs	Sample size (n)	RIZIV-INAMI	Patient costs	Sample size (n)
Patient sample	2578.43	229.80	359	3138.62	223.89	356
Control sample	329.33	29.97	311	361.77	366.23	315
Smoking-adjusted control sample	525.29	48.18	315	472.05	57.42	312

b)	**Chronic phase**			**Aggregated over 21 months**		
	RIZIV-INAMI	Patient costs	Sample size (n)	RIZIV-INAMI	Patient costs	Total cost
Patient sample	4501.00	265.59	254	10,218.05	719.28	10,937.33
Control sample	2000.71	209.14	327	2691.81	275.73	2967.55
Smoking-adjusted control sample	2475.31	212.49	333	3472.65	318.08	3790.73

Table 4.12: Average and incremental cost per lung cancer case (in €)

a)	**Prediagnostic phase**			**Acute phase**		
	RIZIV-INAMI	Patient costs	Total cost	RIZIV-INAMI	Patient costs	Total cost
Average cost	7182.27	640.10	7822.37	8816.35	628.91	9445.26
Incremental cost	6259.76	556.16	6815.91	7802.99	526.33	8329.32
Smoking-adjusted incremental cost	5706.74	504.78	6211.52	7490.36	467.63	7957.99

b)	**Chronic phase**			**Aggregated over 21 months**		
	RIZIV-INAMI	Patient costs	Total cost	RIZIV-INAMI	Patient costs	Total cost
Average cost	17,720.46	1045.62	18,766.09	33,719.07	2314.64	36,033.71
Incremental cost	12,116.23	459.80	12,576.03	26,178.98	1542.28	27,721.26
Smoking-adjusted incremental cost	10,767.35	448.74	11,216.09	23,964.45	1421.15	25,385.60

Table 4.13: Monthly incremental cost per lung cancer case (in €)

a)	**Prediagnostic phase**			**Acute phase**		
	RIZIV-INAMI	Patient costs	Total cost	RIZIV-INAMI	Patient costs	Total cost
Incremental cost	2086.59	185.39	2271.97	2601.00	175.44	2776.44
Smoking-adjusted incremental cost	1902.25	168.26	2070.51	2496.79	155.88	2652.66

b)	**Chronic phase**			**Average monthly cost**		
	RIZIV-INAMI	Patient costs	Total cost	RIZIV-INAMI	Patient costs	Total cost
Incremental cost	807.75	30.65	838.40	1246.62	73.44	1320.06
Smoking-adjusted incremental cost	717.82	29.92	747.74	1141.16	67.67	1208.84

Besides the complete cost estimates, there were also some results obtained for individual nomenclature numbers of the most important interventions relating to lung cancer. The sum of these cost factors explained merely 57%, 51% and 51% of the reimbursed costs and 39%, 37% and 34% of the patient costs made by lung cancer patients in the prediagnostic, acute and chronic phase, respectively (Table 4.14). Because of the size of the unexplained cost section, i.e. the other costs, it is difficult to make any solid judgments on the cost factors that contributed to the cost assessments presented in Table 4.11 to 4.13. The other costs are attributable to interventions and medication which are possibly unrelated to lung cancer – but possibly more to comorbidities – or that were just left out of the list of nomenclature numbers that was appended to the data extraction proposal.

Table 4.14: Principal cost factors for the patient sample

Cost factor	**Prediagnostic phase**		**Acute phase**		**Chronic phase**	
	RIZIV-INAMI	Patient costs	RIZIV-INAMI	Patient costs	RIZIV-INAMI	Patient costs
Hospitalization	46.28%	25.96%	36.66%	24.22%	32.73%	22.65%
Chemotherapy	1.59%	0.00%	4.71%	0.00%	11.64%	0.03%
Radiotherapy	1.32%	0.80%	4.14%	1.30%	2.73%	0.00%
Surgery	0.69%	2.47%	1.03%	3.22%	0.07%	0.56%
Doctor’s fees	1.35%	4.65%	1.92%	5.05%	1.61%	7.90%
Diagnostic tests	3.59%	4.19%	1.37%	2.24%	0.88%	1.60%
Medical imaging	2.33%	1.33%	1.01%	0.87%	0.94%	0.80%
Other costs	42.87%	60.61%	49.15%	63.10%	49.39%	66.45%

However, it is possible to (carefully) derive some implications and trends from the detailed cost assessment. Hospitalization costs are the largest cost factor for the patient sample, ranging from 46% of reimbursed costs in the prediagnostic phase to 33% in the chronic phase. For the prediagnostic phase important cost factors are diagnostic tests and medical imaging, while radiotherapy costs are a substantial expense in the acute phase. Costs for the most important chemotherapy medication increase considerably from the prediagnostic phase (2% of reimbursed costs) to the acute (5%) and chronic phase (12%), indicating its importance in later stages of lung cancer. Patients do not have to contribute to chemotherapy costs because these medicines are completely reimbursed in Belgium (category A). Doctor's fees (including GP, lung specialist, oncologist and MOC) are relatively more burdening to patients.

4.4.3 Hip fracture costs

The cost assessments described in this section are based on the findings of Hiligsmann et al. (2011b). This research used data originating from the RIZIV-INAMI database in 2007 to determine the direct costs of hip fractures in Belgium. Since the assessment is based on the same database as the lung cancer study, the same cost factors as described in section 4.3 are included in and excluded from the cost estimates. The patient population was derived from the database by the ICD-9 code for hip fractures (820) and contained 13,950 persons. The majority of hip fracture patients was female (74.3%) and the mean age was 81.48 years for women and 77.59 years for men. All included patients were older than 50 years. Patients were mostly (93.9%) subjected to a surgical procedure.

Notwithstanding the similarities to the lung cancer study, there are also some important differences between the two study designs. First of all, the study did not distinguish between reimbursed and out-of-pocket costs, which prevents us from calculating the patient's share in hip fracture costs. Furthermore, direct medical costs were only recorded for the hospitalization period after the hip fracture had occurred. Although costs for inpatient care entail the bulk of direct costs due to hip fractures (Borgstrom et al., 2007; Borgstrom et al., 2006), the setup is different from the longitudinal follow-up established in the lung cancer study. Moreover, all costs have been incurred in 2007, while the lung cancer study gathered data from 2007 until 2010. Since the lung cancer costs cannot be discounted to one year, the results are in different economic units. Because of these differences in study design and economic units, it would be inappropriate to combine and add up the results from both studies to estimate the pollution-attributable health care costs. Rather, the cost estimates should be discussed separately.

The hospitalization costs for hip fractures in women and men are shown in Table 4.15. Costs for treating hip fractures are generally higher in older age groups and in men. Costs of hip fractures for men (age-weighted mean: €10,828) were significantly higher than for women (age-weighted mean: €10,389) at the 5%-level, even without adjusting for the lower age at hip fracture incidence for men. Principal cost factors in hip fractures are the costs for inpatient days and doctor’s fees. The small minority of patients that were subjected to a non-operative treatment dealt with substantial lower costs (mean: €6063) in comparison with a surgical procedure.

Table 4.15: Hospitalization costs of hip fractures per age category (in €) (Hiligsmann et al., 2011b)

Age	**Women**	**Men**
50-54	7571	7261
55-59	8230	8317
60-64	8734	9263
65-69	8811	10,038
70-74	9729	9919
75-79	10,485	11,445
80-84	11,521	11,381
85-89	10,552	11,406
90 +	9667	11,364

4.4.4 Pollution-attributable health care costs

In the previous chapter the number of pollution-attributable cases have been estimated (Table 3.6, 3.7 & 3.8). These estimates will be used in order to determine the damage costs attributable to the environmental pollution in the Campine region. However, the multitude of cost estimates forces us to make a (rather arbitrary) choice which cost factors ought to be used to determine the pollution-attributable health care costs. With respect to lung cancer, the smoking-adjusted incremental lung cancer cost over 21 months seems to match our objectives as closely as possible. It has been established before that in this case the costs for lung cancer should be adjusted for smokers since the focus is on the diseases attributable to the environmental pollution. For hip fractures, the average hospitalization costs for men and women are used to estimate damage costs, seeing that no specific age- or gender-related information on hip fracture incidence in the Campine region is available.

In both cases the actual resource costs will probably be (considerably) higher than established here. The higher medical costs for lung cancer patients in comparison with a random control group will probably continue after the 21 month time frame. Moreover, lung cancer patients, as well as hip fracture patients, are more likely to suffer from early mortality. Marks et al. (2003) show that hip fracture incidence leads to considerable functional disability and a decreased capacity for managing daily activities independently. Hence, these patient are likely to face increasing medical and non-medical costs after their hospital stay as well. From this perspective, the pollution-attributable medical costs provided in Table 4.16 will underestimate the true damage costs.

Table 4.16: Pollution-attributable health care costs

	Pollution-attributable cases per year*	**Health care cost per case (in €)**	**Pollution-attributable health care costs per year (in €)***
Lung cancer	22.20 (2.84-33.24)	25,385.60	563,560.31 (72,095.10-843,817.33)
Hip fracture men	8.50 (1.25-13.10)	10,828	92,038.00 (13535.00-141,846.80)
Hip fracture women	31.41 (15.44-42.23)	10,389	326,318.49 (160,406.16-438,727.47)

* The intervals give an indication of the variability which is associated with these estimates. They are an extension of the confidence intervals from the attributable fraction (Table 3.6, 3.7 & 3.8).

The largest part of pollution-attributable health care costs originate from lung cancer, with approximately €560,000 in costs on a yearly basis. However, the lung cancer cost deals with a significant amount of variability due to the uncertainty associated with the epidemiologic research. In case the lower range of the pollution-attributable cases is applied, the costs amount to merely €72,095 per year. With regard to hip fractures, the health care costs add up to €92,038/year for men and €326,318.5/year for women. The lower average health care cost for women in comparison with men is clearly compensated by the extra amount of cases that have occurred in the Campine region. As mentioned before, the results for both illnesses cannot be added since the cost estimates arise from different time frames and different research designs. Obtaining a complete (health-related) damage cost from the environmental pollution is thus economically infeasible.

4.5 Discussion

4.5.1 WTP to avoid mortality and morbidity

As explained in the methodology section of chapter 3 and 4, our economic analysis focused on resource cost flows as a result of the considered illnesses. However, it has been established before that the complete willingness to pay (WTP) to avoid an illness is likely to be at least as large as the direct costs that are lost due to the illness (Dickie, 2003) [25]. For instance, in 2007 the direct health expenditures due to cancer were approximately $89 billion in the United States (US), while the value of the productivity that was lost due to the illness and premature mortality – the largest part of indirect costs attributable to cancer – was estimated to be respectively $18.2 and $112 billion (American Cancer Society, 2008). These economic costs boil down to $750 per person per year. Milligan et al. (2010) followed up on this fact by estimating the median WTP for avoiding cancer and found that the WTP was $1933 per person per year (in 2007 prices). This example merely

[25] See Figure 4.1 for an overview of all cost components related to illnesses.

serves to show that people are eager to pay more than the pure economic costs for avoiding an illness. The disutility associated with disease burdens such as pain, tiredness, suffering, feeling ill, lost leisure time, … is also an important contributor to the societal cost of illnesses. In this discussion section, the objective is to provide more information on studies that have tried to quantify this aspect of illnesses in order to obtain a better understanding of the economic value estimates that can be found in the literature and to put the complete benefits of environmental regulations into perspective. In contrast to the damage cost method, these economic values need to be valued indirectly.

Most of the research on this topic has addressed the monetary value of early mortality (Hammitt and Haninger, 2011). An important concept that has been introduced in this setting is the Value of a Statistical Life (VSL)[26]. This concept is defined as the marginal willingness to pay for a small reduction in the risk of dying. When we put it more theoretically, the VSL is a ratio in which the numerator is the marginal utility of a small reduction in mortality risk, while the denominator is the marginal utility of a small change in income (Cameron, 2010). Since the VSL is designed to evaluate the economic welfare that is lost because people die earlier than could be expected in normal circumstances, VSL estimates provide governments with reference points for assessing the benefits of regulations aimed at improving public health (Viscusi and Aldy, 2003). There is a wide array of public health and safety projects for which the VSL concept is used: transportation safety, occupational health and safety interventions, environmental regulation and

[26] Authors such as Desaigues et al. (2011), Cameron (2010) and Ashenfelter (2006) (and many non-economists with them) find that the term ‘Value of Statistical Life’ does not accurately reflect the economic concept presented here. It seems to be putting a monetary value on the average person’s life, while it is actually an aggregation of all individual marginal willingness to pay estimates for a small mortality risk reduction in a certain population. As a result of this lexicographical misconception, Desaigues et al. (2011) chooses to refer to this concept as the ‘Value of Preventing a Fatality’ (VPF). Cameron (2010) proposes to use the ‘Willingness to Swap’ (WTS) alternative goods and services for a microrisk reduction in the chance of sudden death (or other types of risks to life and health). However, since the term VSL is widely accepted and used in the literature and we do not want to intervene in the expert’s discussion on what should or should not be the correct term for this concept, we will still use VSL throughout this discussion.

so on (Ashenfelter, 2006). In an environmental setting for example, it can concern the WTP for changing the risk of incurring a life-threatening illness that can be achieved by adjusting the population's exposure to substances that cause these illnesses.

With regard to the methodologies used for inferring WTP estimates, researchers have frequently used revealed preference methods, and hedonic wage studies more specifically, to estimate people's WTP for a mortality risk reduction. The hedonic wage method tries to extract the tradeoffs that people are making between wages and safety from the other factors that affect wages. The methodology thus uses a hedonic regression that controls for differences in employee productivity as well as factors that determine job quality in order to reveal the worker's WTP for a certain job fatality risk. See Aldy and Viscusi (2007) for an elaborate discussion on the methodology. One of the most widely cited papers on hedonic labor market studies is the meta-analysis provided by Viscusi and Aldy (2003). In this study, the VSL is estimated to be in the range of $4-9 million for the US labor market (not put to a single year dollar value). For non-US labor markets, somewhat lower VSL estimates are found. Since most other countries have a lower GDP per capita than the US, this corresponds to the income elasticity of 0.5-0.6 that is also found in the study. Mrozek and Taylor (2002) report similar values of income elasticity (0.46-0.49), but a lower VSL range of $1.5-2.5 million (in 1998$). Another meta-analysis of hedonic wage studies finds an average VSL of $6.3 million and a median VSL of $4.6 million (Bellavance et al., 2009). All considered meta-analyses report a substantial amount of variation in the VSL estimates throughout hedonic labor market studies.

However, VSL estimates from hedonic wage studies are considered to be less appropriate for 'benefit transfer' to environmental health policies (Scasny and Alberini, 2012). On a more general level, benefit transfers from 'out of context' values without adjustment are not recommended since it can have serious implications for policy evaluation (Dekker et al., 2011). Viscusi and Aldy (2003) also send out a warning signal themselves in the description of their results: "…transferring the estimates of a value of statistical life to non-labor market contexts, […], should recognize that different populations have different perspectives over risks and different values on life-saving." Study objects in hedonic wage studies are often considered to be largely different from the population that profits from improved environmental programs (Alberini et al., 2011). Hedonic wage studies only include people that are employed, which implies that its study objects are situated in the range of, putting it crudely, 20 to 65-year old individuals. According to Robinson (2007), the average participant in hedonic wage studies is 36-37 years old. The population in these studies is also likely to include more men than women in comparison with the general population. On the other hand, the people that benefit from environmental improvement programs particularly include the elderly and children, two demographic groups that are excluded a priori. This is probably one of the most important reasons why hedonic wage studies tend to report (much) higher VSL estimates than values found using other methods (Alberini and Scasny, 2011; Kluve and Schaffner, 2008; Kochi et al., 2006).

Hedonic wage methods suffer from other restrictions as well, that could be relevant in an environmental program setting. For example, the method is less suitable to deal with the temporal aspects of illnesses (Robinson, 2007). For instance, if the new regulation involves reducing exposure to substances that cause diseases with long periods of latency or pre-mortality morbidity (such as some forms of cancer), the results of hedonic wage studies, which tend to focus on actuarial risks of sudden death, will be not quite appropriate for assessing the health benefits as a results of the environmental regulation. Furthermore, in times of high unemployment the assumption of a fixed supply of jobs and a freely functioning job market is not likely to be fulfilled (Garrod and Willis, 1999). Other, more classic critiques on hedonic models involve problems such as omitted variable bias and information asymmetry. Since they are discussed more extensively in other parts of the dissertation, we will refrain from advancing on these topic here.

The assumptions and restrictions associated with revealed preference methods has increased the popularity of stated preference (SP) methods for estimating VSLs in an environmental setting. As mentioned before (see chapter 1), contingent valuation (CV) and choice modeling approaches such as discrete choice experiments (DCE) are the two most commonly applied stated preference methodologies. Analogous to the hedonic model issues in the previous paragraph, these methodologies are described more thoroughly in other parts of the dissertation. The DCE methodology is covered and applied in the sixth chapter, which also includes a discussion on willingness to pay estimates from SP analyses (see section 6.4.3). Hence, we refer to these chapters for more information on these methods.

With respect to the SP literature on VSL estimation, there have also been a number of meta-analyses on this topic. In one of the most recent analyses (Lindhjem et al., 2011) it has been found that the average VSL is $ 7.4 million and the median VSL of $ 2.4 million (in 2005$). The meta-analysis by Dekker et al. (2011) shows that the VSL estimates from road safety situations are considerably smaller and are dealing with (much) smaller standard deviations in comparison with VSL estimates from air pollution or other situations.

There are still a number of unresolved issues in the VSL literature. For example, are there differences in the WTP for reducing the risk of dying from fatal accidents or from chronic diseases such as cancer? It has been suggested that people are willing to pay more to avoid the dread and suffering from painful chronic illnesses (Sunstein, 1997). When cancer is the illness of interest, this premium is referred to as a 'cancer premium'. Hammitt and Liu (2004) indicate that the WTP for avoiding cancer is roughly one third higher than for a similar degenerative disease. Van Houtven et al. (2008) also finds strong evidence for the presence of a cancer premium with VSL estimates that are approximately 3 times higher in comparison with a fatal accident. The individual's subjective risk assessment of contracting the illness, as well as having a history of health problems have an impact on people’s WTP (Alberini et al., 2006). On the other hand, the duration of the latency period, which is typically around 15 to 20 years but varies with type of cancer among other things, is negatively correlated with the WTP (Tsuge et al., 2005). Although Alberini and Scasny (2011) do not find any discounting up to latency periods of 10 years, Van Houtven et al. (2008) shows that WTP reduces with 50% in case of a 25 year latency period, while Hammitt and Liu (2004) report an annual WTP decline of 1.5% for a latency period of 20 years.

Another topic that gets considerable attention in the VSL literature is the effect of age on VSL estimates. In recent years, more researchers seem to question the likelihood of one unique VSL across different age categories, indicating that the lower life expectancy ought to be reflected in the VSL assessments. This has led to the introduction of the term ‘senior death discount’ (Krupnick, 2007). The application of this discount in US EPA’s calculations for pollution-control regulation benefits has been heavily criticized (Viscusi, 2010). Nevertheless, European government agencies are still applying a lower VSL for older age groups in their cost-benefit analyses (CBA) (Aldy and Viscusi, 2007). A literature review of stated preference analyses on this topic suggests that there is relatively little consistent empirical evidence for such an effect (Krupnick, 2007). If there is an effect noticeable, it is usually observed for individuals older than 65 or 70 years old (Krupnick et al., 2002). Moreover, the effect seems to depend on the type of mortality under investigation. In the hedonic wage literature, which mostly considers accidents with direct mortality consequences, the size of the VSL seems to be related to age in an inverse U-shaped form, implying that younger and older employees have lower VSL estimates than middle-aged persons around 40-50 years old (Aldy and Viscusi, 2007). This is an effect that is also found in a recent DCE approach for sudden deaths situations (Cameron and DeShazo, 2013). However, the latter analysis finds different results with regard to age effects for chronic illnesses using a range of pre-mortality morbidity conditions. In this case, the WTP for a microrisk reduction of incurring an illness often increases in older age groups.

In comparison with the extensive literature on the WTP to avoid the occurrence of a fatality, there has been relatively little research on the economic values associated with morbid conditions. The most commonly applied approach for economically analyzing non-fatal illnesses is by means of the quality-adjusted life years (QALYs) that are lost because of the illness. This approach entails that the loss of QALYs is multiplied with a fixed economic value for one QALY, which is assumed to represent to monetary value society places on a QALY (Robinson, 2007). Conveniently, this approach allows to draw from the vast amount of information on QALYs associated with practically every illness in the cost-utility literature. This has spurred a flow of literature on estimating the willingness to pay for a quality-adjusted life year. In studies estimating the societal burden of osteoporosis, the most widely used values vary in the range of $ 50,000-60,000 per QALY (Borgstrom et al., 2007; Kanis et al., 2002; Landfeldt et al., 2011). However, depending on the considered illnesses, there is a lot of heterogeneity in the economic evaluations of QALYs (Donaldson et al., 2011), both on the lower end (Bobinac et al., 2010; Byrne et al., 2005; Gyrd-Hansen, 2003; King et al., 2005) as on the higher end of this range (Braithwaite et al., 2008; Milligan et al., 2010).

Analogous to the assumption of a fixed VSL across the entire population, it is rather simplistic to assume that there is a single, uniformly distributed value for one QALY. Gyrd-Hansen (2005) even reports that establishing one unique WTP for a QALY is theoretically impossible because it entails overriding individual preferences. Optimizing resource allocation requires considering each gain in QALYs as equal ("a QALY is a QALY") in view of maximizing health benefits from the scarce resources spent on health care, regardless of recipient and intervention characteristics. However, this approach fails to take into account the distributional nature of resource allocation (see chapter 1). Besides allocating resources as efficiently as possible, social welfare might also be increased by favoring an equitable distribution of health and health care (Bobinac et al., 2012). This would imply that QALYs which can be gained by certain subgroups should be assigned more or less weight than others. For example, empirical evidence indicates that people attach more value to helping severely ill patients in comparison with moderately ill individuals (Ubel, 1999). This might result from feelings of fairness and giving 'hope' to the high-cost patients (McKie et al., 2011). Another important aspect in this setting is the patient's own contribution to the ill health situation, in smokers for instance, which might justify assigning relatively less weight QALYs gained by these subgroups (Olsen et al., 2003).

In this light, other authors such as Van Houtven et al. (2006) have proposed to develop a non-linear function for transferring economic values from QALYs. Concave functions of the severity and the duration of the illness might be more appropriate for revealing accurate estimates of the WTP to reduce the risk of morbid conditions (Hammitt and Haninger, 2011). Scasny and Alberini (2012) even take it one step further and advance that, although monetary values of QALYs can be useful in health impact assessments, QALYs should not be used in the first place in a cost-benefit framework of environmental regulations because it is incompatible with economic theory and welfare analysis. They suggest that the only way to integrate health impacts in a CBA is to monetize all case-specific mortality and morbidity risks using non-market valuation methods.

4.5.2 Lung cancer

Although it is explained in section 4.1 that it is difficult to compare cost estimates between countries given the differences in national health care systems and clinical practice for treating illnesses, the findings from other COI studies can be used as a reference point. Particularly the costs of lung cancer have been thoroughly examined in many industrialized countries, since it is the most common cancer in terms of incidence and mortality in the world (Parkin, 2001) and it causes a great economic burden to society (Goodwin and Shepherd, 1998). There have been COI lung cancer studies in Canada (Evans et al., 1996), France (Chouaid et al., 2004), Switzerland (Dedes et al., 2004), the United States (Kutikova et al., 2005), Northern

Ireland (Fleming et al., 2008), the Netherlands (Pompen et al., 2009), South Korea (Park et al., 2010) and Australia (Kang et al., 2012).

The study by Evans et al. (1996) in Canada had an important pioneering function in this setting, being one of the first to establish a direct cost figure for lung cancer. This research used a modeling algorithm to prospectively calculate direct patient care costs for lung cancer patients. As in almost all COI lung cancer studies, a distinction is made between non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC), although the ranges were quite similar for both types of disease. The treatment costs for NSCLC were in the range of CAD$ 6000-18000 (+/- €4225-12676) and for SCLC the range was CAD$ 5000-19000 (+/- €3521-13380) (both in 1988 prices). None of the other previously mentioned studies reported significant differences in average costs between NSCLC and SCLC neither.

However, treatments used in this research are probably no longer applied in current medical practices given the rapid evolution in cancer treatments. Therefore, we should focus on more recent studies for reference points. Fleming et al. (2008) published hospitalization costs (in 2004 prices) for the bulk of lung cancer patients diagnosed in Northern Ireland and found that these costs amount to £5956 (+/- €8509) for NSCLC and £5876 (+/- €8394) for SCLC in the 12 months post diagnosis. Similar to other studies (Dedes et al., 2004; Kutikova et al., 2005; Pompen et al., 2009), hospital residence costs were the dominant cost factor, ranging from 62% to 84% of total costs depending on the disease type. Pompen et al. (2009) gathered all direct medical costs – not merely hospitalization costs – during a time frame of maximum 43 months, but focused on the more frequently occurring non-small cell lung cancer type. The total annual cost per patient from the hospital perspective amounted to approximately €32,000 (in 2005 prices) for this sample of NSCLC patients in the Netherlands. Park et al. (2010) followed up lung cancer patients that were still alive 5 years after diagnosis in a tertiary referral hospital in South Korea. Six percent of the original sample met this condition and incurred on average $21,000 (+/- €15,556) in direct medical costs, which made up 65% of total costs including direct non-medical and indirect costs.

4.5.3 Hip fractures

In western countries there has been a considerable amount of research on the resource costs of hip fractures as well. De Laet et al. (1999) included hip fracture cases in a prospective cohort study in a district of Rotterdam (the Netherlands) from 1991 until 1994. Hip fractures caused approximately €8600 (in 1993 prices) in incremental costs in comparison with a matching control sample in the year following the incidence. Comparable case-control research in the USA found median incremental costs of $11,241 (+/- €12,490) for hip fractures occurring in the time frame 1989-1991 (Gabriel et al., 2002). For the United Kingdom, the hip fracture cost was estimated to be approximately £12,000 (+/- €19,200) (in 1996 prices) (Dolan and Togerson, 1998), while this amounted to €19,700 in Sweden (in 1992 prices) (Zethraeus et al., 1997). Although the percentages differ between studies, a significant amount – and in most cases the greater part – of these costs are attributable to acute hospitalization costs incurred directly after the hip fracture has taken place.

In more recent years, the direct medical costs of hip fractures have also been researched in Belgium. Reginster et al. (1999) collected data from hip fracture patients in 1996 that were affiliated to one Belgian sickness fund. The researchers reported an acute hospitalization cost of €8234 and an incremental outpatient cost of €690 in the year following the hip fracture for these patients. In another Belgian study (Autier et al., 2000; Haentjens et al., 2001) with hip fracture patients from 1995-1996, the acute hospital stay resulted in similar cost estimates (€8667) to the Reginster estimates, but the incremental costs after hospitalization were much higher (€6636). This can be explained by the inclusion of all direct medical costs in the latter study, while the former study excluded important cost factors such as inpatient stays at rehabilitation units and nursing homes. Given that the average hospitalization cost for hip fractures that have occurred in 2007 is around €10,500 (Hiligsmann et al., 2011b), these cost assessments indicate that the hospitalization cost of hip fractures has increased by and large with the general price index in Belgium.

Besides the resource costs associated hip fractures, there is a considerable loss of quality of life as well. Hip fracture patients often do not return completely to their prefracture health state as a result of a loss of function and a slower rate of recovery in later stages of life. Since practically no patients die directly from the trauma, hip fractures particularly suffer from a large morbidity. Nevertheless, 20 to 25% of hip fracture patients die within the first year following the fracture, relating hip fractures to early mortality as well (Lips and van Schoor, 2005). A recent literature review on the utility that is lost due to hip fractures indicates there is a quite broad consensus that utility is most significantly affected in the time frame immediately after the fracture (Peasgood et al., 2009). Hiligsmann et al. (2008b) also reports that on average there is a QALY reduction of 0.17-0.23 in the year following the fracture, which is approximately double the utility reduction in subsequent years. Similar results were found in an earlier literature review (Brazier et al., 2002). The findings by Borgstrom et al. (2006) are situated in the lower range of this confidence interval, while values in the higher end of this range are found in Peasgood et al. (2009) and Tosteson et al. (2001).

4.6 Conclusion

In the second part of the health economic analysis the damage costs attributable to environmental pollution in the Campine region are estimated. The main objective in this chapter was to calculate the direct medical costs relating to lung cancer in Belgium using a longitudinal matched case-control design based on administrative databases. Therefore, a data extraction procedure has been performed on the Permanent Sample of IMA, an organization that comprises databases from all sickness funds in Belgium. This way, a large part of the direct medical costs of patients which have been submitted to the most important interventions leading to the diagnosis of lung cancer – excluding self-employed patients – have been extracted for the period 2007-2010. Two matching control samples were created with similar socio-economic characteristics as the patient sample to determine the incremental costs of lung cancer. Since we are particularly interested in the lung cancer incidence that was caused by environmental pollution and not by tobacco – the principal risk factor for lung cancer – the second control sample adjusted for the main comorbidities of smoking. This way, the costs attributable to smoking are excluded from the cost estimates.

The results show that lung cancer presents a significant burden in direct medical costs to patients and to society. In our 21 month follow up the patient sample (n = 359) made almost €11 million in health care costs. Although the Belgian health care system covers the largest part of these costs, incremental costs reveal that patients incur substantially higher medical costs than matching control samples. More specifically, the average lung cancer patient deals with €27,721.26 in incremental costs over a time frame of 21 months. When the comorbidities of smoking are accounted for, incremental costs are slightly lower (€25,385.60), indicating that smokers experience higher medical costs than a standard control sample. The majority of lung cancer costs are made in the prediagnostic phase – 3 months prior to diagnosis until diagnosis – and the acute phase – from diagnosis until 3 months after diagnosis. In these disease stages, patients are subjected to a lot of inpatient days and considerable costs for cancer treatments are incurred.

Resource costs for hip fractures have not been calculated ourselves, since cost estimates for Belgium have recently been made at University of Liège (Hiligsmann et al., 2011b). This research used data from a Belgian sickness fund to determine mean hospitalization costs for hip fracture treatment, which amounted to €10,828 and €10,389 in men and women, respectively. These cost estimates only include direct medical costs made in the hospitalization period after the hip fracture had occurred. When combining the cost estimates with the pollution-attributable cases determined in the previous chapter, the health care costs attributable to the pollution can be estimated. The results show that the pollution has yearly generated more than €560,000 in lung cancer costs and up to €420,000 (in 2007 prices) in hip fracture costs. As contemplated in the previous chapter, most likely the pollution-attributable cases have been diminishing over the years and will further decrease in the future. Nevertheless, given that the study is designed to provide conservative direct medical cost estimates for both illnesses, the health care costs produced by the pollution are substantial.

Although this study aimed to calculate the attributable health care costs of environmental pollution as correctly as possible, there are a number of constraints that have to be taken into account when interpreting the results. Some of these limitations are inherent to working with administrative-financial databases. For example, the largest cause for uncertainty with respect to the economic aspects of lung cancer is due to the indirect identification of the patient sample because diagnosis codes are unavailable in the IMA database. This way, it is unclear whether only patients diagnosed with lung cancer have been selected or whether other illnesses (cancers) have been included as well. There is no possibility to check this afterwards neither. The same constraint applies to the smoking-adjusted control sample, for which we were unable to check whether it were only smokers that have been selected. Another constraint concerns the delivery of results in aggregated form. Since we were unable to work with individual level data, no extra (variance) analyses could be executed. Moreover, it was unknown when the patient's medical costs have been made, obstructing the possibility of discounting the economic data to a common time frame.

Furthermore, the presented cost estimates will most likely underestimate the true costs produced by lung cancer. This can largely be explained by the (purposely) restricted research design in which we only included direct medical costs related to the illness. This way, a number of elements from the societal costs of illnesses have not been incorporated. For example, sickness fund databases only register the medical costs that are reimbursed by RIZIV-INAMI. This implies that all medical costs that are not covered by the mandatory health care insurance nor costs that fall under the responsibility of regional governments are not included in the cost estimates. Although the (indirect) costs from lost productivity are likely to be quite small given the age distribution of patients incurring the considered illnesses, there still are some opportunity costs for the patient (from lost leisure time, for instance) as well as caregivers. Moreover, the damage cost method itself is unable to include the changes in utility value caused by an illness. This implies that people's WTP to avoid disease burden is excluded from the analysis. In order to compensate for this shortcoming, some studies on the WTP to avoid mortality and morbidity have been discussed in section 4.5.

5 Analyzing the impact of soil contamination on farmland values using hedonic pricing analysis[27]

5.1 Introduction

As shown in chapter 2, the area affected by the zinc smelters included a significant amount of farmland (Schreurs et al., 2011). If agricultural soils have been contaminated, this can create a persistent and harmful problem to farmers as well as to society as a whole. Besides direct health risks for farmers working on these soils, the presence of contaminants can also generate risks to crops grown on this land and thus to food safety (Dudka et al., 1996; Hough et al., 2003b). Some foodstuffs have the capacity to easily transfer metals and other contaminants from the roots to the edible parts which can indirectly cause hazards to public health (Vromman et al., 2008). Since national as well as supranational (European) governments aim to ensure the consumer's food safety, strict guidelines have been established which food producers have to comply to in order to offer their products on the market. Farmers who try to market crops that exceed these food

[27] Parts of this chapter have been presented at international conferences and workshops:

- 8th Conference on Sustainable Development of Energy, Water and Environment Systems (SDEWES): Schreurs, E., Van Passel, S., Peeters, L. and Thewys, T. (2013). Economic impact assessment of farmland contamination: A policy perspective. SDEWES, Dubrovnik (Croatia), 22-27 September 2013.
- 7th World Conference of the Spatial Econometrics Association (SEA): Schreurs, E., Van Passel, S., Peeters, L. and Thewys, T. (2013). Analyzing the impact of soil contamination on farmland values: A spatial hedonic approach using quantile regression. SEA, Washington (DC), 10-12 July 2013.
- 5th European Association of Agricultural Economics (EAAE) PhD Workshop: Schreurs, E., Lizin, S., Van Passel, S. and Thewys, T. (2013). The impact of soil contamination on farmland values: Revealed preference versus stated preference. EAAE PhD Workshop, Leuven (Belgium), 29-31 May 2013.

thresholds are running the risk of being unable to sell their output. Consequently, they might face a severe loss of income.

In order to obtain welfare estimates on the damages caused by environmental pollution, revealed preference techniques such as hedonic pricing analysis are an attractive alternative. The method is based on the premise that the price of a differentiated good is composed of the value that each characteristic appends to the product (Rosen, 1974). Therefore, product prices, preferably from actual market transactions (Cotteleer and van Kooten, 2012), are adopted in order to extract the consumer's marginal willingness to pay for underlying characteristics. Empirical applications of the hedonic price method mainly focus on the valuation of non-marketable variables from real estate prices. Prices for housing and land are apt for application in hedonic models because it concerns multi-attribute and multi-faceted products that can be linked relatively easy to an associated bundle of heterogeneous characteristics. Moreover, the location choice can often be related to neighborhood amenities and data with regard to real estate sales or appraisals are widely available in many countries.

When investigating the impact of environmental pollution on real estate values, generally environmental quality indicators are used as a measure for the variable of interest. Hedonic studies focusing on the value of contamination often relate the level of polluting elements in an environmental medium such as soil (Clauw, 2007; Guignet, 2013), water (Leggett and Bockstael, 2000; Poor et al., 2007) or air (Kim et al., 2003; Yusuf and Resosudarmo, 2009) to property values. If objective risk measures are unavailable, the proxy that is most frequently used to account for the disamenity is the distance to a pollution source (Zabel and Guignet, 2012) or an undesirable land use such as waste sites (Braden et al., 2011). The primary objective in these studies is to estimate the welfare gain that can be achieved by removing the pollution from the environment.

However, a common aspect in these hedonic studies is that all research – for now – has focused on residential property values to estimate the economic impact. To the best of our knowledge, the effect of contamination on agricultural land prices has never been empirically analyzed in hedonic studies[28]. This chapter aims to fill this research gap by applying the hedonic methodology to the farmland market in the Campine region. This way, we want to find out to what extent agricultural land buyers value the presence of soil contaminants. First, a literature review will discuss the findings from previous hedonic studies on agricultural land. In section 5.3 we will elaborate on the hedonic model and spatial econometric techniques that can be applied to incorporate the inherent spatial nature of farmland values. The data that are used in the analysis are described in section 5.4, while the most important results will be reported in section 5.5. Furthermore, the Belgian and Flemish land policies with regard to soil contamination will be analyzed in section 5.6. In the last section, some concluding remarks will be given.

5.2 Literature review

This literature review aims to provide an overview of the most important research topics in the hedonic farmland literature. As mentioned before, hedonic pricing analysis tries to relate land values to a number of factors that are expected to have an influence on the willingness to pay for farmland. By means of this literature review we thus intend to procure the basic elements for creating a solid hedonic farmland model.

In most hedonic farmland models, the dependent variable consists of prices that have been paid for agricultural land in a certain area. While many analyses use micro-level data on individual sales transactions, farmland prices are sometimes aggregated on a higher geographic level, such as the

[28] I have searched in Web of Science for studies that have used hedonic pricing analysis to reveal the impact of soil contamination on farmland values. Therefore, all possible combinations of these keywords were used: 'hedonic (pricing/prices)', '(soil) contamination/pollution' and 'agricultural land/farmland prices/values'. Since these combinations did not return any valuable results, I tried combinations of only two out of three keywords. However, the output of this search strategy did not include any hedonic farmland studies with soil contamination as an explanatory factor neither.

county level, if it concerns large study areas. Examples include Cavailhes and Thomas (2013) and Huang et al. (2006). Another alternative to real sales prices are estimations by tax assessors or land agents. Although this option can be attractive because these data are usually more readily available (Cotteleer and van Kooten, 2012), they can merely present a satisfactory indication of farmland values and should not be used to establish determinants for land value (Ma and Swinton, 2012). Taylor (2003) adds that professional valuations are more susceptible to sample selection bias, because rarely sold, unique properties can be systematically not included in the data. Actual sales prices, on the other hand, are the result of a negotiation between buyers (demand) and sellers (supply) in the market. Hence, they are more likely to reflect the relative power of all market forces at that moment in time.

The most obvious explanatory factors for farmland prices are all inherent farmland characteristics that constitute land productivity such as soil structure and texture, topsoil depth, topographic features and so on. These features can be combined into one land quality score based on a certain algorithm (Maddison, 2000, 2009). The objective of these variables is to capture the agricultural returns that can be expected from the parcel. If no suitable indicator can be found or created using physical attributes, the agricultural income earning capacity can be used as a proxy as well (Patton and McErlean, 2003). Since the equilibrium price for a parcel of land will equal the present value of the stream of rents produced by the land according to the classical rent theory, the differences in agricultural land rents are expected to correspond to productivity differentials (Freeman, 1999). However, in practice land markets often do not comply with the assumptions of a perfectly competitive market that is presumed in theory. Common land market imperfections include transaction costs, information asymmetry and an inelastic demand and/or supply (Ciaian and Swinnen, 2006). That is why factors such as institutional and transaction variables are likely to have an impact on farmland price determination as well.

One of these institutional variables includes the regulatory framework and zoning typology in which the farmland is located. When agricultural land is expected to be redeveloped changed in the future, it has been suggested that farmland prices consist of two components, i.e. a use value reflecting the present discounted value of the future rents from agricultural production and an option value representing the extra value a buyer is willing to pay for the opportunity of redevelopment in a later stage (Plantinga and Miller, 2001). Although there is some pressure from semi-agricultural land uses such as horticulture as well, particularly in densely populated areas such as Flanders (Bomans et al., 2010), these option values are mostly situated at the urban-rural interface for agricultural land (Isgin and Forster, 2006). Generally, farmland prices increase sharply in close proximity to an urban center, but this effect tends to decrease and even disappear when land is situated further away from the city center (Cavailhes and Wavresky, 2003). Since the option value has been shown to play a substantial role in determining the price of farmland, it is nearly always appended in the empirical hedonic literature. See for example Taylor and Brester (2005), Nickerson and Lynch (2001) and Abelairas-Etxebarria and Astorkiza (2012).

However, when the potential for redevelopment is lost as a result of governmental conservation programs, farmland values may be depreciated as well (Wu and Lin, 2010). For instance, Deaton and Vyn (2010) found that imposing the "Greenbelt" legislation – a strict form of agricultural zoning preventing nonagricultural development in more than 7200 km² of farmland in Ontario (Canada) – had a significant negative impact on farmland prices within a range of 5 km from the urban Toronto area. This is likely to be a result of the option premium being cleared from agricultural land prices, seeing that the effect was not found at distance bands further away from the urban area. Therefore, the imposition of agricultural zoning to preserve farmland provokes different reactions among landowners and land users. Since land constitutes the greatest part of total farm asset value (Huang et al., 2006), landowners risk losing substantial asset value from land use restrictions. Moreover, even within the group of landowners, there might be differences with respect to preservation programs. Landowners that are net agricultural land buyers are more likely to support agricultural zoning than net sellers (Deaton et al., 2007).

Besides farming support programs and zoning regulations, there are other institutional factors that can cause variation in land prices. Focusing on member states of the European Union, Ciaian et al. (2012) found that transaction costs play an important role in farmland price setting. Explicit transaction costs contain all administrative costs that are associated with acquiring a parcel of agricultural land. Examples of such costs include notary fees and registration taxes. The higher these transaction costs, the more land buyers are stimulated to negotiate with sellers to agree to a lower 'tax-official' price, which can be complemented by an extra 'non-official' cash amount. Although most countries have installed mechanisms to control whether taxes have been eluded, it still remains a problematic issue in large parts of Europe.

Closely related to this subject is the issue of sellers and potential buyers being acquainted when transferring farmland. This is quite plausible since active farmers are only interested in farmland that is located within an 'acceptable' range from their home dwelling and the community of farmers is generally rather small in industrialized countries. Cotteleer et al. (2008), for example, showed that 90% of agricultural parcels acquired by farmers in the Netherlands are located within a radius of 6.7 km from their home. Therefore, most hedonic studies of agricultural land (e.g. Xu et al. (1993)) only include 'arm's length' transactions because the probability of avoiding transaction costs – and thus agreeing to a below-market price – is larger in non-arm's length transactions. However, a number of studies merely exclude family sales and forced sales, deeming sales between neighbors and acquaintances, for example, to be acceptable arm's length transactions (Perry and Robison, 2001). Nevertheless, excluding non-arm's length transactions seems to be a reasonable assumption seeing that personal relationships are an inherent part of farmland transactions which have a significant depreciating effect on sales prices (Tsoodle et al., 2006). Kostov (2010), building on the framework established by Perry and Robison (2001), found similar results in a more flexible, nonparametric hedonic analysis of farmland in Northern Ireland.

More recently, research has been increasingly paying attention to the contribution of agriculture to environmental amenities. Although the capacity for producing food, fuel and fiber are the main function provided by farmland, it can also add to the ecosystem’s regulating services with respect to water quality and biological pest populations for example and supporting services such as soil formation and nutrient cycling (Swinton et al., 2007). Moreover, in an ever increasing population density, people are migrating towards rural areas in search of environmental amenities such as scenic vistas and outdoor recreation, which can also be provided by agricultural land. This trend may cause a rising demand for land in rural areas as well, in their part increasing the risk of agricultural land being converted to residential areas.

Since environmental characteristics can change the productivity of land, in theory the prices of agricultural parcels will reflect these productivity changes (Freeman, 1999). There have been a number of empirical studies researching the effect of environmental variables on land prices. A substantial part of these studies focused on residential housing prices to estimate the value of open spaces to people (e.g. Ready and Abdalla (2005)). However, also for farmland values effects have been found for environmental characteristics such as climate change (Mendelsohn et al., 1994), water salinity (Koundouri and Pashardes, 2003) and ecosystem services (Ma and Swinton, 2011). Bastian et al. (2002) incorporated a number of environmental amenities into a hedonic model of agricultural land values and found that the effect differed between sorts of amenities. Correspondingly, this study aims to estimate the effect of an environmental disamenity – i.e. soil contamination – on agricultural land prices.

5.3 Methodology

5.3.1 Hedonic model for agricultural land

In Rosen's seminal article hedonic prices are described as the implicit prices of attributes which are revealed to economic agents from observed prices of differentiated products and the specific amounts of characteristics associated with them (Rosen, 1974). A differentiated product can thus be represented by a vector of *k* characteristics X = x_1, x_2, x_3,…, x_k with sales price Y. According to this theory, the equilibrium price schedule emerges from the interactions between many producers and consumers in a perfectly competitive market. The equilibrium price of any particular model is a function of its value-adding and value-reducing characteristics. Hence, the marginal implicit price of any characteristic - i.e. the additional expenditure needed to purchase an extra unit of the characteristic - can be estimated by differentiating the hedonic price function with respect to that characteristic. A standard hedonic model is defined as follows:

$$Y = \alpha\iota + \beta X + \varepsilon \qquad [6]$$

Where Y represents a 1 x n vector of property prices; α is a constant term to be estimated; ι is a 1 x n vector of ones; β is a 1 x k vector of coefficients to be estimated; X is a k x n matrix for property attributes; and ε is a 1 x n vector of normally distributed error terms.

Rosen's model describes the theoretical foundations of hedonic pricing analysis in function of differentiated consumer products based on consumer utility theory. In the real estate market, this involves applications to the residential housing market. However, when differentiated factors of production such as agricultural land are concerned, the case is comparable but conceptually different. In this setting, land serves as an input factor to the agricultural production process. Therefore, land value has to be integrated in the farm's profit function instead of the individual's utility function, which should enable the farmer to select the parcel of land that allows him to maximize the firm's profit.

Palmquist (1989) developed an analytical model for the agricultural land market using agricultural rent prices. Since individuals that rent farmland will only be interested in the land's current productive capabilities, the equilibrium rent schedule will be determined by variables related to the land's productivity. However, in a competitive market the value of the land as an asset is determined by the present value of the future stream of rents produced by this land. Besides productivity variables, other characteristics that indicate a parcel might be more attractive or valuable in the future will affect the land's sales price while rental values will remain unchanged. Correspondingly, factors that are important in an agricultural setting might be discounted if that characteristic is not that important in an alternative land use that is anticipated in the near future (Palmquist and Danielson, 1989). In summary, differences in land values are accounted for by differences in the productivity of the land as well as the buyer's expectations with regard to its future development (Plantinga et al., 2002).

Since our hedonic model includes farmland sales prices, Palmquist's theoretical model should be adjusted in order to deal with sales prices instead of rental prices. Petrie and Taylor (2007) easily resolved this by assuming the rental price R is a simple transformation of the sales price P. More specifically, if it is assumed that all farmers can apply to the same market-clearing interest rate, annuities of sales prices can be considered equivalent to rental prices. The landowner can be considered to rent farmland to himself for agricultural production.

$$R(X) = R(P(x_1, x_2, x_3, \dots, x_k)) \qquad [7]$$

This way, the property costs are described as a rental price schedule R(X) that is dependent on the vector X, which includes all *k* characteristics that affect sales prices and thus is not limited to only those factors that affect rental prices. Theoretically, the derivation is the same.

On the demand side of the property market, there are individuals that are interested in using agricultural land as an input factor for agricultural production. For simplicity, assume there is only one agricultural output product Q. The production function will then correspond to:

$$Q = Q(X, Z, \gamma) \qquad [8]$$

where X represents the vector of land characteristics that affect land value, Z is a vector of all inputs unrelated to land, such as irrigation equipment, and γ is a vector of specific farmer skills.

Initially, it is assumed here that the farmer focuses on maximizing the 'variable profits' π^V, which is defined here as the difference between the value of the output and the value of all non-land related inputs[29].

$$\text{Max } \pi^V = \pi^{V*} = M \cdot Q - C \cdot Z$$

$$\text{Subjected to } Q = Q(X, Z, \gamma) \qquad [9]$$

where M represents the market price for the output Q and C is a vector of the unit costs for non-land inputs Z.

The farmer will then try to optimize his level of non-land inputs (vector Z). From equation [9], it follows that the demand function for these inputs is determined by

$$Z = Z(M, C, X, \gamma)^{30} \qquad [10]$$

Consequently, substituting equation [10] into equation [9] gives π^{V*} and allows the total profit function π to be calculated by subtracting the property costs R(X) for the new piece of farmland from the variable profits.

$$\pi = \pi^{V*} - R(X) = M \cdot Q(X, Z, \gamma) - C \cdot Z(M, C, X, \gamma) - R(X) \qquad [11]$$

[29] As the characteristics of land do not all require fixed expenditures, some of these costs should actually be included as well when calculating the variable profits. However, for the sake of the argument, they are provisionally excluded here.
[30] Since Q = Q(X, Z, γ) is part of equation [9], it follows that Z is also dependent on X and γ.

The profit function is assumed to be strictly concave and at the optimal level of non-land inputs. In order to maximize profits, the farmer will then select a parcel of farmland for which the marginal rent for a characteristic equals its marginal contribution to variable profits. In other words,

$$\delta\pi/\delta x_i = 0$$

$$\delta R(x_i)/\delta x_i = \delta\pi^{V*}/\delta x_i = M\cdot\delta Q/\delta x_i - C\cdot\delta Z/\delta x_i \quad [12]$$

Alternatively, the farmer's bid function θ can be determined by introducing the farmer's desired level of profits π^D. More specifically, the bid function is found by taking the difference between the variable profits and the level of profit the farmer wants to achieve.

$$\theta(X, M, C, \pi^D, \gamma) = \pi^{V*} - \pi^D = M\cdot Q(X, Z, \gamma) - C\cdot Z(M, C, X, \gamma) - \pi^D \quad [13]$$

The optimal bid for a certain land characteristic x_i thus requires taking the partial derivative of the bid function θ.

$$\delta\theta/\delta x_i = M\cdot\delta Q/\delta x_i - C\cdot\delta Z/\delta x_i \quad [14]$$

The marginal bid for that characteristic equals its marginal contribution to variable profits, assuming non-land input factors are at the optimal level. This way, bid curves for all land characteristics can be derived. Marginal bids for a characteristic will be higher than zero in case desirable land attributes such as soil fertility are concerned. It will be negative in case of unwanted land attributes such as soil pollution.

On the supply side of the property market, there are landowners that wish to maximize the profits from the land they are selling. Therefore, the vector of land characteristics X is assumed to be divided into two subvectors $X_1 = x_1,\ldots, x_j$ en $X_2 = x_{j+1}, \ldots, x_k$. The former vector represents all characteristics that are exogenous to the landowner, such as soil type or major topographical features, while the latter vector includes characteristics that can be manipulated by the landowner at some cost, such as soil productivity by means of fertilizers[31].

$$\text{Max } \pi^S = R(X_1, X_2) - C(X_1, X_2, F, \varphi)$$

$$\text{Subjected to } \pi^S \geq 0 \qquad [15]$$

where π^S represents the profits of the landowner, $R(\cdot)$ is the rent schedule from equation [7], $C(\cdot)$ is a joint cost function, F is a vector of input prices and φ is a vector of farmer-specific technical parameters. F includes prices of inputs that the landowner uses to manipulate land characteristics, while φ can include parameters such as the availability of parcels near the parcel of interest.

An offer function ψ can be derived analogous to the bid function θ derived in equation [13].

$$\psi(X_1, X_2, \pi^{S'}, F, \varphi) = \pi^{S'} + C(X_1, X_2, F, \varphi) \qquad [16]$$

where $\pi^{S'}$ is the desired profit level for the landowner.

A landowner can maximize his profits if the marginal offer prices for characteristics in vector X_2 equal the marginal costs for these characteristics in the market and if the supply of land is not completely inelastic. Then, the equilibrium price schedule is determined by the interactions between farmers – i.e. the bid function – and landowners – i.e. the offer function – in the market. In case the supply of land is inelastic – which is a reasonable assumption – the offer functions are superfluous and bid functions are sufficient to derive equilibrium prices (Freeman, 1979). The equilibrium price schedule is exogenous to both the demand as the supply side of the property market. Hence, neither individual farmers nor individual landowners are able to shift the equilibrium schedule from equation [1].

[31] The classification in two subvectors is a simplification of the real life situation. In reality, most land characteristics will not be *purely* exogenous or endogenous to the landowners.

Whether the hedonic model is able to estimate the value of changes in characteristics of farmland is also dependent on the extent of the impact (Palmquist and Danielson, 1989). If the change in a certain characteristic is only affecting an individual landowner or a small number of parcels within the land market, the equilibrium price schedule will remain constant because it will not have a significant impact on other parcels in the land market. In contrast, if the changes affect land prices for the greater part of the land market or the entire land market – for example, when national or regional land policies are altered – the price schedule will shift to a new equilibrium. In this case, it is only possible to derive the upper bound on the value of land improvements. Therefore, the hedonic method is more suitable for localized externalities (Palmquist, 1992).

5.3.2 Spatial effects

Regression analyses using ordinary least squares (OLS) are widely used to estimate explanatory effects in hedonic models. However, since analyses using real estate data are inherently spatial, the presence of spatial effects might cause bias in OLS regressions if assumptions of constant variance across the sample or uncorrelated error terms are violated. Spatial effect issues can generally be classified into two types of effects, spatial heterogeneity and spatial autocorrelation (also referred to as spatial dependence, the terms are used interchangeably). Spatial heterogeneity refers to a variation in relationships over space and usually arises in the form of spatial heteroskedasticity or spatial instability. Spatial autocorrelation concerns the functional relationship between what happens at one point in space and what happens elsewhere.

Generally, spatial autocorrelation presents itself in two forms, i.e. spatial lag dependence and/or spatial error dependence. The models frequently applied for taking into account these forms of spatial dependence are referred to as spatial autoregressive model (SAR) and spatial error model (SEM), respectively. Spatial lag dependence exists when there is a direct spatial relationship between the dependent variables in the model. In case of real estate this usually implies the price in a certain location depends on the prices of neighboring locations. Spatial error dependence occurs when there is spatial interaction in the residuals, i.e. the error terms are correlated across different spatial units. Mostly, this is caused by omitted or unobserved variables that have spatial patterns (Anselin, 1988).

Practically all hedonic real estate applications take into account spatial dependence using a spatial weighting matrix W. This matrix exogenously defines the weight that is attributed to spatially close observations in the averaging procedure. Therefore, it is generated a priori based on ad hoc assumptions about the spatial dependencies underlying the data. There are several ways to create such a weight matrix, all having their advantages and disadvantages. In contiguity matrices, weights take on a value of 1 if the observation is adjacent to another observation. Otherwise, the weights are assigned a value of 0. Nearest neighbors matrices assign a value of 1 to *k* observations that are most close in space to another observation. Actual distances between observations are used to produce inverse (squared) distance matrices.

If spatial autocorrelation tests such as Moran's I test or Lagrange Multiplier tests point out there is spatial dependence present in the model, standard spatial econometric techniques suggest a spatial model should be specified. The general spatial model can be described as:

$$Y = \alpha\iota + \rho WY + \beta X + \varepsilon \qquad [17]$$

$$\varepsilon = \lambda W\varepsilon + u \text{ with } u \sim N(0,\sigma^2 I) \qquad [18]$$

This model builds on the general hedonic model from equation [6]. W is a n x n spatial weight matrix; ρ is the spatial lag coefficient that indicates whether Y is dependent on a weighted average of prices from nearby observations weighted by W; λ is the spatial error coefficient that has been introduced to indicate whether the residuals are spatially dependent; u is an normally distributed error term[32].

The use of spatial weighting matrices has always been somewhat contested because it exogenously imposes a restrictive spatial structure to the model (Brady and Irwin, 2011). Since there is little theoretical guidance for choosing the most appropriate matrix, the choice for a weight matrix will be largely based on the researcher's assumptions and beliefs with respect to the spatial dependencies between observations. Moreover, the spatial weight matrix is a highly influential factor for the regression results (Dubin, 1998). Nevertheless, with the justification that any spatial modeling is better than completely ignoring the problem, the approach is employed in many hedonic models in order to control for spatial autocorrelation.

In recent years, however, the issue whether it is useful to perform standard spatial econometric techniques have been increasingly put forward in critical papers (Pinkse and Slade, 2010). In a recent issue of the Journal of Regional Science (issue 2/2012) a number of papers were published asking the question whether these standard techniques ought to be abandoned completely in empirical applications (Partridge et al., 2012). Although some of these papers admit being set up deliberately provocative (Gibbons and Overman, 2012), they all present some fundamental critiques on these

[32] Besides SAR and SEM models, there are other models to account for spatial autocorrelation as well. One example of such a model is the spatial Durbin model. This model assumes that the dependent variable can be predicted by means of a spatial lag of the dependent variable <u>and</u> spatially lagged independent variables. In algebraic terms, this can be expressed as:

$$Y = \alpha\iota + \rho WY + \xi WX + \beta X + \varepsilon$$

More specifically for this study, the spatial Durbin model assumes that buyers of farmland do not merely consider the prices of other farmland sales nearby, but also the characteristics of these parcels. Of course, there is much to say for this hypothesis, particularly for models in which structural attributes (such as soil quality, drainage systems, house characteristics,…) are important price determinants. However, anticipating the model specification and the results in section 5.5, our hedonic model includes relatively little structural land attributes (partly because of the lack of variation in such a small study area, see section 5.4) and even less that seem to affect farmland prices in the area. That is why we decided not to discuss spatial Durbin models any further.

techniques. Paramount among these criticisms were identification issues and the lack of theoretical foundation associated with spatial models. Furthermore, the inability for standard techniques to differentiate between spatially correlated outcomes and spatial causality (where the outcome in region A causes other outcomes in region B) is another commonly asserted problem.

McMillen (2012) attests that these procedures are merely a convenient way to control for unknown sources of spatial autocorrelation in residuals and dependent variables. Standard techniques will produce significant estimates of spatial lag operator ρ and spatial error operator λ every time models are misspecified. Seeing that this is practically always the case in parametric models, the author asks himself "how predicted values of the dependent variables can prove *not* to be statistically significant as an explanatory variable". This statement clearly refers to the spatial lag model, but is equally applicable to the spatial error model. In his opinion, standard spatial econometric techniques can serve as an additional test to detect various forms of model misspecification, but it is unlikely that these models will impose the correct parametric form to the data and control for spatial autocorrelation. Rather, he advances that nonparametric techniques such as local weighted and geographically weighted regression are more suitable to accurately accomplish spatial smoothing.

In this chapter the primary objective is to discover the effect of environmental risk variables on agricultural land values in the Campine region. With the aspiration of identifying these effects as correctly as possible, spatial econometric techniques are an indispensable tool to control for spatial spillover effects possibly present in the data. However, it is not our goal to establish any groundbreaking modeling approach with regard to the weaknesses in current spatial econometric research. Nevertheless, we will maximize our efforts to take into account the flaws of standard spatial econometric techniques indicated in previously mentioned papers as well as the suggestions made by the same authors.

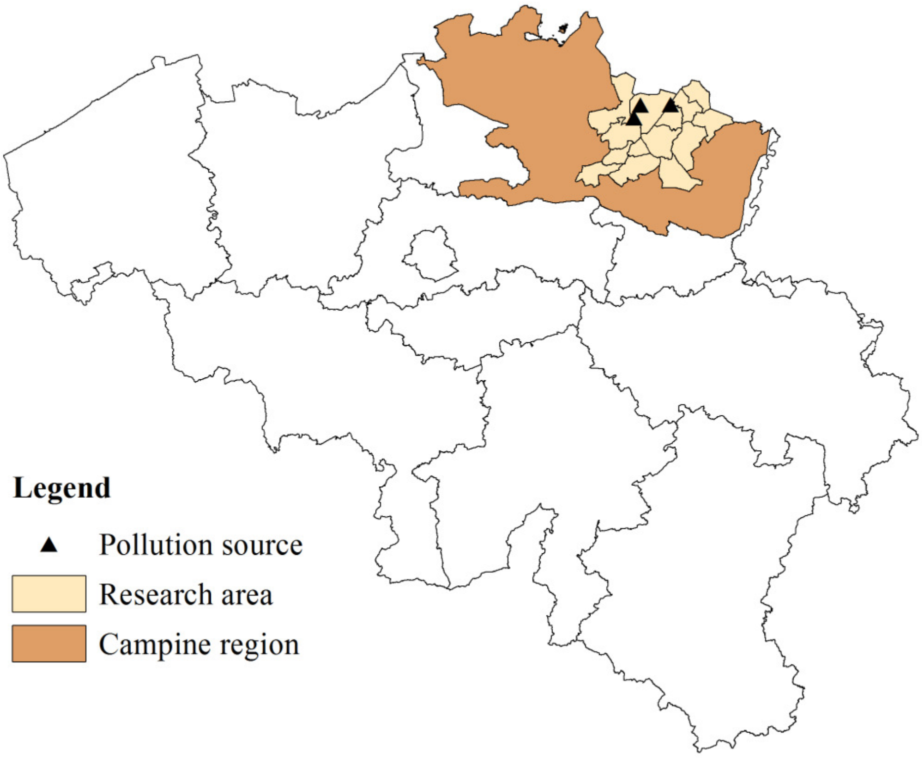
Unlike many empirical applications of hedonic models, we will not simply apply spatial models because spatial autocorrelation tests point out it improves the model's goodness of fit. Supposedly, paying sufficient attention to location variables impacting farmland prices can control for a considerable amount of spatial autocorrelation. Furthermore, it is quite remarkable how some researchers seem to think future observations might be able to impact current observations in price setting. As suggested by other researchers (Maddison, 2009; Pace et al., 2000), adding a temporal weight matrix might be a good idea to particularly focus on the spatial causality of spatial lag dependence. Therefore, we will adopt a spatio-temporal framework in which only preceding sales can have non-zero values assigned in the weight matrix.

In order to test the robustness of the results, linear regression will be complemented by quantile regression (QR). More specifically, conditional median effects of explanatory factors are estimated, while classic linear regression presents conditional mean estimates. The largest advantage of QR is that it provides much more information in comparison with mean estimates. This allows us to get a more comprehensive view of the relationship between variables. Moreover, the results are more robust against heteroskedasticity and outliers in the response variable. Additionally, since QR estimates can be presented alongside OLS estimates (McMillen, 2003), it enables us to compare the results between estimations, both with and without the spatial lag operator. Therefore, the main objective in this setting lays with checking the robustness of the estimates, not favoring one single approach over another.

5.4 Data

The Cadastre, the governmental institution which is responsible for the registration of real estate transactions in Belgium, provided the data for all farmland transactions that have occurred in the research area during the time frame 2004-2011. The research area included fourteen communities in the Campine region that have been affected to some extent by the contamination or that are within a range of 10 km from the pollution sources (Figure 5.1). The dataset that was obtained included price, lot size, cadastral information, sales method and date of 651 farmland transactions. Fifty two transactions were excluded because they lacked information or because the transaction could not be located geographically due to cadastral numbers that have been altered in the meanwhile. In the end, 599 sales transactions were maintained for statistical analysis. The nominal prices were adjusted to real prices of 2011 using a monthly indicator for the Consumer Price Index (CPI) in Belgium to correct for inflation. The cadastral information was used to georeference the agricultural parcels in ArcGIS 10.0.

Figure 5.1: Research area hedonic pricing analysis



The explanatory factors can be divided into four broad subclasses. Seeing that the goal of this chapter is to determine the impact of soil pollution on farmland prices, the principal explanatory factor in the hedonic model is the measure for *environmental risk*, which is provided by the predicted soil Cd concentrations (see chapter 2). Although a lot of studies on this subject use the distance to the pollution source as an environmental quality indicator, in this setting soil contamination levels are probably more useful because this indicator determines the land use restrictions faced by the landowner. Farmland parcels that exceed the agricultural threshold value of 2 ppm can be considered as contaminated and are therefore expected to be depreciated in comparison with 'clean' farmland. The distribution of Cd concentrations is displayed in Table 5.1 and indicates that the amount of observations with elevated soil Cd levels in the dataset is rather limited. Merely 7% of observations exceed the 2 ppm threshold.

Table 5.1: Distribution Cd concentration

Cd concentration (ppm)	**Observations** (#)
< 0,50	69
0,51 - 1	332
1,01 - 1,50	124
1,51 - 2	30
2,01 - 2,50	19
> 2,50	25

Furthermore, *agricultural land characteristics* such as lot size might affect the willingness to pay for farmland. Another potential influencing factor involves the current land use. Flemish farmers have to report what was cultivated on every parcel of land dedicated to agricultural activities as a result of strict fertilizing restrictions. Consequently, it is observable which crop was cultivated on each parcel at the time of sale. Approximately half of them were used as pastures. Most of these parcels were labeled as permanent pasture, which implies that they are not included in a crop rotation scheme for at least five years. This may be due to the ratio of permanent pasture vis-à-vis complete agricultural land the farmer is obliged to maintain as a result of the European Common Agricultural Policy (CAP). Mostly, land which is least suitable for crop production will be allocated to pastures in order to comply to this condition. However, some parcels are required to stay pasture permanently because of other regulations. In Flanders this is, for example, the case in parcels that are considered 'historic permanent pastures'. In any case, the agricultural potential of pastures is rather low. Additionally, when the land transaction accommodated a stable according to cadastral information, this was controlled for by means of a dummy variable as well.

The research area is too small for any substantial variation in climate variables or topographic features. With regard to soil productivity, the analysis is quite similar. As mentioned in section 2.3.2, the Campine region is described as an 'agricultural area' because of the homogeneity of its soil characteristics. Nevertheless, we still attempted to integrate a land quality variable by means of a scoring system that has been developed by the Flemish Land Agency (VLM)[33] for assessing which farmland is most valuable to the agricultural sector. One of the parameters in this scoring system is a measure for agricultural soil productivity based on five physical attributes, i.e. suitability for agriculture, erosion risk, texture, drainage class and flooding risk. However, the inclusion of these land quality scores in the hedonic model did not generate any significant results. There are two possible explanations for this result. Either the indicator was not a suitable proxy for soil productivity, or the farmers did not particularly take into account soil productivity because the soil structure in the region is rather homogenous anyway, i.e. there is too little variation in soil productivity for making a difference in price determination. That is why in the end no productivity variable has been appended in the hedonic model.

Other explanatory factors include variables that are specifically related to *farmland's location*. These variables do not address attributes of the parcel itself, but aspects from its broader surroundings that might have an impact on the willingness to pay for farmland. With respect to these variables the zoning typology farmland is located in might be an important price determinant. Given the scarcity of land in a densely populated region as Flanders, there is an increasing pressure on agricultural land for redevelopment into other land uses such as residential, nature and industrial zoning. Consequently, land that historically has been dedicated to farming can have its zoning type adjusted by means of governmental decisions. Although most observations were still located in an agricultural zoning type at the time of sale, there are a number of parcels that have been sold in a

[33] The Flemish Land Agency is an independent institution within the Flemish Government that aims at improving environmental quality by integrating ecological, economic and societal interests in land investments.

non-agricultural zoning type (Table 5.3). These observations particularly include residential and nature zoning types. This can be an indication that the land is destined to serve other purposes after the transaction.

As already discussed in section 5.2 and relating to the previous subject, an important price effect in farmland is due to urban sprawl and rural redevelopment. The level of urbanization is mostly accounted for by introducing the distance to the central business district as a proxy for potential future redevelopment. However, this variable was found to be inappropriate for our case study, since there is no clear urban center in the vicinity of the research area and practically every piece of farmland is on the urban-rural fringe in Flanders because of the region's high population density. That is why the number of housing units within a radius of 1,000 m (address density) from each observation is inserted to account for the speculation effect. An aspect that is quite specific to this case study is the proximity of the Dutch border. In the Netherlands agricultural land prices are roughly 50 to 100% higher than in Belgium. It might thus seem reasonable for Dutch farmers to cross the border in order to buy agricultural land at a substantially lower prices than in the Netherlands. Agricultural parcels in close proximity to the Dutch border are expected to be more attractive to Dutch farmers.

The last group of *transaction variables* comprise factors relating to the actual transition of land from seller to buyer. Although nominal sales prices are adjusted to real prices, the introduction of a time trend might be helpful to control for yearly influences on land prices. Additionally, the sales method – either private or public – may be an important aspect in Belgium. In private transactions the parties involved might be engaged in mutual interests or having personal relationships. This increases the probability that price negotiations will result in agreeing to a below-market price. Excluding all private transactions, and hence all non-arm's length transactions, was not possible since merely 18% of all observations were public sales.

Table 5.2: Descriptive statistics continuous variables

Variable	**Unit**	**Mean**	**Std. dev.**
Real price	€ ha^{-1}	28248.82	23462.39
Nominal price	€ ha^{-1}	25390.41	20964.53
Cd concentration	ppm	1.01	0.69
Lot size	m²	14005.55	45523.33
Address density	# housing units in 1 km radius	413.68	339.64
Distance to Dutch border	m	10463.51	6135.95

Table 5.3: Descriptive statistics dummy variables

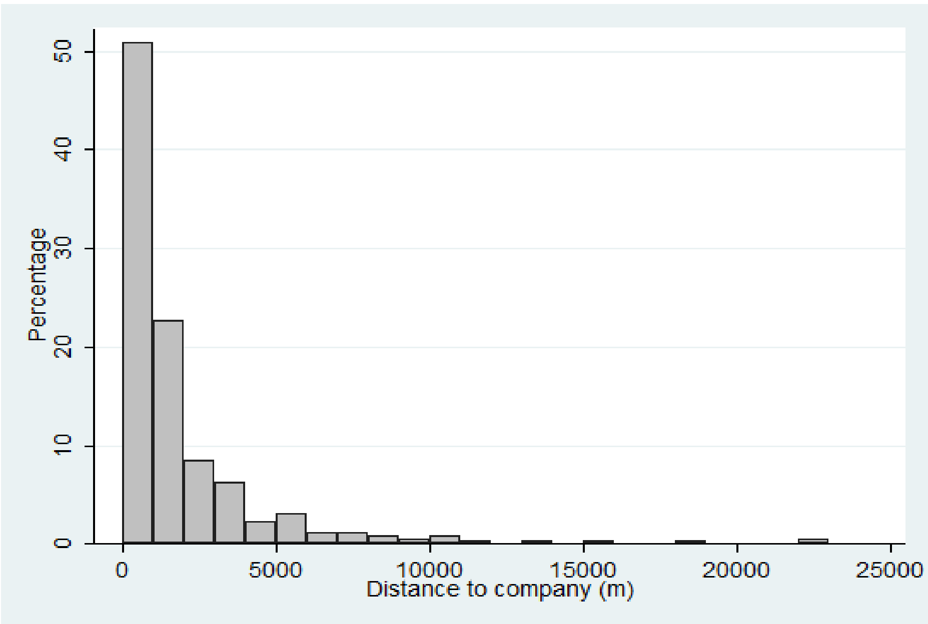
Dummy variables	**Count**
Building	31
Pasture	252
Residential zoning	47
Nature zoning	65
Public sale	110
Year 2004	50
Year 2005	100
Year 2006	91
Year 2007	80
Year 2008	69
Year 2009	68
Year 2010	76
Year 2011	65

5.5 Results

5.5.1 Spatial autocorrelation

As contemplated earlier on, the choice for a spatial weight matrix is rather arbitrary. This has aroused some suspicion with regard to this technique for detecting spatial autocorrelation. In order to partly circumvent this uncertainty and as suggested by McMillen (2012), multiple spatial weight matrices are created. This also allows us to get a better grip on the spatial structure that is underlying the dataset. Therefore, the distance of sold parcels to the farmer's home address was obtained from the Flemish Land Agency (VLM) and used as a guideline in setting up these matrices. However, this distance measure was obtained for only 353 observations[34] because the new owner's location was not always known to VLM. This may be due to – among other things – non-Belgian land ownership or non-agricultural usage after the sale. The histogram of the partial dataset's distance to farm shows that most farmers search for potential new parcels in a rather narrow range from their home address. More than 50% of the parcels are sold within a range of 1 km from the farm, while 80%, 90% and 95% of the sales have occurred within a radius of 3 km, 5 km and 7 km, respectively.

Figure 5.2: Histogram distance parcel to company



Although the farmer's search radius for purchasing farmland is not exactly the same as the range within which he will search for prices of previously

[34] This was also the reason for excluding the variable from the regression.

sold comparable parcels, the histogram in Figure 5.2 does indicate that farmers will primarily be interested in parcels nearby to their residence. Most likely, they will consult prices from parcels sold within a similar range as reference points. That is why it was decided to create an inverse distance matrix and contiguity matrices with a cut off distance of 1 km, 3 km, 5 km and 7 km. The former matrix is less appropriate in large geographic areas, but apt for small datasets of agricultural land values (Cotteleer et al., 2011). All matrices were standardized and created in ordinary spatial form as well as in spatio-temporal form. Spatio-temporal matrices only include observations that have occurred before the observation under consideration. This renders them intuitively more appealing than the ordinary spatial weight matrices, since future land sales cannot be used as a source of comparison for price setting. Moran's I test was used as a global index for spatial autocorrelation in real prices per square meter.

Table 5.4: Moran's I test on real prices per square meter

Weight matrix	**Spatial**		**Spatio-temporal**	
	I	z	I	z
Contiguity 1km	0.189^{***}	5.954	0.274^{***}	8.377
Contiguity 3km	0.030^{***}	2.358	0.081^{***}	4.867
Contiguity 5km	0.011^{*}	1.420	0.059^{***}	4.968
Contiguity 7km	0.006	1.263	0.035^{***}	3.796
Inverse distance	0.051^{***}	7.178	0.103^{***}	12.683

Results in Table 5.4 show that all weight matrices are highly significant, except for the spatial contiguity matrices with a cut off distance of 5 km and 7 km. Positive Moran's I values indicate that farmland prices are spatially clustered. However, judging from Moran's I and its complementary z-values significance is declining from the 1 km contiguity to the 7 km contiguity matrix, indicating that spatial dependence is particularly situated in parcels located nearby to one another. The inverse distance matrix, which is highly significant as well, attributes greater weights to parcels that are located closer to another observation. Therefore, this weight matrix seems to be most suitable to account for the spatial dependence in this dataset.

5.5.2 Regression results

There is little theoretical guidance with regard to the choice of the functional form in hedonic models (Taylor, 2003). A kernel density plot of real prices per square meter (Figure 5.3a) shows that farmland prices are positively skewed, indicating that the dataset includes some outliers with very high prices. When taking the logarithm of real prices per square meter (Figure 5.3b), the values are more closely resembling a normal distribution, which encouraged us to choose a log-linear functional form for our hedonic model.

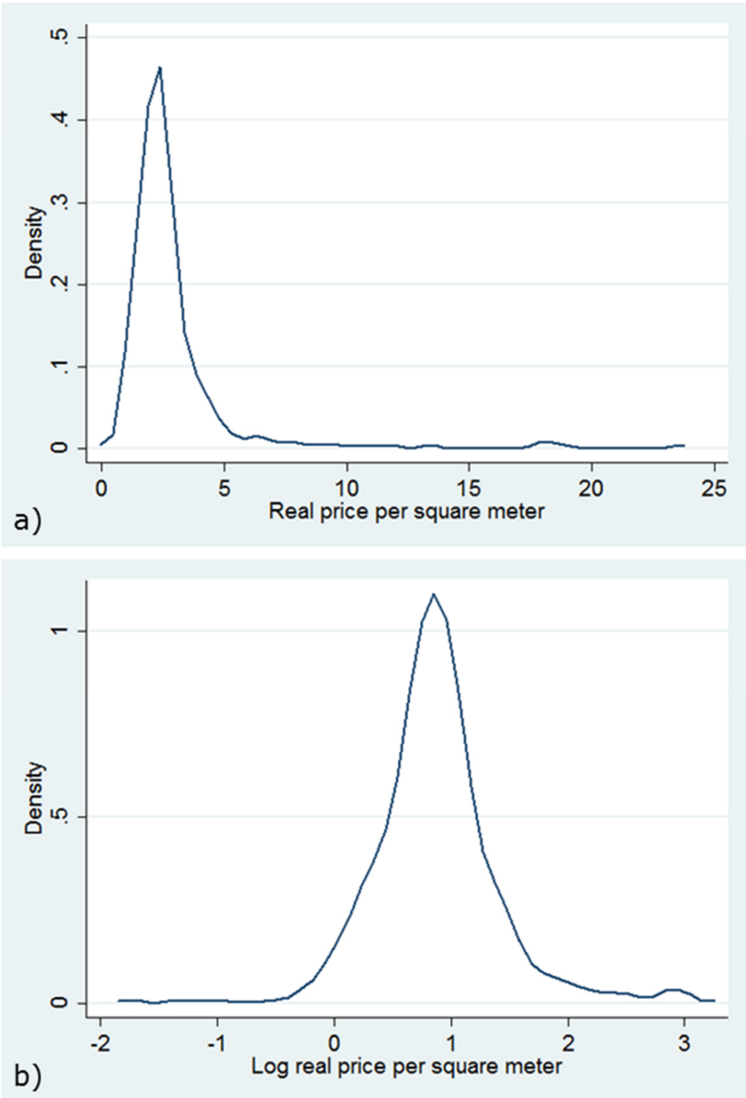
Some additional tests were performed to confirm whether this was likely to be the correct specification of the model. Box-cox transformations pointed out that transforming the dependent variable into its logarithmic form is probably the best option. Although the results in Table 5.5 show that the test using the logarithm of real prices per square meter (theta = 0) is also rejected, its log likelihood ratio is much better than the linear (theta = 1) and the reciprocal (theta = -1) case. Lastly, a Ramsey RESET test on the basic log-linear hedonic model (model 1) confirmed that the model is correctly specified[35].

Table 5.5: Box-cox transformations

Test H0:	Restricted log likelihood	LR statistic	p
theta = -1	-1231.54	614.83	0.000
theta = 0	-927.86	7.47	0.006
theta = 1	-1295.99	743.72	0.000

[35] F(3, 585) = 1.80. This result is conditional on the explanatory factors that are currently included in the model.

Figure 5.3: Kernel density estimate of (a) real prices per square meter and (b) log real prices per square meter



In the regression analyses four models have been estimated. The first two models are classic linear regression models using ordinary least squares (OLS) to estimate the effect of explanatory variables on farmland values. The distinction between both linear models arises from the inclusion of a spatio-temporal lag coefficient θ that aims to control for spatial spillover effects in model 2. As mentioned in section 5.3.2, in a spatio-temporal framework the operator will be particularly focused on explaining the spatial causality effects that nearby observations are enforcing. The inverse distance matrix created in Table 5.4 is used to construct the spatio-temporal lag operator. However, only observations that predate the current observation up to one year are included in its construction. There are two major reasons for this choice. First of all, it is expected that the determination of current sales prices will primarily be influenced by preceding sales that occurred most recently. Secondly, in an ordinary spatio-temporal form the most recent observation will have n-1 nonzero weights in the matrix, while the oldest observation will only contain zero weights. In order to maintain a similar effect on all observations, all observations should be treated equally. This can be achieved by including all observations preceding the current observation by maximum one year.

$$W = S \odot T \qquad [19]$$

$$Y = \alpha\iota + \theta WY + \beta X + \varepsilon \qquad [20]$$

S is a n x n spatial weight matrix where the weight of object e_{ij} is determined by the inverse distance between location *i* and *j*; $\odot$ represents the Hadamard multiplication operator; T is a n x n temporal matrix where the value of object e_{ij} is 1 if observation *i* preceded observation *j* by maximum one year and 0 otherwise; W is a n x n spatio-temporal weight matrix that results from the Hadamard product of matrix S and matrix T; θ is the spatio-temporal lag coefficient. The other parameters in equation [20] have already been used in equation [17].

The proposed construction of the spatio-temporal lag operator also implies that sales transactions for which there was no complete year of previous land sales available are excluded. Since the oldest transaction in our dataset took place in February 2004, all sales prior to February 2005 (n = 64) were left out. Furthermore, it is acknowledged that the temporal restriction of one year is also chosen in an arbitrary fashion. However, this seemed to be the most reasonable way to compromise between losing observations and taking into account spatio-temporal effects. This way, possible spillover effects from previous sales are controlled for, but no further restrictive assumptions are imposed to the model.

The model to be estimated is specified as follows:

Log P = $\alpha + \theta$ W Log P + β_1 Cd concentration + β_2 Lot size + β_3 Structure + β_4 Pasture + β_5 Residential + β_6 Nature + β_7 Address density + β_8 Distance to Netherlands + β_9 Public sale + β_{10} Time trend + ε

In model 3 and 4 conditional median effects are estimated by means of quantile regression. Similar to model 1 and 2, the spatio-temporal lag operator is excluded and included in model 3 and 4, respectively.

Table 5.6: Regression results

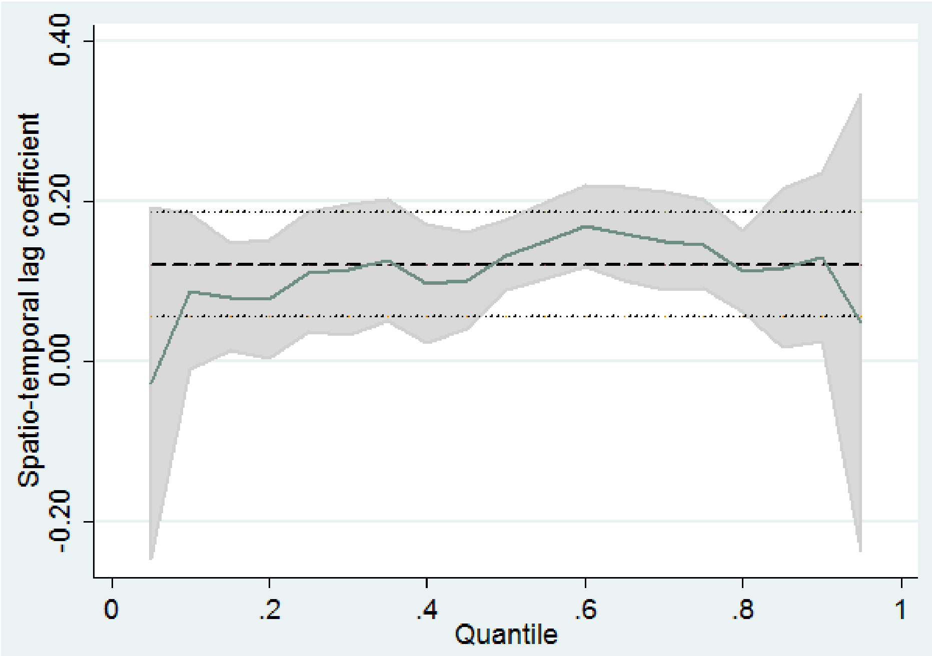
Variables	Mean regression (OLS)		Median regression (QR)	
	(1)	(2)	(3)	(4)
Cd concentration	0.008	0.008	0.017	0.021
Lot size	-1.23E-7	-2.70E-7	-2.74E-7	-3.70E-7
Structure	0.123	0.143	0.216***	0.225***
Pasture	-0.166***	-0.142***	-0.119***	-0.104**
Residential zoning	0.336***	0.266***	0.189***	0.169**
Nature zoning	-0.117	-0.095	-0.152***	-0.136**
Address density	2.881E-4***	2.128E-4***	1.527E-4***	6.679E-5
Distance to Netherlands	-1.717E-5***	-1.521E-5***	-1.322E-5***	-1.191E-5***
Public sale	0.209***	0.240***	0.214***	0.236***
Time trend	0.026***	0.027**	0.033***	0.033***
θ	/	0.121***	/	0.133***
Constant	-50.753***	-53.694**	-65.097***	-66.084***
N	599	535	599	535
R^2	0.190	0.193	/	/
Adjusted R^2	0.176	0.177	/	/
Pseudo R^2	/	/	0.102	0.105

*, **, *** represents significance at 10%, 5% and 1% level, respectively.

In none of the models significant results are found for the environmental risk variable (Table 5.5). The estimates show that soil Cd concentration is not exerting a negative influence on agricultural land prices in the Campine region. Seemingly, farmland buyers do not value the presence of soil risk factors in the prices they are willing to pay for agricultural land. Assuming that all parcels are bought for agricultural purposes, farmers might not intend to use the acquired land for hazardous crops. In this case, farmers would solely be missing out on the option value of converting the soil to different crop cultivations in the future. In fact, a lot of (animal) farmers in the area need a certain amount of farmland to dispose of their manure in order to comply with fertilizing restrictions. For them, it is not particularly relevant whether soil Cd concentrations are elevated. Moreover, the other explanatory factors might be plainly more decisive in determining farmland prices. However, another explanation might be a lack of awareness among farmers with regard to the presence of soil contaminants. If soil certificates lack the necessary information concerning the polluting elements, farmland buyers are unable to take their presence into account in price setting. This explanation will be explored more thoroughly in the policy analysis, which will be provided in the next section.

The introduction of the spatio-temporal lag operator θ seems to strengthen the model's capacity for explaining agricultural land prices. Given that the coefficient of the operator is significant at the 1% level in both models, it is likely that a spatial spillover effect is at play in the regional farmland market. Since only observations that have occurred in the year prior to the sale under consideration are included in θ, it is prevented that implausible relations with other observations such as future sales and sales from a long time ago are made. Quantile estimates of θ in Figure 5.4 show that the effect of θ is relatively constant across the distribution of the dependent variable, except for the lower quantiles. Spatial dependence seems to play a less important role in lower price segments of the farmland market. This operator might be a suitable way to take into account spatial lag effects without imposing the critized spatial lag model.

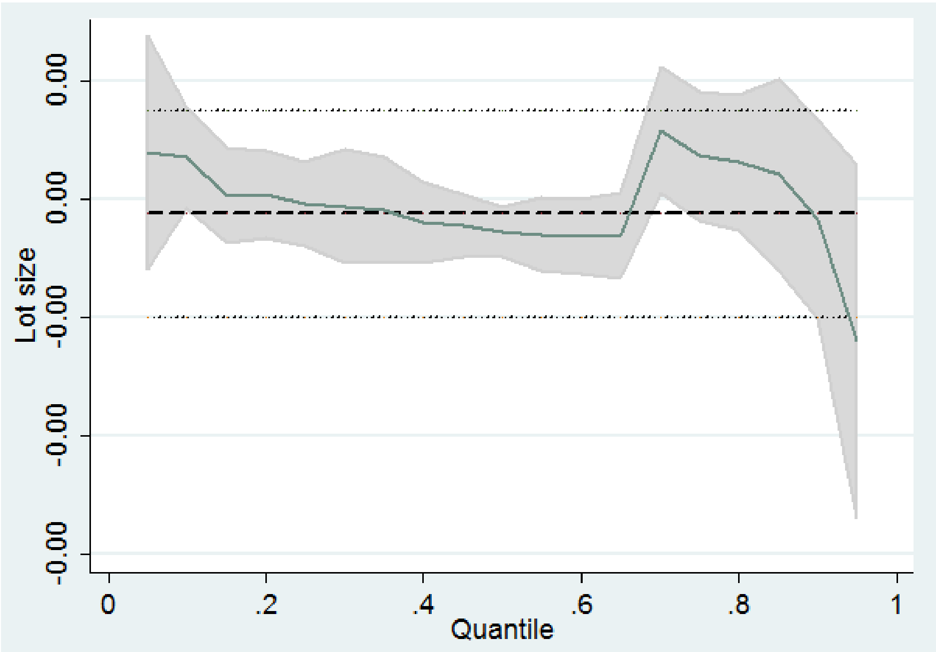
Figure 5.4: Quantile estimates of spatio-temporal lag coefficient θ



The green line represents the quantile estimates, the gray area represents the confidence intervals of quantile estimates and the dotted lines represent mean estimates (linear regression) and its confidence interval.

The estimates of all other explanatory factors show the expected sign and are significant in at least one of the models, except for lot size. In Figure 5.5, quantile estimates for this variable demonstrate that lot size has a positive coefficient in the lowest and the medium high price segment, which is rapidly turned into a negative effect in the highest price ranges. Perhaps the result in upper quantiles can be explained by individuals buying smaller parcels of land for residential purposes. Since the price of developable land is on average 50 times higher than the price of farmland in Belgium, their willingness to pay will be considerably higher. Another explanation can be related to horse farmers, who are not particularly interested in large pieces of farmland neither. Moreover, it can be observed that the variability in the upper quantiles is much higher in comparison with other quantiles. This might result from the divergence in farmland use after the sale. In lower price segments farmers might be willing to pay more for sizeable parcels to benefit from the economies of scale provided by these parcels.

Figure 5.5: Quantile estimates of lot size



Judging from the effect of residential zoning types and the address density in the vicinity of the observations, the redevelopment potential of agricultural parcels is an important price determinant in the Campine region. While land in densely populated areas is generally more apt for redevelopment in the long run, parcels located in residential zoning are more likely to be redeveloped in the short term. However, there does not seem to be a substantial difference between the effects, both of them clearly have a powerful inflating effect on farmland prices. Furthermore, all regression models show a significantly positive effect from the proximity of the Dutch border on farmland values. This might be an indication that Dutch farm holders are crossing the border in search for opportunities to buy relatively cheap agricultural land in Belgium.

As already discussed by Ciaian et al. (2012), high administrative costs associated with buying a piece of (farm)land in Flanders and Belgium have resulted in the emergence of a grey market. Registration taxes in Flanders amount to 10-12.5% of land value. This effect is clearly noticeable in the regression results which show that public sales transactions deal with a price markup of more than 20% in comparison with private sales. Furthermore, the time trend is significant at the 1% level in all models except for model 2. Seeing that prices have been corrected for inflation using the CPI, this illustrates that agricultural land prices have increased with above-inflation percentages from 2004 to 2011. This result might be partially explained by the increasing pressure on farmland from other zoning types such as residential and nature zoning or by an increased demand for farmland from Dutch farmers. Seeing that this will lead to an increasing scarcity in the farmland market, higher land prices would be an economically reasonable consequence.

5.6 Policy analysis

5.6.1 Soil contamination

Belgium is a federal country which implies that governmental responsibilities are spread over the central government and its communities and regions. In Flanders, the Flemish Region manages numerous policy domains including environmental policy. The Flemish government appeals to a number of 'internal independent agencies' that work out and implement the policies set out by the competent ministers. The Public Waste Agency of Flanders (Openbare Vlaamse Afvalstoffen Maatschappij, OVAM) is such an agency which is responsible for waste management and soil remediation in Flanders. With respect to sustainable soil management OVAM aims to prevent soil pollution and realize soil remediation, thereby creating a healthier environment and hence a better quality of life. This is embedded in the government's task to create policies that minimize health risks and ensure and improve public health.

Flemish policy with regard to soil remediation was taken in effect by the Soil Remediation Decree of 22nd February 1995, after which it has been updated by the Soil Decree of 27th October 2006. The implementation resolutions are registered in the Flemish regulation concerning soil remediation (Flemish Government, 2007). Because the Soil Remediation Decree was a great reform from previous soil regulations, a distinction is made between contamination that occurred before and after the day this decree came into effect, namely 29th October 1995. Contamination is considered to be 'recent' if it occurred after this date and 'historic' if it happened before this date. If the contamination occurred before as well as after this date or it is not particularly clear when it was originated, the contamination is termed as 'mixed'.

This regulation also established at what level of contaminants unacceptable health risks arise. For each contaminant limiting values were determined in function of different land uses and soil characteristics. The Flemish Soil Decree discerns three soil standards: background value, target value and threshold value (Table 1.2). Background values aim to protect clean soils from disturbance and contamination by providing the contaminant concentrations in normal, clean circumstances. Target values should at least be reached when a remediation project is set up, while threshold values cannot be exceeded because this would entail excessive health risks.

When confronted with recent contamination, the Soil Decree determines that soil remediation is obligatory when threshold values are exceeded in a certain typology. In case of historic soil contamination, a remediation is only necessary when serious health hazards are concerned. In assessing whether contamination is critical or not, contaminant levels, distribution risks, soil characteristics, land use and exposure risks are taken into account. Soil pollution is considered to cause serious health risks if there is a possibility of direct contact between contamination and humans, plants or animals and this contact is very likely to cause damage to these organisms. If the soil pollution has a negative impact on water collection, it will be considered as a serious threat as well.

5.6.2 Land transactions

Another goal of the Soil Decree is to protect (potential) buyers of land from unknowingly purchasing contaminated land. By transferring land from one landowner to another, the liabilities and the possible associated costs are transferred as well. Here, a transfer of land is assumed to imply a transaction involving a landowner selling his parcel(s) to another individual or organization for a price which both parties agree to. This means that rental agreements or inheritances are not included in this definition of land transfers. In the land transaction procedure we will focus on the way information on soil contamination is included in the procedure as to find out how buyers are notified on this subject.

However, before this procedure can be explained, there is an important distinction that is made by OVAM with regard to soils. If a soil accommodates or used to accommodate hazardous activities which are registered in the list of troublesome facilities[36], the soil will be classified as a 'hazardous soil' (Flemish Government, 2012). This list contains an elaborate amount of factories and establishments using potentially harmful substances in their processes. Examples of installations included in this list range from waste disposal sites to electricity producers and pharmaceutical companies. However, only the pieces of land on which the 'troublesome facilities' are located are considered to be hazardous soils. Hence, a classification as a non-hazardous soil does not entail that the soil cannot be contaminated as well. By means of undertow the contamination can still be transferred from the troublesome facility to other locations.

[36] The complete list can be found as Attachment 1 in the Flemish regulation concerning environmental permits.

The Soil Decree has determined that all land transfers – whether it concerns hazardous soils or not – must include a soil certificate among the transaction documents. A soil certificate is an official document that contains the information on the parcel that is available in OVAM's Ground Information Register (GIR). Soil certificates are designed by the government to inform land purchasers on the soil quality and include information on cadastral data, the identity of the landowner and the land user and in case the parcel is polluted also on the seriousness of the soil contamination and possible usage limitations and precautionary measures. However, in practice the latter characteristics on the soil certificates are only filled in if OVAM has information on soil surveys that have been conducted on that parcel. Otherwise, this information will be left blank.

However, the number of cases in which a soil investigation is obliged by the Soil Decree, is rather limited. An exploratory soil survey is required when hazardous soils are transferred, when the owner of hazardous soils is expropriated or goes bankrupt or when hazardous facilities with an increased risk of soil contamination are terminated. For some hazardous facilities a periodical soil investigation is necessary as well. In other cases or in non-hazardous soils, a soil survey is only performed on a voluntary basis.

There is an established procedure on how soil surveys should be carried out. First, an exploratory soil survey will find out whether there is contamination present in the soil. A parcel is recorded in the GIR at the time a contaminant exceeds 80% of its lowest threshold value (i.e. nature or agricultural land use). If a recent contamination is concerned and the survey indicates there is a significant probability threshold values are or will be exceeded, a descriptive soil survey is needed to find out more about the location and the risks of the contamination. In case of historic contamination there have to be indications of serious health risks before a descriptive soil survey is required.

The descriptive soil survey provides more details on the seriousness of the contamination and determines whether soil remediation is necessary according to the decision rules for a soil remediation. In the next step a soil remediation project is drawn up to elaborate how and using what technique the contamination should optimally be remediated. Moreover, the remediation costs are assessed to forecast the financial consequences of the project for the liable individual(s) or organization(s). In last instance the remediation works are started and a follow up plan is drafted.

5.7 Conclusion and discussion

As a result of the metallurgic industry's former calamitous production processes, large parts of the Campine region are disturbed by increased soil pollution levels. As shown in chapter 2, the affected area also includes a substantial amount of agricultural parcels. The contamination of farmland brings about a particular set of unfavorable consequences. Primarily, the farmer's operations will be affected by soil contamination, but society as a whole might have to deal with some adverse effects as well. It has been established in chapter 1 that certain crops are capable of absorbing contaminants in their systems even at very low soil concentrations. This way, farmers might risk growing crops that are exceeding food thresholds, thereby jeopardizing their own income as well as public food safety. Moreover, farmers might face land use restrictions being imposed to these parcels or in the worst case even soil remediation. Although the risks of contaminated farmland are considerable, the presence of soil contaminants has – to the best of our knowledge – never been incorporated in empirical hedonic analyses of agricultural land before. This chapter aimed to fill this research gap by applying the hedonic methodology to the farmland market in the Campine region.

Classic linear regression presents conditional mean estimates for all parameters in the model. In order to check the robustness of the results, these estimates were supplemented by conditional median estimates, which are provided using quantile regression (QR) techniques. Since QR techniques are capable of supplying estimates over the entire distribution of the response variable, it allows us to get a more comprehensive view of the relationship between variables in comparison with linear regressions. All farmland sales transactions (n = 599) that occurred between 2004 and 2011 in 14 municipalities which are located relatively close to the pollution sources were incorporated in the hedonic analysis. In a literature review, an overview was created of the most important factors affecting farmland price determination. Consequently, four groups of variables are included in the model, i.e. the environmental risk variable, agricultural variables, location variables and transaction variables.

In hedonic regression models it was found that environmental risk variables were insignificant determinants for farmland prices. Apparently, soil contamination is not a critical factor in the price setting of agricultural land in the region. There can be multiple explanations for this result. For example, farmers that do not intend to use the acquired land for crops with a considerable risk of exceeding food thresholds, but rather as pasture or for manure disposal for example, might not take the soil contamination into account. Dairy farms – the most important farming type in the Campine region – are an example of such a farming type. These farmers predominantly need land for pastures and corn as a feed stock for their animals. Since both of these land uses only have a limited capacity for taking up heavy metals, there is little risk that milk thresholds for heavy metals will be exceeded.

Although all observations are labeled as agricultural land in the Cadastre's administration, some parcels were located in different zoning types at the time of sale. This might imply that these pieces of farmland have been bought by individuals with non-agricultural interests such as residential or nature purposes for example. Possibly, the presence of soil contamination might be less of an issue in these cases. Alternatively, in explaining the highly significant result for the time trend in the analysis, it has been suggested that supply scarcity in the Flemish farmland market might have been a cause for rapidly increasing farmland prices. Moreover, farmers who intend to expand their agricultural activities need extra land to comply with strict fertilizing restrictions in Belgium. Hence, these regulations in combination with an increasing farmland scarcity might force farmers to purchase pieces of farmland with inferior or inappropriate land attributes such as soil pollution.

A last explanation can be related to the policy analysis that was performed after the hedonic analysis. Possibly, farmland buyers might have been unaware of the exact heavy metal concentrations present in the acquired parcels and its resulting land use restrictions. An analysis of Flemish land regulations pointed out that a soil investigation is not necessary when a parcel with potential historic contamination is transferred from one owner to another. Since the Campine region fits the description of historic soil contamination, the heavy metal concentrations are not reported in the transaction documents if the OVAM did not have any information with regard to these elements in its GIR or if the buyer did not explicitly asked for it. In case the soil certificates were lacking the necessary information concerning the polluting elements and the resulting land use restrictions, farmland buyers would have been unable to take their presence into account in price setting.

Although the importance of soil contaminants to farmers differs between types of farming, it is the government's responsibility to create policies that protect potential land buyers from buying parcels having undesirable characteristics of which they were unaware. One way of achieving this goal consists of informing land buyers about the condition the soil is in and the potential risks and land use restrictions they have to deal with when facing soil contamination. Whether the contamination is historic or new is not particularly relevant in this setting, the consequences for buyers remain the same. In case agricultural land is concerned, disregarding potential risk factors in soils might lead farmers to use land inappropriately, thereby possibly jeopardizing food safety when their products are considered to be harmful.

In the next chapter, we will try to further unravel the economic and motivational framework underlying farmland price determination in the Campine region, especially with respect to soil contamination. The hedonic price analysis applied in this chapter deals with some restrictions, both from a conceptual as a case-specific perspective. For instance, the hedonic methodology assumes that a perfectly competitive land market is in place and that buyers are completely informed on all product characteristics. However, the results for some explanatory factors pointed out that the farmland market in the area deals with some market imperfections, while the policy analysis showed that it is not particularly clear whether farmers are made entirely aware of soil pollutants and its consequences in sales transactions. Furthermore, a number of factors that possibly affected farmland values were unobservable in the hedonic model. Privacy regulations in Belgium prevented us from taking into account specific characteristics of the individuals involved in the sales transactions. For example, did the buyer purchase the parcel for agricultural purposes? Did he have any personal or other relationship with the seller? Was the parcel extra valuable to the buyer because of specific farm level characteristics or was the seller forced to sell his land? In this setting, a stated preference study might be a helpful complementary tool in order to circumvent some of these issues and better understand the factors specifically impacting the farmer’s land purchase decisions. With regard to the soil contamination issue, a stated preference design allows to explicitly inform the respondents on the topic and its consequences and confirm or oppose the hypotheses put forward in this chapter.

6 Investigating the farmer's land purchase decision in case of soil contamination: A discrete choice experiment[37]

6.1 Introduction

In estimating the benefits of cleaning up hazardous waste and other contaminated sites, the economic literature has focused on property values and complementary taxes that can be regained by remediation (Sigman and Stafford, 2011). Typically, the hedonic pricing model is applied to quantify the effect of contamination on real estate prices and to get an insight into the welfare forgone due to the contamination (Kiel and Williams, 2007). Most of these studies have related the proximity to undesirable facilities and contamination to real estate values in order to determine the depreciating impact on house values (de Vor and de Groot, 2011; Farber, 1998). Studies applying distance measures as a risk dimension implicitly assume that this agent provides all the necessary information about potential risks associated with the externality to individuals interested in buying real estate. The empirical literature on this topic has particularly focused on contaminated (Superfund) sites in the United States (Greenstone and Gallagher, 2008).

[37] Parts of this chapter have been presented at an international PhD workshop and a summer school:

- 5th European Association of Agricultural Economics (EAAE) PhD Workshop: Schreurs, E., Lizin, S., Van Passel, S. and Thewys, T. (2013). The impact of soil contamination on farmland values: Revealed preference versus stated preference. EAAE PhD Workshop, Leuven (Belgium), 29-31 May 2013.
- Summer School on Choice Experiments in Agricultural and Food Economics: Schreurs, E., Lizin, S., Van Passel, S. and Thewys, T. (2013). Analyzing the impact of soil contamination on farmland values. Summer School on Choice Experiments in Agricultural and Food Economics, Heverlee (Belgium), 3 July 2013.

In some cases, actual contaminant concentrations have been applied as a measure for risk as well (Clauw, 2007; Leggett and Bockstael, 2000), However, objective risk measures, as well as expert judgments on health risks, can sometimes differ greatly from perceived risks, because individuals tend to derive their risk perception from more than merely scientific risk assessments (McClelland et al., 1990). Other factors such as a high prior risk perception and an individual's personal assessment on how likely the contamination might impact their health status contribute substantially to perceived risk. Therefore, perceived risk is often a more determinant factor in real estate appraisal than objective risk assessments (Kunreuther and Slovic, 1996).

In case the perceived risk is higher than the objective risk, this is to a large extent attributable to the uncertainty associated with the (health) consequences of the contamination. In this setting, McCluskey and Rausser (2001) have shown that changes in information regarding contamination can have an impact on property values. The way in which media coverage or announcements by responsible government agencies affect the real estate market depends on whether the information decreases or increases the risk perceived by individuals (Gayer et al., 2000). Also when contamination is removed from soils, this property value effect is noticed. If the public is insufficiently informed on the risk reductions after a cleanup has taken place, there might still be a price discount in the areas that have suffered from the contamination. This is referred to as a stigmatization effect (Messer et al., 2006).

In order to provide adequate and comparative information to their citizens, governments often impose information disclosure requirements on goods and services of heterogeneous quality (Fiva and Kirkeboen, 2011). With regard to the problem at hand, information on soil quality is required in most developed countries to prevent individuals from buying properties holding contamination levels of which they were unaware. In this light, it has been established in the policy analysis of chapter 5 that the Flemish Government introduced soil certificates as an obligatory document in real estate transactions. Since these certificates only include information on soil quality and possible land use restrictions in rare cases, it was suggested that the lack of effect from soil contamination on farmland values in the hedonic pricing analysis might have been due to unawareness among farmland buyers.

However, it is difficult to find out a posteriori to what extent farmland buyers were informed on the soil condition at the time of sale, because it is nearly impossible to access the identity of individuals involved in real estate transactions in Belgium. In this setting, it has been put forward that stated preference (SP) methods can facilitate the estimation of parameters that are difficult or impossible to assess using revealed preference (RP) models alone (Holmes and Adamowicz, 2003). Therefore, a SP analysis might be a helpful tool in analyzing the importance of soil contamination in land purchase decisions. SP techniques allow to independently create a hypothetical farmland market and manipulate the information concerning soil quality and contamination that is provided to the respondents. Since respondents can be explicitly informed on the soil condition in this design, the information issue that was at play in the hedonic model can be circumvented.

The objective of this chapter is to investigate the role of soil contaminants in land purchase decisions by means of stated preference techniques. Therefore, a discrete choice experiment is set up among farmers in the Campine region, because they are expected to be the principal buyers of farmland. Furthermore, this research also aims to mark the differences between the results from the SP and RP analysis with regard to soil contamination and provide information on additional characteristics – such as soil productivity and locational factors – that are not incorporated in the hedonic analysis. The chapter is structured as follows. Section 6.2 will focus on stated preference techniques and particularly discrete choice experiments as a methodology for the valuation of environmental (dis)amenities. In section 6.3, we will elaborate on the design of choice experiments and the survey that was used for data collection. The first part of section 6.4 will present descriptive statistics with regard to the respondent sample, while the second part will reveal the principal results from the econometric analysis. In the last section, the results will be interpreted and some concluding remarks will be given.

6.2 Methodology

While RP techniques aim to value nonmarket goods or services such as environmental (dis)amenities implicitly by considering a complementary or substitution good that is traded in the market, SP techniques value the environmental (dis)amenity by explicitly asking individuals to express their preferences with regard to the good. The most compelling argument why many economists tend to favor RP over SP methods is the actual behavioral response – i.e. expenditures – associated with real market data, and correspondingly the hypothetical bias in SP data. However, SP studies have its merits as well. For example, if the research involves a product or an environmental state that is currently not in place yet, market data are unavailable and the usage of RP methods is not possible. In this case, SP methods can indicate the public's preferences and willingness to pay with regard to the product. Moreover, if it is difficult to estimate parameter coefficients using RP models alone or it is simply not clear whether the variable of interest has correctly been taken into account in market transactions – as is the case in this research – SP methods can provide a useful complementary tool to check whether the expressed preferences correspond to the behavioral responses.

In this chapter, discrete choice experiments (DCEs) are adopted as a SP methodology. DCEs aim to identify the individual's indirect utility function associated with attributes of goods or services by examining people's tradeoffs when making choice decisions (Garrod and Willis, 1999). Therefore, multiple alternatives – described by several product characteristics or attributes with varying attribute levels – are presented to respondents in choice sets. The respondents are then asked to pick one single alternative from each choice set, thereby revealing his/her preference for certain attributes or attribute levels. Since the number of profile combinations increases exponentially with the amount of attributes and attribute levels, presenting all choice sets to respondents quickly becomes a too extensive task. The specification of a fractional factorial experimental design allows for a reduced number of choice sets to decrease the cognitive burden faced by respondents (Holmes and Adamowicz, 2003). Subsequently, the choices can be econometrically analyzed in order to estimate attribute coefficients, which will reflect their relative importance.

The microeconomic theory underlying DCEs is based on the 'hedonic' notion that utility or value is derived from attributes of a particular good or situation, which was put on a firm theoretical basis by Lancaster (1966). His theory of consumer demand provides the basic conceptual structure for DCEs in an economic setting (Holmes and Adamowicz, 2003). Based on the conceptual foundation of random utility laid out by Thurstone (1927), McFadden (1974) expanded on the DCE framework and developed an econometric model that combined hedonic analysis of alternatives and random utility maximization. This model is referred to as the multinomial logit (MNL) model, which is considered to be the base model for DCEs (Hensher and Greene, 2003). In short, the derivation from random utility theory (RUT) to the general expression for the MNL model is as follows:

$$U_{ij} = V_{ij} + e_{ij} \qquad [21]$$

$$P_{ij} = P\{V_{ij} + e_{ij} > V_{ik} + e_{ik};\ \forall k \in t\} \qquad [22]$$

$$P_{ij} = \frac{exp(\mu Vij)}{\sum_k^K exp(\mu Vik)} \qquad [23]$$

$$V_{ij} = \sum_q^Q \beta_{jq} * X_{jq} \qquad [24]$$

RUT assumes that an individual i's total latent utility U_{ij} associated with alternative j consists of a deterministic part V_{ij} and a stochastic part e_{ij} (equation [21]). Hence, the probability P_{ij} that individual i prefers alternative j over all other alternatives k in choice set t can also be expressed as the probability that the total latent utility of person i for alternative j exceeds that of all other alternatives k in choice set t. Estimation of equation [22] requires the assumption of independently and identically distributed (IID) error terms. This allows for the convenient closed-form equation of the MNL model (equation [23]). Here, μ is a scale parameter which causes different DCEs not to be directly comparable. Within a single study it is most often assumed to be equal to one (Ben-Akiva and Lerman, 1985). The simplicity of the closed-form comes at a cost, given that the MNL model translates the IID assumption into substitution patterns that are restricted by independence of irrelevant alternatives (IIA) (Tesfaye and Brouwer, 2012). IIA entails that the ratio of the choice probabilities should be independent of the presence or absence of any other alternative in a choice set (Hensher et al., 2005). This means that it is assumed that the choice for one alternative in a choice set is not influenced by other available alternatives that are not considered in the choice set (Garrod and Willis, 1999).

In its simplest form, V_{ij} is assumed to be a linear, additive function with a vector of all attribute levels q of alternative j and their respective attribute parameter weights β_{jq} (equation [24]). V_{ij} transforms the multidimensional attribute vector into a one-dimensional utility measure (Louviere et al., 2000). Consequently, the higher the attribute (level) coefficient, the higher the utility and the higher the probability that an alternative will be chosen. Note that β_{jq} is not indexed for the respondent i, which implies that a homogeneous market is assumed when using the MNL model. For such a function of V_{ij}, it can be shown that that the marginal willingness to pay for an attribute equals the negative of the ratio between that attribute's coefficient and the coefficient for a payment vehicle, typically the price (Bergmann et al., 2006).

In some cases it might be recommended to estimate an intercept – referred to as an alternative specific constant (ASC) – as well in DCEs. This constant is specific to an alternative and represents the unobserved factors influencing choice decisions as they pertain to the particular alternative for which it was estimated (Hensher et al., 2005). ASCs need to be added to labeled alternatives, i.e. alternatives having a title that conveys meaning such as a brand or a classification label (Louviere et al., 2000). With regard to DCEs using generic or unlabeled alternatives, i.e. alternatives having general titles such as 'Alternative' and 'Option', there is some controversy in the literature whether ASCs ought to be included. Although Hensher et al. (2005) asserts ASCs are not required in generic designs claiming it would violate the meaning of 'unlabeled', the exclusion of ASCs would lead other model parameters to capture this effect, resulting in biased attribute parameter estimates (Hoyos, 2010). According to Mogas et al. (2006), ASCs can be interpreted as the utility premium over choosing the status quo option. Therefore, if a DCE design contains a status quo or opt-out option, an ASC is necessary to model the utility of this alternative (Holmes and Adamowicz, 2003).

6.3 Data

Generally, setting up a discrete choice experiment requires following seven steps (Garrod and Willis, 1999; Louviere et al., 2000). These steps are outlined in Table 6.1. Steps 1 to 5 are handled in this section, while steps 6 and 7 are discussed in section 6.4.

Table 6.1: Steps in a discrete choice study

Step	Action
1	Characterize the decision problem
2	Identify key attributes and attribute levels
3	Develop an experimental design
4	Design questionnaire survey
5	Pre-test and undertake survey
6	Estimate model
7	Interpret results

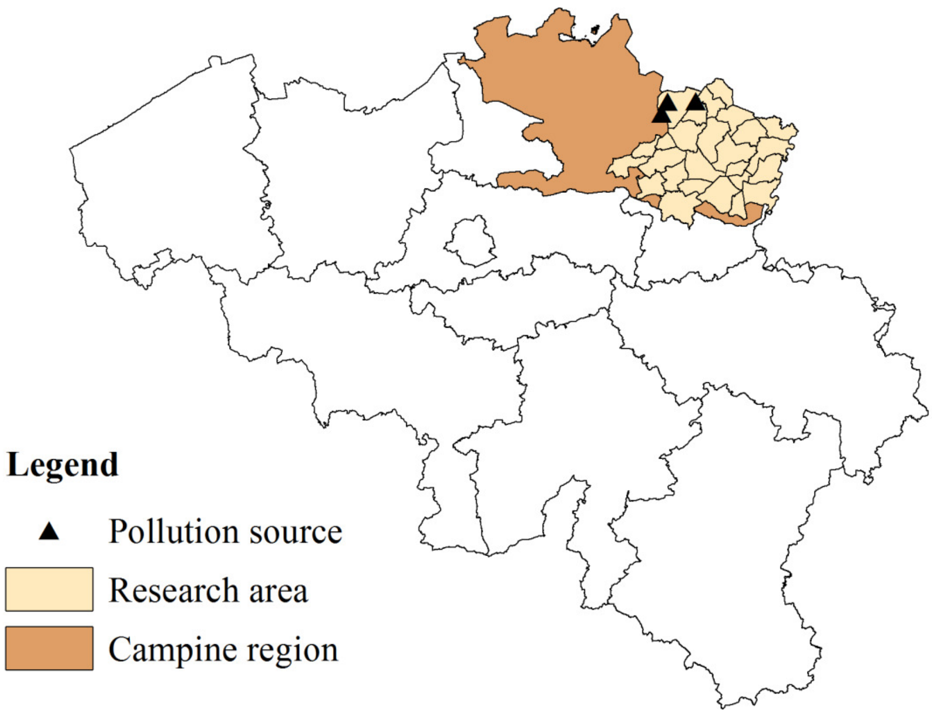
This study aims to estimate the relative importance of soil contamination in the farmer's land purchase decisions and to mark the differences between the hedonic pricing analysis and a stated preference analysis. Therefore, a farmland purchase decision is simulated in the choice experiments. Per choice set, three hypothetical parcels of farmland are offered to farmers, each with a specific price and specific attribute levels. The farmers are then asked to select the alternative that is expected to provide the maximum amount of utility to their operations. Since there are no particular brands or labels associated with farmland, a generic design is created. An opt-out alternative is included in each choice set in order to mimic actual market behavior. This will increase the farmer's degree of familiarity with the choice setting, enhance the theoretical validity of the welfare estimates and improve the statistical efficiency of the estimated choice (Kontoleon and Yabe, 2003).

In light of the different steps required in a DCE, Boerenbond – the largest farmer association in Flanders – was contacted to cooperate in this study. Besides serving as a sounding board and expert panel for selecting the relevant farmland attributes, Boerenbond was also able to provide contact details for a considerable amount of farmers in the study area[38]. Moreover, the involvement of a farmer association might increase the support for the

[38] More than 50% of active farmers in the province of Limburg are member of Boerenbond.

survey among farmers. Since Boerenbond is organized in provincial sections and the principal part of the research area is located in the province of Limburg, the cooperation was set up specifically with Boerenbond Limburg. Contact information was provided for Boerenbond members in all municipalities that were located for at least 50% (of surface area) in the Campine region. More specifically, the study area included the municipalities As, Beringen, Bocholt, Bree, Dilsen-Stokkem, Genk, Ham, Hamont-Achel, Hasselt, Hechtel-Eksel, Heusden-Zolder, Houthalen-Helchteren, Kinrooi, Leopoldsburg, Lommel, Lummen, Maaseik, Maasmechelen, Meeuwen-Gruitrode, Neerpelt, Opglabbeek, Overpelt, Peer, Tessenderlo, Zonhoven and Zutendaal. In Figure 6.1 the municipalities in the study area are displayed in a map of Belgium.

Figure 6.1: Research area discrete choice experiments



The next step in implementing a DCE involves identifying the relevant attributes and levels. Since no hard and fast rules are established to verify whether all relevant attributes are incorporated, it has been suggested to use focus group discussions in this stage of the design (Holmes and Adamowicz, 2003). Therefore, the expert opinions of Boerenbond as well as a focus group of agricultural and environmental economists at Hasselt University were consulted. This led to the selection of five parameters out of an extensive list of factors that could potentially affect the farmer's land purchase decision. These parameters are price, lot size, soil productivity, location and land use restrictions. The location parameter was subdivided into two independent attributes, one that indicates the driving time by tractor from their home to the parcel and one that indicates how far the parcel is located from other farmland that is cultivated by the farmer. So eventually, we ended up with six attributes in the DCE. Both focus groups acknowledged that in general farmers considered these factors as principal determinants in deriving farmland utility.

The land use restrictions attribute incorporated the soil contamination variable as one of its attribute levels. In this attribute level, respondents were explicitly informed that soil thresholds for one contaminant were exceeded and that, as a result, the land user is not allowed to cultivate any vegetables or arable crops (cereal, wheat, potatoes,…) on this land. There were two main reasons for integrating soil contamination in this way. First of all, farmers in the Campine region are generally faced with two kinds of situations. In the first situation, the soil threshold level for Cd is not exceeded and the parcel is considered to be uncontaminated. Hence, the farmer is (theoretically) allowed to cultivate all the crops he prefers. In the other situation, the soil Cd threshold is exceeded and the land user is confronted with some crop restrictions. However, since the most polluted parcels in the area still only slightly exceed this threshold[39], barely any extra restrictions are applied the higher the pollution levels are. Therefore, we chose to simplify and generalize the situation into a (hypothetical) ambivalent status, i.e. an uncontaminated or a contaminated parcel.

[39] In the hedonic pricing analysis, the most polluted farmland parcel had a soil Cd concentration of 5.75 ppm.

Secondly, assigning one attribute to soil contamination with complementary contamination levels would cause two problems. The respondents might feel unfamiliar with soil contamination levels because it pertains a too high level of complexity. If the farmers are incapable of understanding the questions being asked, then they will not be able to reveal their true preferences. In this light, we chose to keep our explanation in the attribute level simple and as close as possible to an agricultural context. Moreover, the introduction of a soil contamination attribute might enforce the respondents to behave strategically and refrain them from answering truthfully to the choice experiments. Therefore, a soil contamination attribute might increase the probability of obtaining biased results.

The other land use restrictions were integrated as part of our deal with Boerenbond to investigate some topics they were interested in, in exchange for their expertise on the topic and contact details from their members. The disadvantage in adding these land use restrictions in the same attribute as soil contamination is that attribute levels are not mutually exclusive anymore. This implies that the land use restrictions are assumed to be applied separately (and thus not as a combination of restrictions) and the coefficients are conditional on the fact that they are compared to exactly three other land use restriction attribute levels. Obviously, this restrains the interpretations that can be made from the results for this attribute and prevents its generalization on a higher level.

Table 6.2: Farmland attributes and levels

Attribute	Level 1	Level 2	Level 3	Level 4
Price (€ ha^{-1})	15,000	25,000	35,000	45,000
Lot size (ha)	0.5	1.5	2.5	3.5
Soil productivity	Low	Rather low	Rather high	High
Driving time to home (min)	5	10	15	20
Distance to other farmland (m)	0	750	1500	2250
Land use restrictions	None	No arable crops and vegetables due to soil contamination	25% less usage of fertilizers	Permanent pasture

Each attribute was assigned four levels, which aimed to reflect the farmland market in the Campine region as closely as possible. For three continuous variables – i.e. price, lot size and driving time to home – level allocation was based on the distribution of these variables in the hedonic analysis (Adamowicz et al., 1994). Therefore, Table 6.3 indicates at which percentile of the hedonic dataset the lower and upper range of the attributes are situated. The upper range for all attributes is situated at approximately the 95th percentile. For the price variable, 15% of the observations deals with sales prices lower than 15,000 €/ha. However, this is partly due to sales transactions in nature area and partly due to the extensive time range 2004-2011 in the hedonic analysis. At the time the survey has been administered (December 2012 - February 2013), farmland prices below 15,000 €/ha were considered as rather exceptional in the area. With regard to lot size, the hedonic dataset included a vast amount of very small parcels, some of which might have been intended for residential purposes. This explains why the lower range of 0.5 ha is situated at the 40th percentile.

Table 6.3: Distribution continuous variables hedonic analysis

	Price		**Lot size**		**Distance to home**	
	Level (€ ha^{-1})	Percentile	Level (ha)	Percentile	Level (km)	Percentile
Lower range	15,000	15th	0.5	40th	/	/
Upper range	45,000	92th	3.5	94th	7[40]	95th

Generally, the range in the attribute levels covers the largest part of hedonic data distribution and excludes outlier values. Since no information was available on the distance to other farmland in the sales data, this attribute was assigned levels on the basis of expert opinions in both focus groups. By spreading the attribute levels using the same intervals, it is possible to estimate intermittent levels of the continuous variables in the econometric analysis. Otherwise, they have to be analyzed as dummy variables, as is the case in non-numeric attributes such as soil productivity and land use restrictions. In the latter attribute, different land use restrictions are compared to the base level in which there are no land use restrictions. The soil contamination case was introduced here as one of the attribute levels in this attribute, together with a fertilizing restriction of 25% and a requirement to permanently use the parcel as a pasture.

The third step in setting up a DCE involves developing an experimental design. Given that 6 attributes are included in the design, each with 4 attribute levels, 4096 possible profiles exist. Consequently, a fractional factorial design is created to reduce the amount of choice sets presented to the respondents. In this study a main effects, D-efficient design for a MNL model was created using SAS. The prevailing argument for selecting a D-efficient design over an orthogonal design is the minimization of standard errors on parameter estimates, which allows for smaller sample sizes (Bliemer and Rose, 2011). This resulted in a design consisting of 16 choice sets, which was blocked over two surveys in order to reduce respondent

[40] If it is assumed a tractor drives at a speed of 20 km/h on average, the tractor will have covered approximately 7 km in 20 minutes.

fatigue. The choice sets in each block were randomized five times to counter order effect bias (Day et al., 2012).

Subsequently, both blocks were introduced in the survey, which was designed to roughly fit the guidelines provided by Bateman et al. (2002): (1) Survey purpose, (2) Farm-level questions, (3) Attitudinal/motivational questions, (4) Choice sets and (5) Socio-economic questions. The second section in the survey contained questions about the agricultural activities the farmer was currently involved and the farm's land allocation, while the third section included statements to assess their risk attitude and environmental awareness. The survey was pre-tested in both focus groups as well as in a subsample of 6 farmers in the area.

The final decision to be made concerned the distribution method. There are only two modes of administration suitable for DCEs, i.e. in-person interviews or (electronic) mail questionnaires. Telephone surveys are excluded because a visual inspection of the choice sets by the respondents is necessary in order to adequately compare the alternatives. In this study the in-person option was preferred because of two major arguments. Although in-person interviews are rather time consuming, this distribution method returns high quality data in which the amount of missing data is strongly reduced. Moreover, it enables the interviewers to provide the respondents with extra information in order to clarify the objective and the interpretation of certain questions. Secondly, given that mail questionnaires have the lowest response rates of all survey methods (Champ, 2003) and the amount of farmers in the study area is rather limited, this method might return a too small sample of respondents. Additionally, farmers are generally not that familiar with electronic mail.

The Boerenbond membership list was used as a sampling list for contacting respondents. This list was corrected by Boerenbond itself to exclude farmers that were classified as having a very limited amount of agricultural activities. The final sampling list contained 684 addresses and telephone numbers from farmers living in the study area. Respondents were selected by simple random sampling from the contact list. The distribution over the municipalities mentioned in Figure 6.1 can be found in Table 6.8 in Appendix 1. Farmers were first contacted by telephone to briefly explain the nature and the objectives of this research, after which they were asked if they were willing to participate in the study. If the respondent agreed to cooperate, an appropriate date and time was arranged for an in-person interview.

6.4 Results

6.4.1 Descriptive statistics

The survey was completed by 200 farmers in the Campine region. A rather high response rate of 67% was obtained, which can probably be explained by a combination of elements. The involvement of a sector organization increased the support for the survey and stimulated farmers to participate. The objective of economically quantifying the consequences of land use restrictions was also deemed as highly interesting by farmers. Moreover, in the telephone conversation prior to the interview, the respondents were told that the survey would take up only 15 minutes of their time. This might have encouraged farmers to cooperate as well, especially since the winter time is the least busy period of time for most farmers.

In Table 6.4 the socio-economic characteristics of the respondents are summarized. It can be observed that the sample almost exclusively contains male farm managers, corresponding to the general conception of the farming population. The table further indicates that predominantly middle-aged farmers are included in the sample. Combined with the fact that 98% of the sample are full time farmers, this points out that the sample mainly includes professional farmers for whom the agricultural activities are the most important source of income. These farmers are most likely to play an active role in the farmland market. Seeing that we are particularly interested in the farmers that have an impact on the farmland demand curve, the sample seems to correspond quite closely to the desired farming population.

Moreover, our sampling list (i.e. the membership list of Boerenbond) did include a number of individuals who were not professional farmers (anymore), but had some agricultural activities as a hobby or as a side activity. Since participants were selected by simple random sampling, these people have also been contacted by telephone in our data collection procedure. However, when the research objectives were briefly explained to these people, they mostly refused to participate in the study because they were not interested in purchasing farmland (anymore) and therefore could not properly contribute to the analysis. That is another reason why our sample predominantly includes professional farmers.

Table 6.4: Socio-economic characteristics of the DCE sample

Socio-economic characteristics		**Percentage**
Sex of farm manager	Male	95.5%
	Female	4.5%
Age of farm manager	>35	6.5%
	35-44	18%
	45-54	58%
	55-64	14%
	>65	3.5%
Employment	Full time	98%
	Part time	2%
Successor (age>50)	Yes	33%
	No	47%
	Not sure	20%

With respect to farming types, the sample primarily includes specialist farms (Table 6.5). A farm is labeled as being specialist if at least two thirds of the farm’s gross margin emanates from one agricultural activity. The sample particularly includes three types of farming, i.e. specialist dairy farms, specialist pig farms and mixed farms. However, the sample is clearly dominated by specialist dairy farmers.

Table 6.5: Types of farming in the DCE sample

Farming type	Percentage
Specialist farms	79,5%
Field crops	2%
Milk	56,5%
Pig	10%
Grazing livestock	4,5%
Vegetables	1,5%
Fruits	3%
Other	2%
Mixed farms	20,5%

6.4.2 Econometric analysis

At first, a basic MNL model was estimated without interaction effects in order to obtain a general insight into the results for the complete respondent sample. If the attribute levels of the quantitative attributes were not significantly different from each other, they were appended as a linear variable. This was the case for the attribute distance to other farmland. The attribute levels of the other quantitative attributes, i.e. lot size and driving time to home, have been introduced as dummy variables because non-linear effects were noticeable[41]. The attribute levels were coded using dummy coding with a lot size of 3.5 ha and a driving time to home of 20 minutes as base levels. For the qualitative attributes, i.e. soil productivity and land use restrictions, high productivity and no land use restrictions were used as base levels. An ASC for the opt-out option is included in the analysis. Following Holmes and Adamowicz (2003), each attribute level of the opt-out

[41] We used Wald tests to verify whether the null hypothesis of equal attribute level coefficients could be rejected.

alternative was handled using zeros. This approach normalizes the utility relative to the opt-out option.

Table 6.6: MNL model with main effects

	Coefficient	**St. err.**
Price	-0.071**	0.030
Lot size 0.5 ha	-0.478***	0.083
Lot size 1.5 ha	-0.078	0.083
Lot size 2.5 ha	-0.351***	0.083
Low productivity	-0.575***	0.081
Rather low productivity	-0.490***	0.086
Rather high productivity	-0.085	0.083
Distance to other farmland	-6.61E-5*	3.43E-5
Driving time to home 5 min	0.133*	0.077
Driving time to home 10 min	0.176**	0.079
Driving time to home 15 min	0.024	0.073
Soil contamination	-0.317***	0.088
Fertilizing restriction	-0.588***	0.109
Permanent pasture	-0.361***	0.099
ASC	-1.574***	0.188
N	6400	
Pseudo R^2	0.056	

*, **, *** represents significance at 10%, 5% and 1% level, respectively.

The main effects model in Table 6.6 shows that the presence of soil contamination and the resulting crop restrictions reduces farmland utility at the 1% level in comparison with the base level in which no land use restrictions are applied. The average farmer in the Campine region appears to prefer parcels of farmland that are not affected by soil contamination in case they are explicitly informed on the (land use) consequences. However, the coefficients of the other two land use restrictions are strongly negative as well, with values that are even lower than the coefficient of soil contamination. Indeed, this indicates that farmers are even less likely to select a parcel of farmland that deals with fertilizing restrictions or that is labeled as permanent pasture. Especially the coefficient of the fertilizing restriction attribute level is remarkable, since it is significantly lower than the other two coefficients[42]. Apparently, farmers prefer farmland that deals with crop restrictions over farmland without crop restrictions but with fertilizing restrictions. This reinforces the argument put forward in section 5.5.2 that farmers in the Campine region are particularly interested in purchasing land to comply with strict fertilizing conditions in Flanders.

All other attribute (level) coefficients exhibit the expected sign. The lot size attribute indicates that farmers are less attracted to smaller pieces of farmland. Especially small parcels of 0.5 ha are not appealing. Surprisingly though, the attribute level of 1.5 ha is not significantly different from the base level of 3.5 ha. In comparison with the high productivity base level, farmers do not expect to experience significantly less utility from a parcel that is slightly less productive, but is still labeled as rather high productive. Possibly, farmers particularly paid attention to the terms 'high' and 'low' in the productivity attribute and tried to avoid parcels that were marked as being low productive. The location of parcels is an important aspect in farmland purchase decisions as well. Farmers are more likely to select farmland that is located more closely to the farmer's home or to other parcels in the farmer's cultivation area. The negative coefficient for the ASC indicates that choosing the opt-out option provides significantly less utility to respondents in comparison with selecting one of the three farmland alternatives.

In a main effects MNL model, it is assumed that every respondent in the sample relies on the same preference structure for making choices. However, farmers differ from each other on a number of characteristics such as demographic and socio-economic factors, but also in their main agricultural activities and their motivation for purchasing farmland. Therefore, it seems to be unlikely that there is effectively preference homogeneity in the respondent sample. That is why we decided to also estimate a random parameter logit (RPL) model, which has been introduced

[42] $\chi^2(1) = 7.33$ (in comparison with the soil contamination coefficient)
$\chi^2(1) = 5.57$ (in comparison with the permanent pasture coefficient)

by Train (2003) and allows parameters weights to differ between respondents. This model can be estimated using an experimental design that was intended for a MNL model without substantial efficiency losses. The advantage of a RPL model is that it allows to capture unobserved heterogeneity by estimating whether the attributes' standard deviation is significantly different from zero. The RPL model was estimated with simulated maximum likelihood using 1000 Halton draws. Parameters are assumed to be normally distributed, except for the ASC and the price parameter which were considered to be fixed.

Table 6.7: Random parameter logit model

	Mean		**Standard deviation**	
	Coefficient	**St. err.**	**Coefficient**	**St. err.**
Price	-0.102***	0.035	/	
Lot size 0.5 ha	-0.510***	0.097	-0.101	0.168
Lot size 1.5 ha	-0.127	0.103	0.613***	0.142
Lot size 2.5 ha	-0.407***	0.100	0.467***	0.168
Low productivity	-0.615***	0.094	0.016	0.039
Rather low productivity	-0.586***	0.102	0.352*	0.187
Rather high productivity	-0.141	0.102	0.751***	0.133
Distance to other farmland	1.10E-4***	4.07E-5	3.60E-4***	5.47E-5
Driving time to home 5 min	0.170*	0.091	0.130	0.293
Driving time to home 10 min	0.177*	0.091	-0.086	0.241
Driving time to home 15 min	-0.034	0.091	0.009	0.047
Soil contamination	-0.479***	0.110	0.521***	0.138
Fertilizing restriction	-0.820***	0.133	0.764***	0.150
Permanent pasture	-0.535***	0.128	0.605***	0.146
ASC	-1.879***	0.205	/	
N	6400			
Pseudo R^2	0.071			

The results of the RPL model in Table 6.7 show that the mean effects are similar to the MNL model. The coefficients of some attributes shift a little bit, but the findings from the MNL model are generally intact. As for the heterogeneity in parameter estimates, the RPL model indicates that there are a number of attribute(s) (levels) that deal with standard deviations that are significantly different from zero. More specifically, the respondents seem to have divergent preferences with respect to the attributes lot size, distance to other farmland, land use restrictions and to a lesser extent the productivity attribute. In order to find out to what this heterogeneity is attributable, interaction effects with specific characteristics that are constant across the choice occasions for all individuals are introduced in the MNL model. This way, the heterogeneity that is observed in the RPL model can be specified, which should improve model fit and the reliability of parameter estimates. Moreover, interaction effects can provide some identification of attribute parameter differences in response to changes in individual factors.

Table 6.8: MNL model with main and interaction effects		
	Coefficient	**Std. Error**
Main effects		
Price	-0.298***	0.108
Lot size 0.5 ha	-0.412***	0.093
Lot size 1.5 ha	0.023	0.092
Lot size 2.5 ha	-0.252***	0.091
Low productivity	-0.585***	0.081
Rather low productivity	-0.501***	0.086
Rather high productivity	-0.090	0.083
Distance to other farmland	-1.13E-4***	4.01E-5
Driving time to home 5 min	0.123	0.078
Driving time to home 10 min	0.185**	0.080
Driving time to home 15 min	0.016	0.074
Soil contamination	-0.601***	0.139
Fertilizing restriction	-0.860***	0.156
Permanent pasture	-0.359***	0.099
ASC	-1.619***	0.187
Interaction effects		
Price - price estimation	7.00E-5**	3.02E-5
Lot size 0.5 ha - food producer	-0.392*	0.234
Lot size 1.5 ha - food producer	-0.683***	0.254
Lot size 2.5 ha - food producer	-0.665***	0.255
Distance to other farmland - expansive farms	1.78E-4**	8.29E-5
Soil contamination - specialist dairy farms	0.458***	0.148
Fertilizing restriction - agricultural eduction	0.444***	0.167
N	6400	
Pseudo R^2	0.069	

Understandably, interaction effects are found in attributes that showed significant heterogeneity levels in the RPL model (Table 6.8). Further, the results show that the overall model fit increases when the interaction terms are appended in comparison with the basic MNL model (Table 6.6). The coefficients of main effects have (somewhat) changed, but have remained roughly the same in terms of significance. With respect to the soil contamination attribute level, specialist dairy farms seemed to attach below-average importance to its presence in farmland. This might be explained by the fact that in case of soil contamination farmers are restrained from cultivating arable crops and vegetables, but not from using it as a pasture or to cultivate fodder crops, which are the two main land allocations for this type of farming. This result also implies that the principal farming type in the Campine region (Table 6.5) considers soil contamination to be less of an issue than the average farmer.

The positive effect for the interaction term 'price – price estimation' indicates that the importance of the price attribute depends on the farmer's assessment of average farmland prices in their municipality. In the survey, the respondents were asked to provide an estimate of the current per hectare price for a parcel of farmland in their municipality[43]. The higher farmland prices were assessed by the respondents, the less attention was paid to selecting the alternative with the lowest price in the choice sets. With regard to size of the parcel, the main effect was enhanced by including an interaction effects with the farming type that primarily used land for producing food crops (field crops, vegetables, fruit). These farmers were even more responsive to obtaining the largest amount of farmland area compared to the average farmer.

Next, there is a significant interaction effect between the attribute distance to other farmland and 'expansive farms'. Farmers were considered to be interested in expanding their activities if the amount of farmland they purchased in the last 5 years was situated in the upper quartile of the sample[44]. The interaction effect reveals that this category of farmers did not seem to care much whether newly acquired farmland is located close to other farmland in their possession. The last interaction term involves the

[43] The average farmland price estimate was 32,062.5 €/ha with a standard deviation of 9082.14 €/ha. The median estimate was 30,000 €/ha. A histogram of estimated prices can be found in Appendix 2.
[44] One of the questions in the survey asked how much farmland the respondent has purchased in the last 5 years.

effect of the farmer's education on fertilizing restrictions. If farmers received an agricultural education, fertilizing restrictions were considered to be less of a burden in comparison with the average farmer. Possibly, these farmers were better informed on techniques to manage these restrictions in their education.

6.4.3 WTP estimates

Although it was not the primary objective in this analysis, DCEs allow to determine the respondent's willingness to pay for the included attributes. The marginal WTP is the negative ratio of the attribute coefficient and the coefficient for the payment vehicle, i.e. the price attribute in this case (Bergmann et al., 2006). For simplicity, these WTP estimates are based on the basic MNL model, but the results for the RPL model are quite similar.

Table 6.9: Marginal willingness to pay estimates from MNL model

Attribute levels	**mWTP** (€/ha)
Lot size 0.5 ha	-67,019.10
Lot size 1.5 ha	-10,929.33
Lot size 2.5 ha	-49,205.24
Low productivity	-80,722.18
Rather low productivity	-68,782.42
Rather high productivity	-11,899.10
Distance to other farmland (per 1000m)	-9271.69
Driving time to home 5 min	18,611.34
Driving time to home 10 min	24,743.19
Driving time to home 15 min	3353.83
Soil contamination	-44,517.60
Fertilizing restriction	-82,524.96
Permanent pasture	-50,654.83

While it was mentioned before that the farmers in the DCE sample assessed farmland prices at 32,062.5 €/ha on average, the hedonic pricing analysis reveals that the actual average farmland sales price in 2011 (the last year for which data were available) was 30,625.71 €/ha[45]. The setup of such a revealed preference – stated preference study design conveniently allows to compare the results from the SP analysis with market-based values. That is why we are compelled to conclude that the results in Table 6.9 indicate that the WTP estimates for all attributes are clearly overestimated.

To pick out just one example, the application of fertilizing restrictions to a parcel of farmland would result in a negative WTP of 82,524.96 €/ha in comparison with a parcel of farmland without any land use restrictions. Given that the average price for a piece of farmland at the moment the survey has been administered was approximately 32,000 €/ha, this would imply that there is a negative WTP of 50,000 €/ha for farmland with fertilizing restrictions. However, the farmland sales data in the HPA contain a number of parcels that have to deal with fertilizing restrictions. It even includes farmland that is located in nature area, which implies that practically no fertilizers or herbicides can be used, compared to the 25% reduction simulated in the choice experiments. For all other attributes, the analysis is about the same, i.e. the implicit prices are way too high for them to be realistic in the current Campine region farmland market.

Of course, it is not the first time that marginal WTP values derived from SP analyses tend to turn out higher than expected and seem to be overestimated. The difference between WTP estimates from hypothetical (stated preference) experiments and real world experiments with actual financial consequences is referred to as hypothetical bias (Ozdemir et al., 2009). Especially in CV analyses the presence of hypothetical bias has been reported several times, using different elicitation methodologies, on the valuation of private as well as public goods and in settings all across the world. List and Gallet (2001) and Murphy et al. (2005) present two meta-analyses on this topic and find that SP analyses overstate WTP values 2 to 3 times in comparison with revealed WTP. Since economists are usually

[45] Note that the research area in the HPA is slightly different from the research area in the DCE. Nevertheless, this price merely serves as a reference point for the general farmland price level in the Campine region.

primarily interested in the values of nonmarket goods that can be derived from SP analyses, these issues have raised some questions about the external validity of the use of CV induced benefit measures for public policy (Diamond and Hausman, 1994).

Although hypothetical bias has been seriously debated in the CV literature, there have been few studies testing for hypothetical bias in a DCE format (Carlsson et al., 2005). Probably, this can be related to the relatively recent adoption of DCE in empirical applications in comparison with CV. Furthermore, it was also assumed that CE is less prone to hypothetical bias as a result of its closer resemblance to actual consumer purchasing decisions (Murphy et al., 2005). Early efforts to study hypothetical bias in DCE returned mixed results. Carlsson and Martinsson (2001) and Lusk and Schroeder (2004) found no evidence of significant differences in marginal WTP, but Lusk and Schroeder (2004) did report different total WTPs in a hypothetical and actual choice experiments. Carlsson et al (2005) and Hensher (2010) make notice of studies that reject the equality between marginal WTPs. Yet, these studies have not been published in international peer-review journals.

In the later literature, authors advanced on the topic of hypothetical bias in DCEs. Especially the study design was a subject of interest, since the previously mentioned studies that did not find differences in marginal WTP between hypothetical and actual CE were conducted in closely controlled field experiments (Carlsson et al., 2005). Johansson-Stenman and Svedsater (2008), for example, argued that a within-subject study design (such as used by Carlsson and Martinsson (2001)) is less likely to include a hypothetical bias because people strive for consistency in their behavior. This will urge them to act in a way that corresponds to prior hypothetical statements, but which is different from the way they would behave in an independent setting. Therefore, between-subject designs are considered to be more valuable for researching hypothetical bias in DCEs. The presence or absence of an opt-out option has also been suggested to make a difference in the results of DCEs (Hensher, 2010; Ready et al., 2010)[46].

There are other arguments why researchers started to question WTP estimates from DCEs. For example, in the health economic literature DCEs have been shown to provide even higher WTP values than (particularly open-ended) CV studies (Balistreri et al., 2001; Chestnut et al., 2012; Ryan and Watson, 2009). A recent experiment by Taylor et al. (2010) reported similar findings in an environmental economic setting, but suggests that more research on this topic is needed before more general conclusions on the validity of choice modeling approaches can be made. Nevertheless, the empirical literature seems to contradict the preliminary assumption that CE approaches are more capable of eliminating hypothetical bias than CV methodologies.

However, there is still much to know about the underlying causes of hypothetical bias and how to eliminate it from SP analyses (Mitani and Flores, 2009). Murphy et al. (2005) confirms that there are a number of factors affecting hypothetical bias, but states that there will probably be no single technique that will be the 'magic bullet' that eliminates hypothetical bias completely. That is why a flow of literature has arisen that tries to deals with this issue in SP analyses (Loomis, 2011).

One of the most cited mechanisms for reducing hypothetical bias is the use of cheap talk. Cheap talk informs respondents that previous SP analyses have reported the existence of an upward bias in their value estimates and reminds them that people are prone to overstating their actual WTP in a hypothetical setting. This strategy has been introduced by Cummings and Taylor (1999), who found that it is a successful way of reducing the hypothetical bias in a CV study of valuing environmental goods. Although the results of cheap talk are not uniformly positive (Blumenschein et al., 2008), cheap talk scripts have been successfully used in DCE settings as well (e.g.

[46] Carlsson and Martinsson (2001) did not include an opt-out option in their study design, which has been described as a serious design flaw by some (Harrison, 2006).

Ozdemir et al. (2009)). Tonsor and Shupp (2011) adds that including cheap talk might increase the reliability of the produced WTP estimates as well. However, at the same time the effectiveness of cheap talk has been shown to depend on the subset of consumer samples to which the DCEs have been presented. Especially respondents that are unfamiliar with the attribute being evaluated seem to be prone for overestimating their WTP (Champ et al., 1997). That is why Schlapfer and Fischhoff (2012) recommend that SP studies should only be conducted when the good and the context can be made familiar and meaningful to the respondents.

A second ex ante approach to reduce hypothetical bias suggests that the simulated market should be consequential for the respondents to be incentive compatible (Carson and Groves, 2007). That is, if the respondents perceive no gain or losses from answering a preference survey, they are not likely to always answer truthfully. This implies that the study design should urge respondents to reveal their true preferences by introducing a potential effect on their future utility. For example, respondents in a simulated market situation might be committed to actually paying one of the stated WTP values after the experiment, such as in Lusk and Schroeder (2004). This approach has been found to reduce the degree of hypothetical bias in a number of empirical applications (Landry and List, 2007; Mitani and Flores, 2009).

Among ex post methods to reduce hypothetical bias are certainty indexes that can serve to recalibrate the values from SP analyses. Champ et al. (1997) was one of the first to include respondent uncertainty in their analyses, which appeals to the intuition that as people’s certainty with regard to the answers given decreases, the choices become less deterministic (Beck et al., 2013). These studies ask the respondents to indicate how certain they are of their answer, usually on a 1 to 10 scale (Loomis, 2011). However, at which threshold point a respondent can be labeled as certain is chosen somewhat arbitrarily (Morrison and Brown, 2009). The limited evidence that is currently available suggests that this approach is capable of converging hypothetical WTP estimates to actual WTP (Ethier et al., 2000; Johannesson et al., 1999; Morrison and Brown, 2009). Beck et al. (2013) recently renewed the interest in certainty indexes for reducing hypothetical bias, but advocates that a more general theoretical framework is needed before the importance of this approach can be determined. This has been suggested before by Ready et al. (2010) as well.

6.5 Conclusion and discussion

The topic of interest in this chapter was the farmer's perspective on farmland purchase decisions, and particularly the role of soil contamination in this setting. The analysis aimed to reveal the importance farmers attached to the presence of soil contaminants in agricultural land as to get a better view on the motivational framework underlying farmland purchase decisions. Therefore, a stated preference (SP) analysis was set up in order to complement the revealed preference analysis carried out in the previous chapter and get an insight into the awareness and perception with regard to soil contamination among farmers. The discrete choice experiment (DCE) technique was preferred as an SP methodology, because DCEs are more appropriate to reveal preferences for environmental attributes that are integrated into other products. Moreover, choice modeling techniques are generally considered to be superior to contingent valuation techniques, since respondents are not asked to assign monetary values directly.

One of the most important aspects in designing a DCE study is selecting the attributes and its associated attribute levels. In this study, attributes were selected on the basis of complementarity to the hedonic price analysis and the outcomes of two focus group meetings. Soil productivity, farmland location and fertilizing restrictions were found to be important features in farmland purchase decisions that have not been included in the hedonic pricing analysis. Therefore, these characteristics were integrated as attributes or attribute levels in the DCE. Fertilizing restrictions have been incorporated in the attribute of land use restrictions jointly with soil contamination and permanent pasture. With respect to soil contamination, respondent were informed that the soil threshold for one contaminant has been exceeded because of which they were not allowed to cultivate any field crops or vegetables on these parcels. The other attributes, i.e. price and lot size, have been assigned attribute levels based on their distributions in the hedonic dataset.

Subsequently, a main effects, D-efficient design for a multinomial logit (MNL) model was developed in order to reduce the number of choice sets in the survey and to minimize the standard errors on parameter estimates. This resulted in 16 choice sets that were broken down into two blocks. Subsequently, each block of 8 choice sets was randomized five times before inclusion in the survey. Respondents were selected by simple random sampling from the Boerenbond membership list. Surveys were administered in face-to-face interviews among 200 farmers in all municipalities in the province of Limburg that were located at least for 50% in the Campine region. The sample seemed to include particularly professional farmers with an active role in the local farmland market.

The results of the MNL model show that the presence of soil contamination and the resulting crop restrictions significantly reduces farmland utility compared to the situation in which no land use restrictions are applied. Apparently, soil contamination does play a decisive role in the farmers' land purchase decision. In a discrete choice setting farmers indicate to favor unrestricted farmland over farmland that has been affected by soil contamination. When compared to the results of the hedonic price analysis, this result seems to confirm a number of hypotheses suggested in the previous chapter. Firstly, in case farmers are explicitly informed on the soil contamination and the resulting land use restrictions are effectively applied, farmers are unlikely to choose contaminated parcels. This might be an indication that farmland buyers have been improperly informed on the soil contamination and its consequences in farmland transactions. Secondly, the introduction of interaction terms in the MNL model suggested that specialist dairy farms pay less attention to the presence of soil contaminants. Since there are no resulting restrictions with regard to the production of fodder crops and usage as pasture, these farmers might not experience contaminated parcels as 'restricted'. Seeing that specialist dairy farms are the principal type of farming in the Campine region, this might justify the lack of a depreciating effect in the hedonic analysis as well. Thirdly, the coefficient for a situation in which the use of fertilizers is restricted by 25% is significantly lower than the coefficient for the soil contamination attribute level. This finding reinforces the argument that farmers in the Campine region are mainly concerned about their manure disposal capacity. If farmers do not have sufficient land at their disposition, they cannot spread the manure from their activities across fields and need to pay high treatment costs. That is why these farmers tend to be particularly interested in fertilizing restrictions when buying farmland, which might encourage them to ignore potential contamination risks.

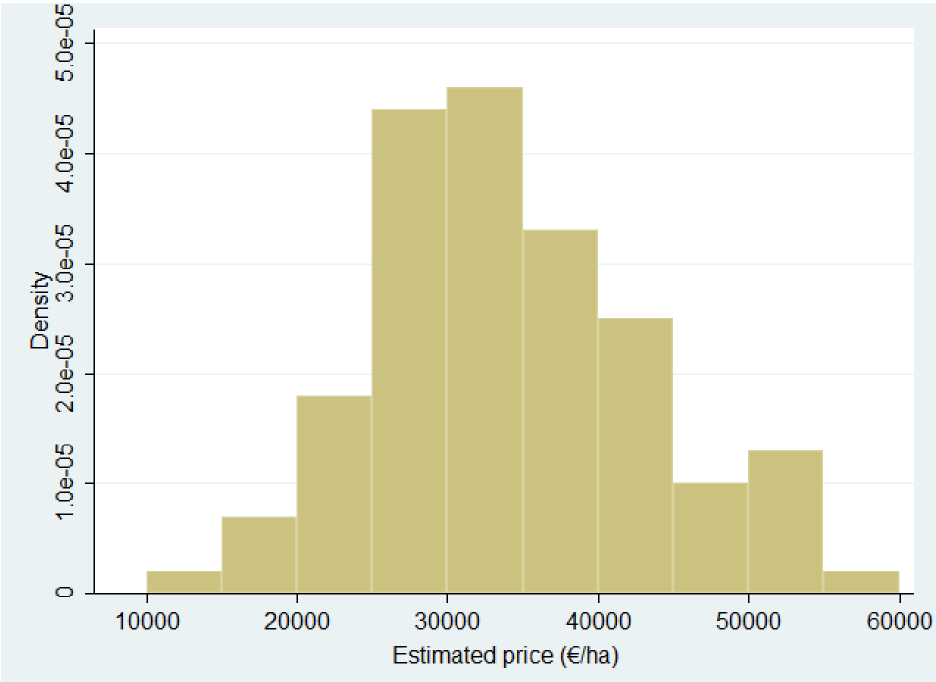
The other attributes in the DCE generally showed the expected results, but diverged in terms of significance. Notably, the attributes that seemed to be relevant for farmers besides the land use restrictions were soil productivity and the size of the parcel. Farmland that was marked as being low productive was considered to be quite undesirable in comparison with (rather) high productive land. With respect to lot size the sample of farmers in the DCE seemed to prefer larger farmland parcels because of the economies of scale that can be acquired from these parcels. A similar finding was reported for farmland located close to other parcels in the farmer's cultivation area. Small parcels of 0.5 ha were deemed as most unattractive, while these kinds of parcels were frequently occurring in the hedonic dataset. Moreover, in the quantile regression analysis it was found that lot size had a negative impact in the farmland's upper price range. This can be interpreted as a confirmation of the hypothesis that these pieces of farmland might have been bought for residential purposes or by horse farmers, a type of farming that was not included in the DCE sample.

In the last section the farmer's willingness to pay for farmland attributes is estimated. In comparison with results from the hedonic pricing analysis and the real and estimated farmland prices in the area, the resulting marginal WTP estimates are too high for them to be realistic. Since it concerns the valuation of a private good (farmland) which the respondents are quite familiar with, this seems to be related in the first place to the study methodology and not to the ignorance of the respondents with the good to be valued. Our discussion on hypothetical bias in SP analyses, and DCEs more in particular, indicated that the overestimation of marginal WTP values is a rather common phenomenon. Therefore, a flow of literature has emerged that concentrates on approaches to eliminate this bias. Although the results for some of these techniques are promising, future research needs to focus on improving the reliability and the validity of these SP induced WTP estimates and specifying the essential conditions for applying SP analyses.

Table 6.10: Distribution of responding farmers over the municipalities

Municipality	# farmers
As	1
Beringen	2
Bocholt	25
Bree	22
Dilsen-Stokkem	9
Ham	2
Hamont-Achel	17
Hasselt	7
Hechtel-Eksel	2
Heusden - Zolder	1
Kinrooi	18
Lommel	7
Lummen	6
Maaseik	13
Maasmechelen	5
Meeuwen-Gruitrode	10
Neerpelt	15
Opglabbeek	4
Overpelt	7
Peer	23
Tessenderlo	4

Figure 6.2: Histogram of respondents' estimated farmland prices



7 Conclusion

7.1 General conclusion

In the introduction we started out by stating that given the relatively recent interest in environmental aspects in the global economic development agenda, there is still much to know about the true value environmental systems provide to society. The aspiration of this dissertation was to provide a contribution on the way towards a better understanding of the economic value of resource-environmental systems and how disturbances can have an economic impact on a societal level. Conveniently, we had the opportunity to advance on a case of environmental pollution that has been the subject of many interdisciplinary research projects in the last couple of decades, i.e. the Campine region in Belgium. The work of numerous researchers from different backgrounds has generated a lot of data and knowledge on the situation, which we were able to benefit from in a number of occasions throughout the dissertation.

This dissertation proceeded on this firm research basis to discover to what extent there are currently still costs associated with the remaining soil contamination in the area. Since the soil pollution can be labeled as an external cost from zinc refinery processes, the economic damage produced by the soil pollution can be measured using nonmarket valuation techniques. In case the externality appears to bring about significant costs, these costs should be internalized in the cost-benefit analysis of remediation strategies, which will allow for a more accurate analysis. In this setting, the dissertation basically focused on two main aspects of the pollution, namely the economic impact on farmland in the area and the health economic impact on the population level. Herein, an important role is played by nonmarket valuation techniques, the environmental economic methodologies that are applied to analyze the external costs from different perspectives.

The first research chapter, i.e. **chapter 2**, had a double objective in this setting. On the one hand, the prediction maps of the soil contamination in the Campine region that were created in this chapter served as a basis for both the health related and farmland research in the next chapters. By means of the prediction maps goal was to obtain a better understanding of the extent and the austerity of the soil contamination. On the other hand, the chapter also aimed to determine the amount of farmland that has been affected by the contamination. Seeing that the problems associated with soil contamination can affect the agricultural potential of farmland (which is discussed more extensively in chapter 1), this can provide a justification for investigating the impact on farmland values further in the next chapters.

With respect to the first objective, a dataset of soil Cd concentrations that was gathered in previous research projects, was at our disposal. This dataset included 11,885 sample measurements from fourteen Belgian and Dutch municipalities in a wide range from the pollution sources. These samples were used to predict soil Cd levels in the study area using ordinary kriging, a geostatistical interpolation methodology based on regression techniques. This method enables us to infer soil Cd values at unsampled locations by relating values of neighboring observations to the distance to the unobserved point. The resulting prediction maps show that the soil contamination is closely related to the presence of pollution sources. In the vicinity of three of the four Zn smelters located in the northern Campine, there are clear Cd deposition plumes noticeable in northeastern directions. This conforms to what was expected, since wind current flows in Belgium predominantly arise from southwestern directions.

The second objective of chapter 2 was to determine the agricultural area that has been affected by soil contamination as to derive the amount of farmland that can potentially dedicated to phytoremediation. It was found that more than 3000 hectares of farmland exceeded the Cd guide value, while almost 2000 hectares surpassed the Cd threshold value set by the Flemish Government. These results clearly indicated that there is a rationale for researching the economic impact on farmland values in the next chapters. To determine which land can be dedicated to phytoremediation, we relied on data from a phytoremediation field experiment that has recently been conducted in the Campine region. With respect to crop choices, the focus was solely on willow and a rotation of tobacco and maize, because these cultivations achieved the best results in the field experiment. The fact that 60% and 80% of the contaminated farmland area (>1.2 ppm) can be cleaned within respectively 21 and 42 years using willow confirms the potential of phytoremediation as a promising strategy to redevelop mildly contaminated soils suitable for crop cultivation.

Deviating from the dissertation's structure, we will immediately focus on our research on the economic impact of soil contamination on agricultural land value. Given the numerous adverse effects farmers are faced with as a result of farmland pollution and the extent of the problem in the Campine region, it is hypothesized that the presence of pollutants will have an influence on the value of farmland. In **chapter 5**, the objective was to examine to what extent farmland buyers value soil contamination based on a dataset of 599 sales transactions that have taken place in the Campine region between 2004 and 2011. Hedonic pricing analysis (HPA) is a revealed preference technique that uses prices of differentiated products (i.e. farmland values in this case) to identify the marginal implicit prices of value-adding and value-reducing product characteristics (i.e. soil contamination in this case). Therefore, the sales prices are linked to soil Cd concentrations by means of the prediction maps created in chapter 2. The inherent spatial nature of real estate transactions and its associated issues are taken into account by extending classic spatial econometric techniques. Moreover, classic linear regression, which estimates conditional mean parameters, is complemented by a quantile regression approach providing conditional median estimates in order to test the robustness of the results.

The results of the hedonic regressions indicated that soil Cd levels were insignificant determinants for farmland prices. Apparently, the presence of soil contamination did not affect the value farmland buyers were willing to pay for agricultural land in the area and thus neither the asset value of contaminated landowners. The other independent variables in the hedonic model did provide some insight into clarifications for this result and the way farmland price formation in general. A number of significant price determinants suggest that the farmland market in the Campine region has dealt with an increasing supply scarcity in the last couple of years. For example, while all observations were labeled as agricultural land in the administration, a considerable amount of parcels were sold in other zoning types such as residential and nature zoning, indicating there is a pressure for redeveloping farmland. In case agricultural land is (expected to be) transformed for environmental purposes, a substantial part of its use value to farmers is lost, which was reflected in farmland prices. On the other hand, if farmland is expected to be used for residential purposes in the future, powerful inflating effects in land prices were observed.

Several other potential explanations arose which could possibly explain the lack of a significant effect from soil contamination in the HPA. Therefore, we have tried to find confirmation for some of these hypotheses by further unraveling the motivational framework underlying farmland price determination in **chapter 6**. The objective of this chapter was to investigate the role of soil contaminants in farmland purchase decisions by means of stated preference (SP) techniques. More specifically, a discrete choice experiment (DCE) was set up to get an insight into the awareness and perception with regard to soil contamination among farmers and to complement the revealed preference HPA where possible. In a DCE multiple alternatives, which are described by different attributes with varying attribute levels, are presented to respondents in choice sets. The respondent is then asked to pick one single alternative from each choice set, thereby stating his/her preference for certain attributes or attribute levels. Since SP techniques allow to independently create a hypothetical farmland market and manipulate the information that is provided to the respondents, the respondents can be explicitly informed on the soil contamination and its resulting land use restrictions. Furthermore, the DCE allows for the identification of potential heterogeneity among farmers with regard to the issue and facilitates the estimation of parameters that are difficult or impossible to assess using revealed preference models alone.

Data for the DCE were obtained from in-person interviews with 200 farmers, all members of the largest farmer association in Flanders. The results of the basic multinomial logit (MNL) model showed that the presence of soil contamination and the resulting crop restrictions significantly reduces farmland utility in comparison with a situation in which no land use restrictions are applied. Hence, in case farmers are explicitly informed on the land use restrictions resulting from soil contamination, they seem to favor parcels of farmland that are not affected by soil contamination. When we tried to take into account the heterogeneity in the sample by introducing interaction terms in the MNL model, it was found that specialist dairy farmers attach significantly less value to the presence of soil contaminants in comparison with the average farmer.

If the results of the HPA and the DCE with respect to the effect of soil contamination on the value of farmland are combined, some potential explanations can be put forward. For example, in contrast to the DCE, the farmland transactions in the HPA might not just concern farmers alone. If farmland is bought for other purposes than agriculture, the buyers might not consider the soil contamination to be harmful and their willingness to pay will be unaffected. Additionally, it has been suggested before that the Flemish farmland market is increasingly experiencing farmland scarcity. It seems that the continuous, above-inflation rise in farmland prices has spurred farmers to delay on offering farmland on the market in hopes of higher returns in the future. This way, farmers with the intention of expanding their professional activities are experiencing difficulties to find suitable farmland to achieve their objectives. Consequently, the market situation might force these farmers to buy pieces of farmland with inferior or inappropriate land attributes such as elevated soil Cd levels.

Another explanation for the discrepancy in results might be related to the information on the soil contaminants, and the lack thereof in the HPA. The policy analysis in chapter 5 pointed out that soil certificates may not provide the necessary information on soil pollutants to (farm)land buyers, which might have made them unaware of the exact concentrations of heavy metals present in the acquired parcels. In this case, farmland buyers would have been unable to take their presence into account in price setting. When farmers were explicitly informed on this matter in the DCE, they indicated to prefer uncontaminated parcels. Moreover, the DCE also mentioned that as a result of the soil contamination, farmers were not allowed to cultivate field crops and vegetables for food safety reasons. However, since these land use restrictions are not strictly enforced in practice, farmers might be unaware of these consequences as well.

Lastly, the DCE found evidence that the importance of soil contaminants to farmers differs between farm types. Specialist dairy farmers were significantly less interested in the presence of soil contaminants, although their net effect on farmland utility was still negative. These farmers do not intend to use the acquired land for crops with a considerable risk of exceeding food thresholds, but rather for the production of fodder crops and as pasture. Since there are no restrictions with respect to these land uses, this type of farming might not consider contaminated parcels as 'restricted' and their willingness to pay might not be affected. Seeing that specialist dairy farms are the principal type of farming in the Campine region, this might justify the lack of a negative effect in the hedonic analysis as well. The DCE analysis also revealed the importance of fertilizing restrictions to farmers, which reinforces the argument that farmers in the area are particularly interested in farmland without fertilizing restrictions rather than farmland without land use restrictions.

In the two remaining chapters, i.e. **chapter 3 and 4**, the health economic damages brought about by the pollution are investigated. In this setting, a damage function approach (DFA) is used to estimate the damage costs attributable to the environmental Cd pollution. Typically, a DFA is divided into three sections. In the first stage, epidemiologic research has established an association between environmental pollution levels and the incidence of certain illnesses. In three study periods from the middle of the 1980's until 2004, two major cohort studies in the Campine region indicated that the presence of environmental Cd has led to an increased risk of fractures related to osteoporosis and lung cancer. This research provided us with the relative risks needed to start the next stage of the DFA. In the second stage (chapter 3), a health risk assessment aimed to determine the pollution-attributable illnesses on a population level, while the third stage (chapter 4) estimated the costs related to those illnesses using the damage cost method.

The first step of the health risk assessment in **chapter 3** involved estimating the amount of people that have been exposed to certain levels of pollution by means of the prediction maps created in the chapter 2. It has been suggested before that soil Cd values are satisfactory proxies for exposure to Cd. Each level of soil contamination was then linked to the relative risks of fractures and lung cancer that have been estimated in the former epidemiologic research and combined with the most recently available incidence data of these illnesses in Belgium or Flanders in order to estimate the number of people that have experienced adverse health effects. Assuming the relative risks have remained constant throughout the years, there are each year approximately 22 cases of lung cancer, 8.5 hip fractures in men and 32 in women attributable to the pollution in the Campine region. However, it is quite plausible that the relative risks have decreased ever since the epidemiologic studies have been finalized. On the condition that the measures for preventing the population's exposure to Cd are maintained, it can be expected that the effect of the environmental pollution on public health will fade out in the future, as the fraction of the population that has been exposed to high levels of Cd passes away.

In **chapter 4**, the economic evaluation of the health effects caused by the pollution is the research topic. The damage cost method – in this context also referred to as Cost Of Illness (COI) – was employed to estimate the average cost of hip fractures and lung cancer in Belgium. Recently, a COI study at University of Liège analyzed the hospitalization costs of hip fracture patients in Belgium and found that mean hospitalization costs amounted to €10,828 and €10,389 in men and women, respectively. Since the data emerged from the Belgian health care system and are up to date, these results were transferred to our research. Lung cancer costs were extracted from the Intermutualistic Agency's (IMA) permanent sample, a database comprising health care expenditures of a sample of 300,000 Belgian citizens, using a longitudinal matched case-control design. This way, all costs covered by the mandatory health care insurance were administered from 3 months prior to diagnosis until 18 months after diagnosis or until mortality. The incremental costs of lung cancer were determined by creating two matching control samples with similar socio-economic characteristics as the patient sample. The second control sample additionally controlled for the main comorbidities of smoking, i.e. COPD and high cholesterol levels. Since it assumed that the lung cancer incidence in the Campine region is caused by the environmental pollution, the adjustment allows for an exclusion of health care costs due to smoking, the principal risk factor for lung cancer.

The results of the COI analysis show that lung cancer patients incur incremental costs in the range of €25,000 - €30,000 per patient due to the illness. Especially in the first stages of lung cancer, the period of time shortly before and after diagnosis, patients have to deal with extensive medical costs. Although the Belgian health care system covers the largest part of these costs, the out-of-pocket costs can present a considerable burden to patients. For calculating the pollution-attributable health care costs, the incremental costs of lung cancer in comparison with the smoking-adjusted control sample over the 21 month time window are selected, which amount to €25,385.60. This results in more than half a million euro per year in costs due to pollution-attributable lung cancer cases. For hip fractures, the health care costs yearly add up to €92,053 for men and €326,334 for women.

Notably, these cost estimations are based on the assumptions of constant relative risks in the health risk assessment.

In conclusion, the environmental pollution in the Campine region has generated a long-term burden to the area, which is still at play up to this day. More than 25 years after the Zn smelters in the area have stopped emitting heavy metals, the population is still facing illnesses which are attributable to the pollution, while farmers have to deal with the constraints imposed by the contamination of their land. This exemplary case must be a strong incentive for governments all across the world, but especially in industrializing countries, to take environmental aspects into account in framing and guiding economic development. The implementation of stringent environmental regulations combined with a closely controlled monitoring system can prevent the pollution from occurring and save future generations from bearing this burden. Although this strategy calls for relatively high start-up costs, it is more likely to fit the environment’s capacity to absorb waste products and protect the environment from disturbances. Moreover, it has the potential to reduce or possibly eliminate the long-term (external) costs associated with environmental pollution, which might turn out to be economically more beneficial in the end.

The governments in Belgium seem to have selected such a strategy for preventing similar pollution cases from happening in the future. The health damages in the area are likely to decrease further because aerial Cd concentrations have been brought down to insignificant levels and information campaigns have recommended the population to refrain from cultivating and eating homegrown vegetables. Both measures have reduced the population’s exposure to heavy metals. However, the question remains what the exact contribution is from soil contamination to human exposure and thus to these health damages and whether action needs to be undertaken in order to deal with this issue. From an agricultural perspective, there does not appear to be a significant economic impact of soil contamination since no discount in farmland values is found in contaminated parcels. The most important explanatory factors for this result are probably the (farm)land scarcity in Flanders and the specific agricultural situation in the area. Nevertheless, the precautionary measures that have currently been implemented by the different governments must be upheld and the situation needs to be monitored to prevent the contaminants from leaching into groundwater reserves, which can bring about another set of external costs.

7.2 Constraints

In this section an overview will be presented of the most important constraints and limitations that have been dealt with in this dissertation. Most of these constraints concern restrictions in data (availability) and methodological issues. They will be discussed according to the order in which they have occurred in the text.

Although it was a great advantage for us to be able to work with a study area that has been extensively researched in the past, this also confined the research to some extent. In chapter 2, for example, there was a significant amount of soil Cd sample measurements at our disposition. However, it was delivered to us ‘as is’, implying that only the soil Cd concentration and the geographic coordinates of all measurements were given, not the point in time at which the sample had been taken or by whom. According to the prediction maps that resulted from this dataset, the vicinity of the Zn smelter that was dismantled in the 1970s (Lommel Maatheide) seems to deal with lower contaminant levels. However, this might be explained by the samples on which the predictions were based. A considerable part of them were probably taken at the Zn smelter site of 8000 m² that has been rehabilitated by soil excavation in 2003-2004 in order to confirm its clean status. Since relatively little samples were taken in the surrounding areas, these predictions might be somewhat biased towards lower soil Cd levels. Nevertheless, this potential bias is not expected to significantly impact the estimations in chapter 3 and 5, for which the prediction maps provided the starting point. The neighborhood is dominated by nature and reclamation zoning types, while we were particularly interested in residential (for the health related chapters) and agricultural land use (for the farmland chapters).

In the health related research, we were constrained by the epidemiologic research and the evidence provided by Staessen and Nawrot. At present, there was little we were able to do (as economists) about the choices made in this research, for example with regard to the health endpoints included in and excluded from the analysis. Since no detailed information on the data used in this research was available, the assumption was made that lung cancer and osteoporotic fractures were the only health endpoints for which significant evidence was found. If the doctors and researchers noticed a systematic reoccurrence of other health endpoints in the study samples, it is assumed that they would have reported it. From a research perspective, we needed to rely on the judgments made by the epidemiologists to create research designs that are most appropriate for finding out the effect of environmental exposure to heavy metals on the incidence of illnesses. There is no reason for us to assume that the researchers did not follow the guidelines for conducting proper scientific research or held back information on other health related issues. Furthermore, the transferred relative risks were assumed to have remained constant in the health risk assessment. Seeing that the epidemiologic research has been performed a number of years ago, this assumption is likely to have caused an overestimation of the amount of pollution-attributable illnesses.

As has been extensively discussed in chapter 3 and 4, the application of the damage function approach as well as our decision to focus solely on direct medical costs reduces the scope of the health economic research. Since the damage cost method focuses on resource costs and excludes the utility value of changes in service flows, a substantial part of the societal WTP for illnesses is left out of the estimations. The decision to restrict our research design further was primarily made to comply with our objective to provide a rough estimate of the health related pollution-attributable costs to be expected in the future. Moreover, time constraints also forced us to focus on what is probably the most relevant part of health economic damages, i.e. the direct medical costs. Determining the non-medical and indirect costs due to the illnesses would have required a much more elaborate research design and simply falls outside the scope of this dissertation in environmental economics. However, as a result of this restricted design, the cost estimates provided in chapter 4 are likely to be quite conservative in comparison with the true societal costs of these illnesses.

With respect to the hedonic model, the main constraint emerged from the limited extent of the study area on the one hand, and the relatively small contaminated area on the other hand. The former aspect might have caused relatively little variation in farmland prices and independent variables across the study area. This prevented us, for example, from adding important structural variables such as soil quality and climate variables in the hedonic model specification. The latter aspect implied that there were only a limited amount of contaminated parcels in the dataset (see section 5.4), which might have distorted the regression analysis to some extent. It would have been interesting to find out whether the results would change in case the dataset included more contaminated parcels and in case the predominant farming type in the area was constrained in terms of land use due to the contamination. At present, it seems as though the specialist dairy farmers do not seem to mind the presence of soil contaminants as long as it does not restrict their land allocation. Furthermore, if the farmland was not facing the supply scarcity that is currently at play in the Campine region and Flanders as a whole, soil contaminants might have been more likely to have an impact on farmland values. However, this can only be confirmed by hedonic farmland research using a dataset containing a greater share of contaminated parcels.

The research design was the most important limiting factor in the discrete choice experiment. In contrast to the hedonic pricing analysis, there is little that can be done a posteriori once the data collection is finalized. The attributes and its associated levels need to be chosen in a rather arbitrarily fashion, without the reassurance that the selected characteristics will actually be the dominating factors in farmland purchasing decisions. Therefore, we thoroughly tested the design in panel sessions with two independent and experienced focus groups. A follow-up multiple choice question at the end of our survey (which was not used in the analysis) confirmed that the selected attributes were likely to be the most important price determinants among farmers in the area. On the other hand, the experimental design that has been used to introduce the choice experiments in the surveys is created for estimating a multinomial logit model. This design assumes, among other things, that the attributes are independent from each other. Although it still allows to estimate a random parameter logit model, it is impossible to specify a non-linear functional form in which interaction effects between attributes are introduced. This is a factor that could possibly explain some of the results from this analysis as well.

7.3 Suggestions for future research

As in every research, there are a number of questions that have been answered, but at the same time some questions remain and some new questions have been raised. Our main concern in this setting goes out to methodological issues, because there still seems to be some significant uncertainties associated with a number of environmental economic methodologies. When research on environmental damages is concerned, the most significant impact generally relates to the effects on public health. However, the two predominant methods for estimating and economically quantifying these effects, i.e. the averting behavior method and the damage function approach, both deal with considerable shortcomings that have been extensively discussed in section 3.2. Although there often is some overlap between environmental and health economic studies in this setting, it seems as though the two research domains are still somewhat separated from each other. It appears to me that an intensified relationship between environmental and health economics has the potential to result in a more comprehensive framework for estimating environmental damage costs by cross-pollination. Environmental economists can for example benefit from the experience of health economists in dealing with one of society's most important nonmarket goods, i.e. health, and transfer some of the concepts and ideas to environmental nonmarket goods.

This collaboration can also be useful in the aspiration of improving economic estimates from stated preference (SP) analyses, since both research domains frequently make use of SP methodologies. As the discussion in section 6.4.3 points out, the overestimation of WTP values is common practice in SP research. The transfer from predominantly contingent valuation studies to choice modeling approaches has not generated the expected increase in external validity, as the limited empirical evidence indicates that hypothetical bias is even more existent in choice modeling. Although we have tried to take into account a number of elements for estimating realistic economic values from the DCE research (familiar good and context, pre-testing in focus groups, no order effects or starting point bias, minimal chances of strategic behavior,…), the WTP values turn out to be markedly overestimated. That is why future research should focus on the external validity of SP derived WTP estimates. One of the most suitable approaches for researching this topic is comparing the results from SP analyses with RP analyses, analogous to this research. However, instead of complementing each other, researchers should try to create similar hypothetical and real research designs. The application of hypothetical and real choice experiments is one example of ways to achieve this goal.

As mentioned in chapter 5, to the best of our knowledge a hedonic pricing analysis of contaminated farmland has never been done before. Since the pollution of agricultural land with anomalous elements occurs quite frequently across the world, this research gap can probably be explained by a limited availability of individual level farmland values and the major data requirements on soil contamination. It has been shown in chapter 5 that collecting an adequate dataset for this kind of research is not particularly a walk in the park. A substantial part of hedonic farmland research focuses on farmland prices on the municipality level, which is not suitable for a very specific, small-scale characteristic such as soil contamination. Future hedonic studies should therefore focus on collecting individual farm-level data and determining the impact on farmland prices shortly after the contamination has taken place. Moreover, the farmland market situation needs to be discussed in order to obtain a better understanding of the particular effects in a certain area.

Lastly, analyses of environmental damage situations should prefer using social cost benefit analyses (CBA) to economically assess the need for cleaning up the environment or preventing the damage from occurring in the first place. The more widely used private CBA does not take into account the external benefits that can be achieved or the external costs that can be reduced by protecting environmental systems from disturbances. It has been shown in this research that these external effects can be substantial. Governments and policy makers have an important example function in this setting by implementing regulations that require the integration of environmental effects in investment decisions. Obviously, this strategy will be most valuable in case these effects are economically quantified so that they can be compared to resource cost flows for a complete overview of all policy alternatives. This might be another incentive to upgrade and optimize environmental economic research methods.

References

Abelairas-Etxebarria, P. and Astorkiza, I. (2012). Farmland prices and land-use changes in periurban protected natural areas. Land Use Policy 29(3): 674-683.

Aboyans, V., Thomas, D. and Lacroix, P. (2010). The cardiologist and smoking cessation. Current Opinion in Cardiology 25(5): 469-477.

Adamowicz, W., Louviere, J. and Williams, M. (1994). Combining revealed and stated preference methods for valuing environmental amenities. Journal of Environmental Economics and Management 26(3): 271-292.

Adams, S. V., Newcomb, P. A., Shafer, M. M., Atkinson, C., Bowles, E. J. A., Newton, K. M. et al. (2011). Sources of cadmium exposure among healthy premenopausal women. Science of the Total Environment 409(9): 1632-1637.

Ahmadi, S. H. and Sedghamiz, A. (2007). Geostatistical analysis of spatial and temporal variations of groundwater level. Environmental Monitoring and Assessment 129(1-3): 277-294.

Akesson, A., Bjellerup, P., Lundh, T., Lidfeldt, J., Nerbrand, C., Samsioe, G. et al. (2006). Cadmium-induced effects on bone in a population-based study of women. Environmental Health Perspectives 114(6): 830-834.

Akesson, A., Julin, B. and Wolk, A. (2008). Long-term dietary cadmium intake and postmenopausal endometrial cancer incidence: A population-based prospective cohort study. Cancer Research 68(15): 6435-6441.

Akesson, A., Lundh, T., Vahter, M., Bjellerup, P., Lidfeldt, J., Nerbrand, C. et al. (2005). Tubular and glomerular kidney effects in Swedish women with low environmental cadmium exposure. Environmental Health Perspectives 113(11): 1627-1631.

Alberini, A., Gans, W. and Alhassan, M. (2011). Individual and public-program adaptation: Coping with heat waves in five cities in Canada. International Journal of Environmental Research and Public Health 8(12): 4679-4701.

Alberini, A., Hunt, A. and Markandya, A. (2006). Willingness to pay to reduce mortality risks: Evidence from a three-country contingent valuation study. Environmental & Resource Economics 33(2): 251-264.

Alberini, A. and Scasny, M. (2011). Context and the VSL: Evidence from a stated preference study in Italy and the Czech Republic. Environmental & Resource Economics 49(4): 511-538.

Aldy, J. E. and Viscusi, W. K. (2007). Age differences in the value of statistical life: Revealed preference evidence. Review of Environmental Economics and Policy 1(2): 241-260.

Alfven, T., Elinder, C. G., Carlsson, M. D., Grubb, A., Hellstrom, L., Persson, B. et al. (2000). Low-level cadmium exposure and osteoporosis. Journal of Bone and Mineral Research 15(8): 1579-1586.

American Cancer Society. (2008). Cancer facts & figures 2007. Atlanta, Georgia: American Cancer Society.

Anselin, L. (1988). Spatial Econometrics: Methods and Models. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Ashenfelter, O. (2006). Measuring the value of a statistical life: Problems and prospects. Economic Journal 116(510): C10-C23.

Autier, P., Haentjens, P., Bentin, J., Baillon, J. M., Grivegnee, A. R., Colson, M. C. et al. (2000). Costs induced by hip fractures: A prospective controlled study in Belgium. Osteoporosis International 11(5): 373-380.

Baker, D. and Nieuwenhuijsen, M. (2008). Environmental Epidemiology: Study Methods and Application. Oxford: Oxford University Press.

Balistreri, E., McClelland, G., Poe, G. and Schulze, W. (2001). Can hypothetical questions reveal true values? A laboratory comparison of dichotomous choice and open-ended contingent values with auction values. Environmental & Resource Economics 18(3): 275-292.

Bastian, C. T., McLeod, D. M., Germino, M. J., Reiners, W. A. and Blasko, B. J. (2002). Environmental amenities and agricultural land values: a hedonic model using geographic information systems data. Ecological Economics 40(3): 337-349.

Bateman, I., Carson, R., Day, B., Hanemann, M., Hanley, N., Hett, T. et al. (2002). Economic Valuation with Stated Preference Techniques: A Manual. Cheltenham: Edward Elgar.

Beck, M. J., Rose, J. M. and Hensher, D. A. (2013). Consistently inconsistent: The role of certainty, acceptability and scale in choice. Transportation Research Part E-Logistics and Transportation Review 56: 81-93.

Bellavance, F., Dionne, G. and Lebeau, M. (2009). The value of a statistical life: A meta-analysis with a mixed effects regression model. Journal of Health Economics 28(2): 444-464.

Ben-Akiva, M. and Lerman, S. (1985). Discrete Choice Analysis: Theory and Application to Travel Demand. Cambridge, Massachusetts: MIT Press.

Bergmann, A., Hanley, M. and Wright, R. (2006). Valuing the attributes of renewable energy investments. Energy Policy 34(9): 1004-1014.

Bliemer, M. C. J. and Rose, J. M. (2011). Experimental design influences on stated choice outputs An empirical study in air travel choice. Transportation Research Part a-Policy and Practice 45(1): 63-79.

Blumenschein, K., Blomquist, G. C., Johannesson, M., Horn, N. and Freeman, P. (2008). Eliciting willingness to pay without bias: Evidence from a field experiment. Economic Journal 118(525): 114-137.

Boardman, A. E., Greenberg, D. H., Vining, A. R. and Weimer, D. L. (2011). Cost-Benefit Analysis: Concepts and Practice. Upper Saddle River, New Jersey: Pearson Education.

Bobinac, A., van Exel, N. J. A., Rutten, F. F. H. and Brouwer, W. B. F. (2010). Willingness to pay for a Quality-Adjusted Life-Year: The individual perspective. Value in Health 13(8): 1046-1055.

Bobinac, A., van Exel, N. J. A., Rutten, F. F. H. and Brouwer, W. B. F. (2012). Inquiry into the relationship between equity weights and the value of the QALY. Value in Health 15(8): 1119-1126.

Bomans, K., Duytschaever, K., Gulinck, H. and Van Orshoven, J. (2010). Tare land in Flemish horticulture. Land Use Policy 27(2): 399-406.

Borgstrom, F., Sobocki, P., Strom, O. and Jonsson, B. (2007). The societal burden of osteoporosis in Sweden. Bone 40(6): 1602-1609.

Borgstrom, F., Zethraeus, N., Johnell, O., Lidgren, L., Ponzer, S., Svensson, O. et al. (2006). Costs and quality of life associated with osteoporosis-related fractures in Sweden. Osteoporosis International 17(5): 637-650.

Boyd, D. R. and Genuis, S. J. (2008). The environmental burden of disease in Canada: Respiratory disease, cardiovascular disease, cancer, and congenital affliction. Environmental Research 106(2): 240-249.

Boyle, K. (2003). Contingent valuation in practice. In Champ, P., Boyle, K. & Brown, T. (eds.), A Primer on Nonmarket Valuation. Dordrecht: Kluwer Academic Publishers, 111-176.

Braden, J. B., Feng, X. and Won, D. (2011). Waste sites and property values: A meta-analysis. Environmental & Resource Economics 50(2): 175-201.

Brady, M. and Irwin, E. (2011). Accounting for spatial effects in economic models of land use: Recent developments and challenges ahead. Environmental & Resource Economics 48(3): 487-509.

Braithwaite, R. S., Meltzer, D. O., King, J. T., Leslie, D. and Roberts, M. S. (2008). What does the value of modern medicine say about the $50,000 per Quality-Adjusted Life-Year decision rule? Medical Care 46(4): 349-356.

Brazier, J. E., Green, C., Kanis, J. A. and Comm Sci, A. (2002). A systematic review of health state utility values for osteoporosis-related conditions. Osteoporosis International 13(10): 768-776.

Briggs, D. (2003). Environmental pollution and the global burden of disease. British Medical Bulletin 68: 1-24.

Byrne, M. M., O'Malley, K. and Suarez-Almazor, M. E. (2005). Willingness to pay per quality-adjusted life year in a study of knee osteoarthritis. Medical Decision Making 25(6): 655-666.

Callan, S. J. and Thomas, J. M. (2010). Environmental Economics and Management: Theory, Policy, and Applications. Mason, OH: South-Western Cengage Learning.

Cambardella, C. A., Moorman, T. B., Novak, J. M., Parkin, T. B., Karlen, D. L., Turco, R. F. et al. (1994). Field-scale variability of soil properties in central Iowa soils. Soil Science Society of America Journal 58(5): 1501-1511.

Cameron, T. A. (2010). Euthanizing the value of a statistical life. Review of Environmental Economics and Policy 4(2): 161-178.

Cameron, T. A. and DeShazo, J. R. (2013). Demand for health risk reductions. Journal of Environmental Economics and Management 65(1): 87-109.

Carlsson, F., Frykblom, P. and Lagerkvist, C. J. (2005). Using cheap talk as a test of validity in choice experiments. Economics Letters 89(2): 147-152.
Carlsson, F. and Martinsson, P. (2001). Do hypothetical and actual marginal willingness to pay differ in choice experiments? Application to the valuation of the environment. Journal of Environmental Economics and Management 41(2): 179-192.
Carson, R. T. and Groves, T. (2007). Incentive and informational properties of preference questions. Environmental & Resource Economics 37(1): 181-210.
Cavailhes, J. and Thomas, I. (2013). Are agricultural and developable land prices governed by the same spatial rules? The case of Belgium. Canadian Journal of Agricultural Economics-Revue Canadienne D Agroeconomie 61(3): 439-463.
Cavailhes, J. and Wavresky, P. (2003). Urban influences on periurban farmland prices. European Review of Agricultural Economics 30(3): 333-357.
Champ, P. (2003). Collecting survey data for nonmarket valuation. In Champ, P., Boyle, K. & Brown, T. (eds.), A Primer on Nonmarket Valuation. Dordrecht: Kluwer Academic Publishers, 59-98.
Champ, P. A., Bishop, R. C., Brown, T. C. and McCollum, D. W. (1997). Using donation mechanisms to value nonuse benefits from public goods. Journal of Environmental Economics and Management 33(2): 151-162.
Chestnut, L. G., Rowe, R. D. and Breffle, W. S. (2012). Economic valuation of mortality-risk reduction: Stated preference estimates from the United States and Canada. Contemporary Economic Policy 30(3): 399-416.
Chouaid, C., Molinier, L., Combescure, C., Daures, J. P., Housset, B. and Vergnenegre, A. (2004). Economics of the clinical management of lung cancer in France: an analysis using a Markov model. British Journal of Cancer 90(2): 397-402.
Ciaian, P., Kancs, A., Swinnen, J., Van Herck, K. and Vranken, L. (2012). Institutional factors affecting agricultural land markets. Brussels: Centre for European Policy Studies (CEPS).
Ciaian, P. and Swinnen, J. F. M. (2006). Land market imperfections and agricultural policy impacts in the new EU Member States: A partial equilibrium analysis. American Journal of Agricultural Economics 88(4): 799-815.
Cipriano, L. E., Romanus, D., Earle, C. C., Neville, B. A., Halpern, E. F., Gazelle, G. S. et al. (2011). Lung cancer treatment costs, including patient responsibility, by disease stage and treatment modality, 1992 to 2003. Value in Health 14(1): 41-52.
Clauw, F. (2007). *The effect of soil contamination on real estate value: A hedonic pricing approach based on different risk measures.* Ph.D. dissertation, Hasselt University, Diepenbeek.
Cleemput, I., Neyt, M., Van de Sande, S. and Thiry, N. (2012). Belgische richtlijnen voor economische evaluaties en budget impact analyses:

tweede editie *Health Technology Assessment (HTA)*. Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE).
Cotteleer, G., Gardebroek, C. and Luijt, J. (2008). Market power in a GIS-based hedonic price model of local farmland markets. Land Economics 84(4): 573-592.
Cotteleer, G., Stobbe, T. and van Kooten, G. C. (2011). Bayesian model averaging in the context of spatial hedonic pricing: An application to farmland values. Journal of Regional Science 51(3): 540-557.
Cotteleer, G. and van Kooten, G. C. (2012). Expert opinion versus actual transaction evidence in the valuation of non-market amenities. Economic Modelling 29(1): 32-40.
Cressie, N. A. C. (1993). Statistics for Spatial Data. New York: Wiley.
Cummings, R. G. and Taylor, L. O. (1999). Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method. American Economic Review 89(3): 649-665.
Day, B., Bateman, I. J., Carson, R. T., Dupont, D., Louviere, J. J., Morimoto, S. et al. (2012). Ordering effects and choice set awareness in repeat-response stated preference studies. Journal of Environmental Economics and Management 63(1): 73-91.
De Laet, C., van Hout, B. A., Burger, H., Weel, A., Hofman, A. and Pols, H. A. P. (1999). Incremental cost of medical care after hip fracture and first vertebral fracture: The Rotterdam Study. Osteoporosis International 10(1): 66-72.
de Vor, F. and de Groot, H. L. F. (2011). The impact of industrial sites on residential property values: A hedonic pricing analysis from the Netherlands. Regional Studies 45(5): 609-623.
Deaton, B. J., Hoehn, J. P. and Norris, P. E. (2007). Net buyers, net sellers, and agricultural landowner support for agricultural zoning. Land Economics 83(2): 153-165.
Deaton, B. J. and Vyn, R. J. (2010). The effect of strict agricultural zoning on agricultural land values: The case of Ontario's Greenbelt. American Journal of Agricultural Economics 92(4): 941-955.
Dedes, K. J., Szucs, T. D., Bodis, S., Joerger, M., Lowy, A., Russi, E. W. et al. (2004). Management and costs of treating lung cancer patients in a university hospital. Pharmacoeconomics 22(7): 435-444.
Dekker, T., Brouwer, R., Hofkes, M. and Moeltner, K. (2011). The effect of risk context on the value of a statistical life: A Bayesian meta-model. Environmental & Resource Economics 49(4): 597-624.
Desaigues, B., Ami, D., Bartczak, A., Braun-Kohlova, M., Chilton, S., Czajkowski, M. et al. (2011). Economic valuation of air pollution mortality: A 9-country contingent valuation survey of value of a life year (VOLY). Ecological Indicators 11(3): 902-910.
Diamond, P. A. and Hausman, J. A. (1994). Contingent valuation - Is some number better than no number? Journal of Economic Perspectives 8(4): 45-64.
Dickie, M. (2003). Defensive behavior and damage cost methods. In Champ, P., Boyle, K. & Brown, T. (eds.), A Primer on Nonmarket Valuation. Dordrecht, The Netherlands: Kluwer Academic Publishers, 395-444.

Dickinson, N. and Pulford, I. (2005). Cadmium phytoextraction using short-rotation coppice: the evidence trail. Environment International 31(4): 609-613.
Dietz, S. and Maddison, D. J. (2009). New frontiers in the economics of climate change. Environmental & Resource Economics 43(3): 295-306.
Dolan, P. and Togerson, D. J. (1998). The cost of treating osteoporotic fractures in the United Kingdom female population. Osteoporosis International 8(6): 611-617.
Donaldson, C., Baker, R., Mason, H., Jones-Lee, M., Lancsar, E., Wildman, J. et al. (2011). The social value of a QALY: raising the bar or barring the raise? Bmc Health Services Research 11: 8.
Dubin, R. A. (1998). Spatial autocorrelation: A primer. Journal of Housing Economics 7(4): 304-327.
Dudka, S., Piotrowska, M. and Terelak, H. (1996). Transfer of cadmium, lead, and zinc from industrially contaminated soil to crop plants: A field study. Environmental Pollution 94(2): 181-188.
Eberly, S., Swall, J., Holland, D. S., Cox, B. and Baldridge, E. (2004). *Developing spatially interpolated surfaces and estimating uncertainty*. Durham, North Carolina: US Environmental Protection Agency.
Egan, S. K., Bolger, P. M. and Carrington, C. D. (2007). Update of USFDA's Total Diet Study food list and diets. Journal of Exposure Science and Environmental Epidemiology 17(6): 573-582.
Engstrom, A., Michaelsson, K., Suwazono, Y., Wolk, A., Vahter, M. and Akesson, A. (2011). Long-term cadmium exposure and the association with bone mineral density and fractures in a population-based study among women. Journal of Bone and Mineral Research 26(3): 486-495.
EPA. (2006). *Cost of Illness Handbook.* Washington DC: Environmental Protection Agency
Ethier, R. G., Poe, G. L., Schulze, W. D. and Clark, J. (2000). A comparison of hypothetical phone and mail contingent valuation responses for green-pricing electricity programs. Land Economics 76(1): 54-67.
Evangelou, M. W. H., Daghan, H. and Schaeffer, A. (2004). The influence of humic acids on the phytoextraction of cadmium from soil. Chemosphere 57(3): 207-213.
Evans, W. K., Will, B. P., Berthelot, J. M. and Wolfson, M. C. (1996). The economics of lung cancer management in Canada. Lung Cancer 14(1): 19-29.
Ezzati, M. and Lopez, A. D. (2003). Estimates of global mortality attributable to smoking in 2000. The Lancet 362(9387): 847-852.
Farber, S. (1998). Undesirable facilities and property values: A summary of empirical studies. Ecological Economics 24(1): 1-14.
Fiva, J. H. and Kirkeboen, L. J. (2011). Information shocks and the dynamics of the housing market. Scandinavian Journal of Economics 113(3): 525-552.

Fleming, I., Monaghan, P., Fleming, I., Gavin, A. and O'Neill, C. (2008). Factors influencing hospital costs of lung cancer patients in Northern Ireland. European Journal of Health Economics 9(1): 79-86.
Besluit van de Vlaamse Regering van 14 december 2007 houdende vaststelling van het Vlaams reglement betreffende de bodemsanering en de bodembescherming (2007).
Besluit van de Vlaamse Regering van 6 februari 1991 houdende vaststelling van het Vlaams reglement betreffende de milieuvergunning (2012).
Flores, N. E. (2003). Conceptual framework for nonmarket valuation. In Champ, P., Boyle, K. & Brown, T. (eds.), A Primer on Nonmarket Valuation. Dordrecht, The Netherlands: Kluwer Academic Publishers, 27-58.
Freeman, A. M. (1979). Hedonic prices, property values and measuring environmental benefits: Survey of the issues. Scandinavian Journal of Economics 81(2): 154-173.
Freeman, A. M. (1999). The Measurement of Environmental and Resource Values: Theory and Methods. Washington DC: Resources for the future.
Freeman, A. M. (2003). Economic valuation: What and why. In Champ, P., Boyle, K. & Brown, T. (eds.), A Primer on Nonmarket Valuation. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1-26.
Freeman, A. M., Havemen, R. H. and Kneese, A. V. (1973). The Economic of Environmental Policy. New York: John Wiley.
Gabriel, S. E., Tosteson, A. N. A., Leibson, C. L., Crowson, C. S., Pond, G. R., Hammond, C. S. et al. (2002). Direct medical costs attributable to osteoporotic fractures. Osteoporosis International 13(4): 323-330.
Garrod, G. and Willis, K. G. (1999). Economic Valuation of the Environment. Northampton, MA: Edward Elgar Publishing.
Gayer, T., Hamilton, J. T. and Viscusi, W. K. (2000). Private values of risk tradeoffs at superfund sites: Housing market evidence on learning about risk. Review of Economics and Statistics 82(3): 439-451.
Gibbons, S. and Overman, H. G. (2012). Mostly pointless spatial econometrics? Journal of Regional Science 52(2): 172-191.
Gillet, P. and Reginster, J. Y. (1999). Increased number of hip fractures. Lancet 353(9170): 2160-2161.
Gillet, P., Reginsterhaneuse, G. and Reginster, J. Y. (1995). Incidence and direct social costs of hip fracture from 1984 to 1993 in Belgium. Journal of Bone and Mineral Research 10: S463-S463.
Godt, J., Scheidig, F., Grosse-Siestrup, C., Esche, V., Brandenburg, P., Reich, A. et al. (2006). The toxicity of cadmium and resulting hazards for human health. Journal of Occupational Medicine and Toxicology 1(1): 1-6.
Goodwin, P. J. and Shepherd, F. A. (1998). Economic issues in lung cancer: A review. Journal of Clinical Oncology 16(12): 3900-3912.
Goovaerts, P. (1997). Geostatistics for Natural Resource Evaluation. New York: Oxford University Press.
Greenstone, M. and Gallagher, J. (2008). Does hazardous waste matter? Evidence from the housing market and the superfund program. Quarterly Journal of Economics 123(3): 951-1003.

Guignet, D. (2013). What do property values really tell us? A hedonic study of underground storage tanks. Land Economics 89(2): 211-226.
Gyrd-Hansen, D. (2003). Willingness to pay for a QALY. Health Economics 12(12): 1049-1060.
Gyrd-Hansen, D. (2005). Willingness to pay for a QALY - Theoretical and methodological issues. Pharmacoeconomics 23(5): 423-432.
Haentjens, P., Autier, P., Barette, M., Boonen, S. and Belgian Hip Fracture Study, G. (2001). The economic cost of hip fractures among elderly women - A one-year, prospective, observational cohort study with matched-pair analysis. Journal of Bone and Joint Surgery-American Volume 83A(4): 493-500.
Hammer, D., Kayser, A. and Keller, C. (2003). Phytoextraction of Cd and Zn with Salix viminalis in field trials. Soil Use and Management 19(3): 187-192.
Hammitt, J. K. and Haninger, K. (2011). Valuing morbidity risk: Willingness to pay per Quality-Adjusted Life Year *LERNA Working Papers*. Toulouse: LERNA, University of Toulouse.
Hammitt, J. K. and Liu, J. T. (2004). Effects of disease type and latency on the value of mortality risk. Journal of Risk and Uncertainty 28(1): 73-95.
Harrison, G. W. (2006). Experimental evidence on alternative environmental valuation methods. Environmental & Resource Economics 34(1): 125-162.
Hensher, D. and Greene, W. (2003). The mixed logit model: The state of practice. Transportation 30: 133-176.
Hensher, D., Rose, J. and Greene, W. (2005). Applied Choice Analysis: A Primer. Cambridge: Cambridge University Press.
Hensher, D. A. (2010). Hypothetical bias, choice experiments and willingness to pay. Transportation Research Part B-Methodological 44(6): 735-752.
Hiligsmann, M., Bruyere, O., Detilleux, J., Gillet, P., Parmentier, Y., Dercq, J. P. et al. (2010). Trend of hip fracture incidence in Belgium between 2000 and 2007 and future projections. Value in Health 13(7): A303-A304.
Hiligsmann, M., Bruyere, O., Detilleux, J., Gillet, P., Parmentier, Y., Dercq, J. P. et al. (2011a). Incidence of hip fracture in Belgium between 2000 and 2007 and future projections. Osteoporosis International 22: 145-145.
Hiligsmann, M., Bruyere, O., Ethgen, O., Gathon, H. J. and Reginster, J. Y. (2008a). Lifetime absolute risk of hip and other osteoporotic fracture in Belgian women. Bone 43(6): 991-994.
Hiligsmann, M., Ethgen, O., Bruyere, O., Richy, F., Gathon, H. J. and Reginster, J. Y. (2009). Development and validation of a Markov microsimulation model for the economic evaluation of treatments in osteoporosis. Value in Health 12(5): 687-696.
Hiligsmann, M., Ethgen, O., Richy, F. and Reginster, J. Y. (2008b). Utility values associated with osteoporotic fracture: A systematic review of the literature. Calcified Tissue International 82(4): 288-292.

Hiligsmann, M., Gathon, H. J., Bruyere, O., Daubie, M., Parmentier, Y., Dercq, J. P. et al. (2011b). Hospitalisation costs of hip fractures in Belgium. Osteoporosis International 22: 332-332.

Hogervorst, J., Plusquin, M., Vangronsveld, J., Nawrot, T., Cuypers, A., Van Hecke, E. et al. (2007). House dust as possible route of environmental exposure to cadmium and lead in the adult general population. Environmental Research 103(1): 30-37.

Holland, D. S., Johnston, R., Sanchirico, J. N. and Joglekar, D. (2010). Economic Analysis for Ecosystem-based Management: Application to Marine and Coastal Environments. Washington DC: Resources for the future.

Holmes, T. P. and Adamowicz, W. L. (2003). Attribute-based methods. In Champ, P., Boyle, K. & Brown, T. (eds.), A Primer on Nonmarket Valuation. Dordrecht: Kluwer Academic Publishers, 171-220.

Hough, R. L., Breward, N., Young, S. D., Crout, N. M. J., Tye, A. M., Moir, A. M. et al. (2003a). Assessing potential risk of heavy metal exposure from consumption of home-produced vegetables by urban populations. Environmental Health Perspectives 112(2): 215-221.

Hough, R. L., Young, S. D. and Crout, N. M. J. (2003b). Modelling of Cd, Cu, Ni, Pb and Zn uptake, by winter wheat and forage maize, from a sewage disposal farm. Soil Use and Management 19(1): 19-27.

Houlding, S. W. (2000). Practical Geostatistics: Modeling and Spatial Analysis. New York: Springer.

Hoyos, D. (2010). The state of the art of environmental valuation with discrete choice experiments. Ecological Economics 69(8): 1595-1603.

Huang, H. X., Miller, G. Y., Sherrick, B. J. and Gomez, M. I. (2006). Factors influencing Illinois farmland values. American Journal of Agricultural Economics 88(2): 458-470.

Inaba, T., Kobayashi, E., Suwazono, Y., Uetani, M., Oishi, M., Nakagawa, H. et al. (2005). Estimation of cumulative cadmium intake causing Itai-itai disease. Toxicology Letters 159(2): 192-201.

Isaaks, E. H. and Srivastava, R. M. (1989). Applied Geostatistics. New York: Oxford University Press.

Isgin, T. and Forster, D. L. (2006). A hedonic price analysis of farmland option premiums under urban influences. Canadian Journal of Agricultural Economics-Revue Canadienne D Agroeconomie 54(3): 327-340.

James, K. and Meliker, J. (2013). Environmental cadmium exposure and osteoporosis: a review. International Journal of Public Health: 1-9.

Jarup, L. and Akesson, A. (2009). Current status of cadmium as an environmental health problem. Toxicology and Applied Pharmacology 238(3): 201-208.

Jarup, L. and Alfven, T. (2004). Low level cadmium exposure, renal and bone effects - the OSCAR study. Biometals 17(5): 505-509.

Johannesson, M., Blomquist, G., Blumenschein, K., Johansson, P.-o., Liljas, B. and O'Conor, R. (1999). Calibrating hypothetical willingness to pay responses. Journal of Risk and Uncertainty 18(1): 21-32.

Johansson-Stenman, O. and Svedsater, H. (2008). Measuring hypothetical bias in choice experiments: The importance of cognitive consistency. B E Journal of Economic Analysis & Policy 8(1).
Johnston, K., Ver Hoef, J. M., Krivoruchko, K. and Lucas, N. (2001). Using ArcGIS Geostatistical Analyst (1st ed.). Redlands: ESRI.
Kabata-Pendias, A. and Pendias, H. (1992). Trace Elements in Soils and Plants. Boca Raton, FL: CRC Press.
Kamh, M., Wiesler, F., Ulas, A. and Horst, W. J. (2005). Root growth and N-uptake activity of oilseed rape (Brassica napus L.) cultivars differing in nitrogen efficiency. Journal of Plant Nutrition and Soil Science 168(1): 130-137.
Kang, S., Koh, E. S., Vinod, S. K. and Jalaludin, B. (2012). Cost analysis of lung cancer management in South Western Sydney. Journal of Medical Imaging and Radiation Oncology 56(2): 235-241.
Kanis, J. A., Johnell, O., Oden, A., De Laet, C., Oglesby, A. and Jonsson, B. (2002). Intervention thresholds for osteoporosis. Bone 31(1): 26-31.
Keller, C. (2006). Factors limiting efficiency of phytoextraction at multi-metal contaminated sites. In Morel, J.-L., Echevarria, G. & Goncharova, N. (eds.), Phytoremediation of Metal-contaminated Soils. Dordrecht: Springer, 241-266.
Keller, C., Hammer, D., Kayser, A., Richner, W., Brodbeck, M. and Sennhauser, M. (2003). Root development and heavy metal phytoextraction efficiency: Comparison of different plant species in the field. Plant and Soil 249(1): 67-81.
Keller, C., Ludwig, C., Davoli, F. and Wochele, J. (2005). Thermal treatment of metal-enriched biomass prduced from heavy metal phytoextraction. Environmental Science & Technology 39(9): 3359-3367.
Khuder, S. A. (2001). Effect of cigarette smoking on major histological types of lung cancer: a meta-analysis. Lung Cancer 31(2–3): 139-148.
Kiel, K. A. and Williams, M. (2007). The impact of Superfund sites on local property values: Are all sites the same? Journal of Urban Economics 61(1): 170-192.
Kim, C. W., Phipps, T. T. and Anselin, L. (2003). Measuring the benefits of air quality improvement: a spatial hedonic approach. Journal of Environmental Economics and Management 45(1): 24-39.
King, J. T., Tsevat, J., Lave, J. R. and Roberts, M. S. (2005). Willingness to pay for a quality-adjusted life year: Implications for societal health care resource allocation. Medical Decision Making 25(6): 667-677.
Kluve, J. and Schaffner, S. (2008). The value of life in Europe – A meta-analysis. Sozialer Fortschritt 57(10-11): 279-287.
Kobayashi, E., Suwazono, Y., Dochi, M., Honda, R. and Kido, T. (2009). Influence of consumption of cadmium-polluted rice or Jinzu river water on occurrence of renal tubular dysfunction and/or Itai-itai disease. Biological Trace Element Research 127(3): 257-268.
Kochi, I., Hubbell, B. and Kramer, R. (2006). An empirical Bayes approach to combining and comparing estimates of the value of a statistical life for environmental policy analysis. Environmental & Resource Economics 34(3): 385-406.

Kolstad, C. D. (2000). Environmental Economics. Oxford: Oxford University Press.

Kontoleon, A. and Yabe, M. (2003). Assessing the impacts of alternative 'opt-out'formats in choice experiment studies: Consumer preferences for genetically modified content and production information in food. Journal of Agricultural policy and Resources 5(1): 1-43.

Kostov, P. (2010). Do buyers' characteristics and personal relationships affect agricultural land prices? Land Economics 86(1): 48-65.

Koundouri, P. and Pashardes, P. (2003). Hedonic price analysis and selectivity bias - Water salinity and demand for land. Environmental & Resource Economics 26(1): 45-56.

Krupnick, A. (2007). Mortality-risk valuation and age: Stated preference evidence. Review of Environmental Economics and Policy 1(2): 261-282.

Krupnick, A., Alberini, A., Cropper, M., Simon, N., O'Brien, B., Goeree, R. et al. (2002). Age, health and the willingness to pay for mortality risk reductions: A contingent valuation survey of Ontario residents. Journal of Risk and Uncertainty 24(2): 161-186.

Krutilla, J. V. (1967). Conservation reconsidered. American Economic Review 57(4): 777-786.

Kunreuther, H. and Slovic, P. (1996). Science, values, and risk. Annals of the American Academy of Political and Social Science 545: 116-125.

Kuppens, T. (2012). *Techno-economic assessment of fast pyrolysis for the valorisation of short rotation coppice cultivated for phytoextraction.* Ph.D. dissertation, Hasselt University, Diepenbeek.

Kutikova, L., Bowman, L., Chang, S., Long, S. R., Obasaju, C. and Crown, W. H. (2005). The economic burden of lung cancer and the associated costs of treatment failure in the United States. Lung Cancer 50(2): 143-154.

Lancaster, K. J. (1966). New approach to consumer theory. Journal of Political Economy 74(2): 132-157.

Landfeldt, E., Lundkvist, J. and Strom, O. (2011). The societal burden of poor persistence to treatment of osteoporosis in Sweden. Bone 48(2): 380-388.

Landry, C. E. and List, J. A. (2007). Using ex ante approaches to obtain credible signals for value in contingent markets: Evidence from the field. American Journal of Agricultural Economics 89(2): 420-429.

Lauwerys, R., Amery, A., Bernard, A., Bruaux, P., Buchet, J. P., Claeys, F. et al. (1990). Health effects of environmental exposure to cadmium - Objectives, design and organization of the Cadmibel study - A cross-sectional morbidity study carried our in Belgium from 1985 to 1989. Environmental Health Perspectives 87: 283-289.

Leggett, C. G. and Bockstael, N. E. (2000). Evidence of the effects of water quality on residential land prices. Journal of Environmental Economics and Management 39(2): 121-144.

Lindhjem, H., Navrud, S., Braathen, N. A. and Biausque, V. (2011). Valuing mortality risk reductions from environmental, transport, and health policies: A global meta-analysis of stated preference studies. Risk Analysis 31(9): 1381-1407.

Lips, P. and van Schoor, N. M. (2005). Quality of life in patients with osteoporosis. Osteoporosis International 16(5): 447-455.
List, J. A. and Gallet, C. A. (2001). What experimental protocol influence disparities between actual and hypothetical stated values? Environmental & Resource Economics 20(3): 241-254.
Liu, D. W., Wang, Z. M., Zhang, B., Song, K. S., Li, X. Y., Li, J. P. et al. (2006). Spatial distribution of soil organic carbon and analysis of related factors in croplands of the black soil region, Northeast China. Agriculture Ecosystems & Environment 113(1-4): 73-81.
Loomis, J. (2011). What's to know about hypothetical bias in stated preference valuation studies? Journal of Economic Surveys 25(2): 363-370.
Lopez-Encuentra, A. and Bronchogenic Carcinoma, C.-o. (2002). Comorbidity in operable lung cancer - A multicenter descriptive study on 2992 patients. Lung Cancer 35(3): 263-269.
Louviere, J., Hensher, D. and Swait, J. (2000). Stated Choice Methods: Analysis and Applications. Cambridge: Cambridge University Press.
Lusk, J. L. and Schroeder, T. C. (2004). Are choice experiments incentive compatible? A test with quality differentiated beef steaks. American Journal of Agricultural Economics 86(2): 467-482.
Ma, S. and Swinton, S. M. (2011). Valuation of ecosystem services from rural landscapes using agricultural land prices. Ecological Economics 70(9): 1649-1659.
Ma, S. and Swinton, S. M. (2012). Hedonic valuation of farmland using sale prices versus appraised values. Land Economics 88(1): 1-15.
Maddison, D. (2000). A hedonic analysis of agricultural land prices in England and Wales. European Review of Agricultural Economics 27(4): 519-532.
Maddison, D. (2009). A spatio-temporal model of farmland values. Journal of Agricultural Economics 60(1): 171-189.
Marks, R., Allegrante, J. P., MacKenzie, C. R. and Lane, J. M. (2003). Hip fractures among the elderly: causes, consequences and control. Ageing Research Reviews 2(1): 57-93.
Maxted, A. P., Black, C. R., West, H. M., Crout, N. M. J., McGrath, S. P. and Young, S. D. (2006). Phytoextraction of cadmium and zinc by Salix from soil historically amended with sewage sludge. Plant and Soil 290(1-2): 157-172.
McClelland, G. H., Schulze, W. D. and Hurd, B. (1990). The effect of risk beliefs on property values: A case study of a hazardous waste site. Risk Analysis 10(4): 485-497.
McCluskey, J. J. and Rausser, G. C. (2001). Estimation of perceived risk and its effect on property values. Land Economics 77(1): 42-55.
McElroy, J. A., Shafer, M. M., Trentham-Dietz, A., Hampton, J. M. and Newcomb, P. A. (2006). Cadmium exposure and breast cancer risk. Journal of the National Cancer Institute 98(12): 869-873.
McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In Zarembka, P. (ed.), Frontiers in Econometrics. New York: Academic Press, 105-142.

McKie, J., Shrimpton, B., Richardson, J. and Hurworth, R. (2011). The monetary value of a life year: Evidence from a qualitative study of treatment costs. Health Economics 20(8): 945-957.
McMillen, D. P. (2012). Perspectives on spatial econometrics: Linear smoothing with structured models. Journal of Regional Science 52(2): 192-209.
Mendelsohn, R., Nordhaus, W. D. and Shaw, D. (1994). The impact of global warming on agriculture: A Ricardian analysis. American Economic Review 84(4): 753-771.
Menke, A., Muntner, P., Silbergeld, E. K., Platz, E. A. and Guallar, E. (2009). Cadmium levels in urine and mortality among US adults. Environmental Health Perspectives 117(2): 190-196.
Messer, K. D., Schulze, W. D., Hackett, K. F., Cameron, T. A. and McClelland, G. H. (2006). Can stigma explain large property value losses? The psychology and economics of Superfund. Environmental & Resource Economics 33(3): 299-324.
Milligan, M. A., Bohara, A. K. and Pagan, J. A. (2010). Assessing willingness to pay for cancer prevention. International Journal of Health Care Finance & Economics 10(4): 301-314.
Mitani, Y. and Flores, N. E. (2009). Demand revelation, hypothetical bias, and threshold public goods provision. Environmental & Resource Economics 44(2): 231-243.
Mitchell, R. C. and Carson, R. (1989). Using Surveys to Value Public Goods: The Contingent Valuation Method. Washington, DC: Resources for the Future.
Mogas, J., Riera, P. and Bennett, J. (2006). A comparison of contingent valuation and choice modelling with second-order interactions. Journal of Forest Economics 12(1): 5-30.
Molinier, L., Combescure, C., Chouaid, C., Daures, J. P., Housset, B., Fabre, D. et al. (2006). Cost of lung cancer - A methodological review. Pharmacoeconomics 24(7): 651-659.
Morrison, M. and Brown, T. (2009). Testing the effectiveness of certainty scales, cheap talk, and dissonance-minimization in reducing hypothetical bias in contingent valuation studies. Environmental and Resource Economics 44(3): 307-326.
Mrozek, J. R. and Taylor, L. O. (2002). What determines the value of life? A meta-analysis. Journal of Policy Analysis and Management 21(2): 253-270.
Munasinghe, M. (1992). Macroeconomics and the Environment. Cheltenham, UK: Edward Elgar Publishers.
Murphy, J. J., Allen, P. G., Stevens, T. H. and Weatherhead, D. (2005). A meta-analysis of hypothetical bias in stated preference valuation. Environmental & Resource Economics 30(3): 313-325.
Nawrot, T., Plusquin, M., Hogervorst, J., Roels, H. A., Celis, H., Thijs, L. et al. (2006). Environmental exposure to cadmium and risk of cancer: a prospective population-based study. Lancet Oncology 7(2): 119-126.
Nawrot, T., Staessen, J. A., Roels, H., Munters, E., Cuypers, A., Richart, T. et al. (2010). Cadmium exposure in the population: from health risks to strategies of prevention. Biometals 23(5): 769-782.

Nickerson, C. J. and Lynch, L. (2001). The effect of farmland preservation programs on farmland prices. American Journal of Agricultural Economics 83(2): 341-351.

Nordberg, G., Nogawa, K., Nordberg, M. and Friberg, L. (2007). Cadmium. In Nordberg, G., Fowler, B., Nordberg, M. & Friberg, L. (eds.), Handbook on the Toxicology of Metals. Amsterdam: Elsevier, 445-486.

Olsen, J. A., Richardson, J., Dolan, P. and Menzel, P. (2003). The moral relevance of personal characteristics in setting health care priorities. Social Science & Medicine 57(7): 1163-1172.

Olsson, I. M., Bensryd, I., Lundh, T., Ottosson, H., Skerfving, S. and Oskarsson, A. (2002). Cadmium in blood and urine - Impact of sex, age, dietary intake, iron status, and former smoking - Association of renal effects. Environmental Health Perspectives 110(12): 1185-1190.

Ozdemir, S., Johnson, F. R. and Hauber, A. B. (2009). Hypothetical bias, cheap talk, and stated willingness to pay for health care. Journal of Health Economics 28(4): 894-901.

Pace, R. K., Barry, R., Gilley, O. W. and Sirmans, C. F. (2000). A method for spatial-temporal forecasting with an application to real estate prices. International Journal of Forecasting 16(2): 229-246.

Palmquist, R. B. (1989). Land as a differentiated factor of production - A hedonic model and its implications for welfare measurement. Land Economics 65(1): 23-28.

Palmquist, R. B. (1992). Valuing localized externalities. Journal of Urban Economics 31(1): 59-68.

Palmquist, R. B. and Danielson, L. E. (1989). A hedonic study of the effects of erosion control and drainage on farmland values. American Journal of Agricultural Economics 71(1): 55-62.

Park, Y. S., Kim, S. H., Park, S. K., Park, B. J., Kim, Y. T., Lee, S. M. et al. (2010). Costs for 5-year lung cancer survivors in a tertiary care hospital in South Korea. Lung Cancer 68(2): 299-304.

Parkin, D. M. (2001). Global cancer statistics in the year 2000. Lancet Oncology 2(9): 533-543.

Partridge, M. D., Boarnet, M., Brakman, S. and Ottaviano, G. (2012). Introduction: Whither spatial econometrics? Journal of Regional Science 52(2): 167-171.

Patton, M. and McErlean, S. (2003). Spatial effects within the agricultural land market in Northern Ireland. Journal of Agricultural Economics 54(1): 35-54.

Pearce, D. and Turner, R. K. (1990). Economics of Natural Resources and the Environment. Hemel Hempstead: Harvester Weatsheaf.

Peasgood, T., Herrmann, K., Kanis, J. A. and Brazier, J. E. (2009). An updated systematic review of Health State Utility Values for osteoporosis related conditions. Osteoporosis International 20(6): 853-868.

Perin, B., Zaric, B., Jovanovic, S., Matijasevic, J., Stanic, J., Kopitovic, I. et al. (2012). Patient-related independent clinical risk factors for early

complications following Nd: YAG laser resection of lung cancer. Annals of Thoracic Medicine 7(4): 233-237.
Perman, R., Ma, Y. and McGilvray, J. (1996). Natural Resource & Environmental Economics. New York: Longman Publishing.
Perry, G. M. and Robison, L. J. (2001). Evaluating the influence of personal relationships on land sale prices: A case study in Oregon. Land Economics 77(3): 385-398.
Petrie, R. A. and Taylor, L. O. (2007). Estimating the value of water use permits: A hedonic approach applied to farmland in the southeastern united states. Land Economics 83(3): 302-318.
Pinkse, J. and Slade, M. E. (2010). The futute of spatial econometrics. Journal of Regional Science 50(1): 103-117.
Plantinga, A. J., Lubowski, R. N. and Stavins, R. N. (2002). The effects of potential land development on agricultural land prices. Journal of Urban Economics 52(3): 561-581.
Plantinga, A. J. and Miller, D. J. (2001). Agricultural land values and the value of rights to future land development. Land Economics 77(1): 56-67.
Pompen, M., Gok, M., Novak, A., van Wuijtswinkel, R., Biesma, B., Schramel, F. et al. (2009). Direct costs associated with the disease management of patients with unresectable advanced non-small-cell lung cancer in The Netherlands. Lung Cancer 64(1): 110-116.
Poor, P. J., Pessagno, K. L. and Paul, R. W. (2007). Exploring the hedonic value of ambient water quality: A local watershed-based study. Ecological Economics 60(4): 797-806.
Ready, R. C. and Abdalla, C. W. (2005). The amenity and disamenity impacts of agriculture: Estimates from a hedonic pricing model. American Journal of Agricultural Economics 87(2): 314-326.
Ready, R. C., Champ, P. A. and Lawton, J. L. (2010). Using respondent uncertainty to mitigate hypothetical bias in a stated choice experiment. Land Economics 86(2): 363-381.
Reginster, J. Y. and Brandi, M. L. (2010). Fracture prevention in osteoporosis: assessing risks, providing solutions. Osteoporosis International 21: 401-402.
Reginster, J. Y., Gillet, P., Ben Sedrine, W., Brands, G., Ethgen, O., de Froidmont, C. et al. (1999). Direct costs of hip fractures in patients over 60 years of age in Belgium. Pharmacoeconomics 15(5): 507-514.
Reginster, J. Y., Gillet, P. and Gosset, C. (2001). Secular increase in the incidence of hip fractures in Belgium between 1984 and 1996: need for a concerted public health strategy. Bulletin of the World Health Organization 79(10): 942-946.
Reginster, J. Y., Minne, H. W., Sorensen, O. H., Hooper, M., Roux, C., Brandi, M. L. et al. (2000). Randomized trial of the effects of risedronate on vertebral fractures in women with established postmenopausal osteoporosis. Osteoporosis International 11(1): 83-91.

Robinson, L. A. (2007). Policy monitor how US government agencies value mortality risk reductions. Review of Environmental Economics and Policy 1(2): 283-299.
Rosen, S. (1974). Hedonic prices and implicit markets - Product differentiation in pure competition. Journal of Political Economy 82(1): 34-55.
Ruttens, A., Vangronsveld, J., Meiresonne, L., Van Slycken, S., Meers, E., Tack, F. et al. (2008). Sustainable use of metal contaminated agricultural soils: Cultivation of energy crops as an alternative for classical agriculture. In OVAM (Ed.), (pp. 59). Mechelen: (Public Flemish Waste Agency).
Ryan, M. and Watson, V. (2009). Comparing welfare estimates from payment card contingent valuation and discrete choice experiments. Health Economics 18(4): 389-401.
Satarug, S., Garrett, S. H., Sens, M. A. and Sens, D. A. (2010). Cadmium, environmental exposure, and health outcomes. Environmental Health Perspectives 118(2): 182-190.
Scasny, M. and Alberini, A. (2012). Valuation of mortality risk attributable to climate change: Investigating the effect of survey administration modes on a VSL. International Journal of Environmental Research and Public Health 9(12): 4760-4781.
Schlapfer, F. and Fischhoff, B. (2012). Task familiarity and contextual cues predict hypothetical bias in a meta-analysis of stated preference studies. Ecological Economics 81: 44-47.
Schreurs, E., Voets, T. and Thewys, T. (2011). GIS-based assessment of the biomass potential from phytoremediation of contaminated agricultural land in the Campine region in Belgium. Biomass & Bioenergy 35(10): 4469-4480.
Schutte, R., Nawrot, T., Richart, T., Thijs, L., Vanderschueren, D., Kuznetsova, T. et al. (2008). Bone resorption and environmental exposure to cadmium in women: A population study. Environmental Health Perspectives 116(6): 777-783.
Schwartz, G. G., Il'yasova, D. and Ivanova, A. (2003). Urinary cadmium, impaired fasting glucose, and diabetes in the NHANES III. Diabetes Care 26(2): 468-470.
Sigman, H. and Stafford, S. (2011). Management of hazardous waste and contaminated land. In Rausser, G. C., Smith, V. K. & Zilberman, D. (eds.), Annual Review of Resource Economics, Vol 3 (Vol. 3), 255-275.
Smolders, E., Janson, G., Van Laer, L., Ruttens, A., Vangronsveld, J., Römkens, P. F. A. M. et al. (2007). *Teeltadvies voor de landbouw in kader van het interreg project BeNeKempen*. Mechelen (BE): OVAM.
Souhami, R. and Tobias, J. (2003). Cancer and its Management. Oxford: Blackwell Science.
Staessen, J. A., Lauwerys, R. R., Ide, G., Roels, H. A., Vyncke, G. and Amery, A. (1994). Renal function and historical cadmium pollution from zinc smelters. Lancet 343(8912): 1523-1527.
Staessen, J. A., Roels, H., Vangronsveld, J., Clijsters, H., De Schrijver, K., De Temmerman, L. et al. (1995). Preventiemaatregelen vooor

bodemverontreiniging met cadmium. Tijdschrift voor Geneeskunde 51(20): 1387-1395.
Staessen, J. A., Roels, H. A., Emelianov, D., Kuznetsova, T., Thijs, L., Vangronsveld, J. et al. (1999). Environmental exposure to cadmium, forearm bone density, and risk of fractures: prospective population study. Lancet 353(9159): 1140-1144.
Staessen, J. A., Vyncke, G., Lauwerys, R. R., Roels, H. A., Celis, H. G., Claeys, F. et al. (1992). Transfer of cadmium from a sandy acidic soil to man: A population study. Environmental Research 58(1): 25-34.
Sunstein, C. R. (1997). Bad deaths. Journal of Risk and Uncertainty 14(3): 259-282.
Swinton, S. M., Lupi, F., Robertson, G. P. and Hamilton, S. K. (2007). Ecosystem services and agriculture: Cultivating agricultural ecosystems for diverse benefits. Ecological Economics 64(2): 245-252.
Taylor, J. D. (2010). COPD and the response of the lung to tobacco smoke exposure. Pulmonary Pharmacology & Therapeutics 23(5): 376-383.
Taylor, L. O. (2003). The hedonic method. In Champ, P., Boyle, K. & Brown, T. (eds.), A Primer on Nonmarket Valuation. Dordrecht, The Netherlands: Kluwer Academic Publishers, 331-394.
Taylor, L. O., Morrison, M. D. and Boyle, K. J. (2010). Exchange rules and the incentive compatibility of choice experiments. Environmental & Resource Economics 47(2): 197-220.
Taylor, M. R. and Brester, G. W. (2005). Noncash income transfers and agricultural land values. Review of Agricultural Economics 27(4): 526-541.
Tellez-Plaza, M., Guallar, E., Howard, B. V., Umans, J. G., Francesconi, K. A., Goessler, W. et al. (2013). Cadmium exposure and incident cardiovascular disease. Epidemiology 24(3): 421-429.
Tesfaye, A. and Brouwer, R. (2012). Testing participation constraints in contract design for sustainable soil conservation in Ethiopia. Ecological Economics 73: 168-178.
Thewys, T., Witters, N., Van Slycken, S., Ruttens, A., Meers, E., Tack, F. M. G. et al. (2010). Economic viability of phytoremediation of a cadmium contaminated agricultural area using energy maize. Part I: Effect on the farmer's income. International Journal of Phytoremediation 12(7): 650-662.
Thomas, L. D. K., Hodgson, S., Nieuwenhuijsen, M. and Jarup, L. (2009). Early kidney damage in a population exposed to cadmium and other heavy metals. Environmental Health Perspectives 117(2): 181-184.
Thomas, L. D. K., Michaelsson, K., Julin, B., Wolk, A. and Akesson, A. (2011). Dietary cadmium exposure and fracture incidence among men: A population-based prospective cohort study. Journal of Bone and Mineral Research 26(7): 1601-1608.
Thurstone, L. L. (1927). A law of comparative judgment. Psychology Review 34: 273-286.
Tonsor, G. T. and Shupp, R. S. (2011). Cheap talk scripts and online choice experiments: "Looking beyond the mean". American Journal of Agricultural Economics 93(4): 1015-1031.

Tosteson, A. N. A., Gabriel, S. E., Grove, M. R., Moncur, M. M., Kneeland, T. S. and Melton, L. J. (2001). Impact of hip and vertebral fractures on quality-adjusted life years. Osteoporosis International 12(12): 1042-1049.

Train, K. E. (2003). Discrete Choice Methods with Simulation. Cambridge: Cambridge University Press.

Tsoodle, L. J., Golden, B. B. and Featherstone, A. M. (2006). Factors influencing Kansas agricultural farm land values. Land Economics 82(1): 124-139.

Tsuge, T., Kishimoto, A. and Takeuchi, K. (2005). A choice experiment approach to the valuation of mortality. Journal of Risk and Uncertainty 31(1): 73-95.

Ubel, P. A. (1999). How stable are people's preferences for giving priority to severely ill patients? Social Science & Medicine 49(7): 895-903.

Valent, F., Little, D., Bertollini, R., Nemer, L. E., Barbone, F. and Tamburlini, G. (2004). Burden of disease attributable to selected environmental factors and injury among children and adolescents in Europe. Lancet 363(9426): 2032-2039.

Van Cauwenbergh, R., Bosscher, D., Robberecht, H. and Deelstra, H. (2000). Daily dietary cadmium intake in Belgium using duplicate portion sampling. European Food Research and Technology 212(1): 13-16.

Van Houtven, G., Powers, J., Jessup, A. and Yang, J. C. (2006). Valuing avoided morbidity using meta-regression analysis: what can health status measures and QALYs tell us about WTP? Health Economics 15(8): 775-795.

Van Houtven, G., Sullivan, M. B. and Dockins, C. (2008). Cancer premiums and latency effects: A risk tradeoff approach for valuing reductions in fatal cancer risks. Journal of Risk and Uncertainty 36(2): 179-199.

Van Nevel, L., Mertens, J., Oorts, K. and Verheyen, K. (2007). Phytoextraction of metals from soils: How far from practice? Environmental Pollution 150(1): 34-40.

Vangronsveld, J., Herzig, R., Weyens, N., Boulet, J., Adriaensen, K., Ruttens, A. et al. (2009). Phytoremediation of contaminated soils and groundwater: lessons from the field. Environmental Science and Pollution Research 16(7): 765-794.

Vangronsveld, J., Van Assche, F. and Clijsters, H. (1995). Reclamation of a bare industrial area contaminated by non-ferrous metals: in situ metal immobilization and revegetation. Environmental Pollution 87(1): 51-59.

Verougstraete, V., Lison, D. and Hotz, P. (2003). Cadmium, lung and prostate cancer: A systematic review of recent epidemiological data. Journal of Toxicology and Environmental Health-Part B-Critical Reviews 6(3): 227-255.

Viscusi, W. K. (2010). The heterogeneity of the value of statistical life: Introduction and overview. Journal of Risk and Uncertainty 40(1): 1-13.

Viscusi, W. K. and Aldy, J. E. (2003). The value of a statistical life: A critical review of market estimates throughout the world. Journal of Risk and Uncertainty 27(1): 5-76.

Vromman, V., Saegerman, C., Pussemier, L., Huyghebaert, A., De Temmerman, L., Pizzolon, J. C. et al. (2008). Cadmium in the food chain near non-ferrous metal production sites. Food Additives and Contaminants 25(3): 293-301.

Wackernagel, H. (2003). Multivariate Geostatistics: An Introduction with Applications. New York: Springer.

Wang, M., Zou, J. H., Duan, X. C., Jiang, W. S. and Liu, D. H. (2007). Cadmium accumulation and its effects on metal uptake in maize (Zea mays L.). Bioresource Technology 98(1): 82-88.

Warren, J. L., Yabroff, K. R., Meekins, A., Topor, M., Lamont, E. B. and Brown, M. L. (2008). Evaluation of trends in the cost of initial cancer treatment. Journal of the National Cancer Institute 100(12): 888-897.

Welinder, H., Skerfving, S. and Henriksen, O. (1977). Cadmium metabolism in man. British Journal of Industrial Medicine 34(3): 221-228.

Wildemeersch, D. (2008). *Blootstellingsonderzoek Noorderkempen: Geïntegreerd rapport*. Brussels: Administration Public Health Retrieved from http://www.ovam.be/jahia/Jahia/cache/offonce/pid/176?actionReq=actionPubDetail&fileItem=1806.

Witters, N. (2011). *Phytoremediation: An alternative remediation technology and a sustainable marginal land management option.* Ph.D. dissertation, Hasselt University, Diepenbeek.

Witters, N., Slycken, S., Ruttens, A., Adriaensen, K., Meers, E., Meiresonne, L. et al. (2009). Short-rotation coppice of willow for phytoremediation of a metal-contaminated agricultural area: a sustainability assessment. Bioenergy Research 2(3): 144-152.

Wu, J. J. and Lin, H. X. (2010). The effect of the Conservation Reserve Program on land values. Land Economics 86(1): 1-21.

Xu, F., Mittelhammer, R. C. and Barkley, P. W. (1993). Measuring the contributions of site characteristics to the value of agricultural land. Land Economics 69(4): 356-369.

Yabroff, K. R., Lund, J., Kepka, D. and Mariotto, A. (2011). Economic burden of cancer in the United States: Estimates, projections, and future research. Cancer Epidemiology Biomarkers & Prevention 20(10): 2006-2014.

Young, R. A. A. (2005). Determining the Economic Value of Water: Concepts and Methods. Washington DC: Resources for the future.

Young, R. P., Hopkins, R. J., Christmas, T., Black, P. N., Metcalf, P. and Gamble, G. D. (2009). COPD prevalence is increased in lung cancer, independent of age, sex and smoking history. European Respiratory Journal 34(2): 380-386.

Yusuf, A. A. and Resosudarmo, B. P. (2009). Does clean air matter in developing countries' megacities? A hedonic price analysis of the Jakarta housing market, Indonesia. Ecological Economics 68(5): 1398-1407.

Zabel, J. E. and Guignet, D. (2012). A hedonic analysis of the impact of LUST sites on house prices. Resource and Energy Economics 34(4): 549-564.

Zethraeus, N., Stromberg, L., Jonsson, B., Svensson, O. and Ohlen, G. (1997). The cost of a hip fracture - Estimates for 1,709 patients in Sweden. Acta Orthopaedica Scandinavica 68(1): 13-17.
Zhao, F. J., Lombi, E. and McGrath, S. P. (2003). Assessing the potential for zinc and cadmium phytoremediation with the hyperaccumulator Thlaspi caerulescens. Plant and Soil 249(1): 37-43.

Academic bibliography

Contributions in peer reviewed journals

Van Dael, M., Van Passel, S., Pelkmans, L., Guisson, R., Swinnen, G. and Schreurs, E. (2012). Determining potential locations for biomass valorization using a macro screening approach. Biomass & Bioenergy 45: 175-186.

Witters, N., Mendelsohn, R., Van Slycke, S., Weyens, N., Schreurs, E., Meers, E., Tack, F., Carleer, R. and Vangronsveld, J. (2012). Phytoremediation, a sustainable remediation technology? Conclusions from a case study. I: Energy production and carbon dioxide abatement. Biomass & Bioenergy 39: 454-469.

Witters, N., Mendelsohn, R., Van Passel, S., Van Slycken, S., Weyens, N., Schreurs, E., Meers, E., Tack, F., Vanheusden, B. & Vangronsveld, J. (2012). Phytoremediation, a sustainable remediation technology? II: Economic assessment of CO2 abatement through the use of phytoremediation crops for renewable energy production. Biomass & Bioenergy 39: 470-477.

Schreurs, E., Voets, T. and Thewys, T. (2011). GIS-based assessment of the biomass potential from phytoremediation of contaminated agricultural land in the Campine region in Belgium. Biomass & Bioenergy 35(10): 4469-4480.

Manuscripts in preparation for submission in peer reviewed journals

Schreurs, E., Van Passel, S., Peeters, L. and Thewys, T. (xxxx). Analyzing the impact of soil contamination on farmland values: A spatial hedonic approach using quantile regression. (Journal to be selected)

Schreurs, E., Lizin, S., Van Passel, S. and Thewys, T. (xxxx). Investigating farmland purchasing decisions in case of soil contamination: A discrete choice experiment. (Journal to be selected)

Presentations at international conferences and workshops

Peeters, L., Schreurs, E. and Van Passel, S. (2013). Determinants of farmland values in the Campine region (Belgium): Conditional and unconditional quantile regression estimates. 6th Seminar 'Jean Paelinck' of Spatial Econometrics, Madrid (Spain), 18-19 October 2013.

Schreurs, E., Van Passel, S., Peeters, L. and Thewys, T. (2013). Economic impact assessment of farmland contamination: A policy perspective. 8th Conference on Sustainable Development of Energy, Water and Environment Systems (SDEWES), Dubrovnik (Croatia), 22-27 September 2013.

Schreurs, E., Cleemput, I., Nawrot, T. and Thewys, T. (2013). Economic impact assessment of environmental contamination: A damage function approach [Poster presentation]. 8th Conference on Sustainable Development of Energy, Water and Environment Systems (SDEWES), Dubrovnik (Croatia), 22-27 September 2013.

Schreurs, E., Van Passel, S., Peeters, L. and Thewys, T. (2013). Analyzing the impact of soil contamination on farmland values: A spatial hedonic approach using quantile regression. 7th World Conference of the Spatial Econometrics Association (SEA), Washington (DC), 10-12 July 2013.

Schreurs, E., Lizin, S., Van Passel, S. and Thewys, T. (2013). Analyzing the impact of soil contamination on farmland values. Summer School on Choice Experiments in Agricultural and Food Economics, Heverlee (Belgium), 3 July 2013.

Schreurs, E., Lizin, S., Van Passel, S. and Thewys, T. (2013). The impact of soil contamination on farmland values: Revealed preference versus stated preference. 5th European Association of Agricultural Economics (EAAE) PhD Workshop, Leuven (Belgium), 29-31 May 2013.

Schreurs, E., Voets, T. and Thewys, T. (2011). GIS-based assessment of the agricultural land suitable for heavy metal phytoextraction: A case study of the Campine region (Belgium). 8th International Phytotechnologies Society (IPS) conference, Portland (OR), 13-16 September 2011.

Other material

Schreurs, E., Cleemput, I., Thomeer, M. and Thewys, T. (2013). Economic burden of illnesses attributable to environmental pollution in the Campine region. EPS Werkgroep, Brussels (Belgium), 2 July 2013.

Schreurs, E., Hoogmartens, R. and Thewys, T. (2012). The impact of soil contamination on agricultural land prices in the Campine region. 14th PhD symposium of the Belgian Association of Agricultural Economists, Brussels (Belgium), 18 April 2012.

Schreurs, E., Voets, T. and Thewys, T. (2012). Biomassapotentieel van energiegewassen op licht vervuilde landbouwgronden: Gevalstudie uit de Belgische Kempen. Nieuwsbrief Milieu & Economie 26(1): 11-12.