Information Technology and Urbanization Economies

de Vos, Duco

DOI
10.4233/uuid:45114889-f29f-4964-9a98-56e37a736170

Publication date
2020

Document Version
Final published version

Citation (APA)
de Vos, D. (2020). *Information Technology and Urbanization Economies*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:45114889-f29f-4964-9a98-56e37a736170

Information Technology and Urbanization Economies

Duco de Vos

Delft University of Technology

Information Technology and Urbanization Economies

Proefschrift

ter verkrijging van de graad van doctor

aan de Technische Universiteit Delft,

op gezag van de Rector Magnificus, prof.dr.ir. T.H.J.J. van der Hagen

voorzitter van het College voor Promoties

in het openbaar te verdedigen op

dinsdag 17 maart 2020 om 12:30

door

Duco Willem DE VOS

Master of Science in Spatial, Transport and Environmental Economics

Vrije Universiteit, Amsterdam

geboren te Dordrecht

This dissertation has been approved by the promotors

Composition of the doctoral committee:

Rector Magnificus	Chairperson
Dr. E.J. Meijers	Delft University of Technology, promotor
Prof. dr. M. van Ham	Delft University of Technology, promotor

Independent members:

Prof. dr. Willem Korthals Altes	Delft University of Technology
Prof. dr. Caspar Chorus	Delft University of Technology
Prof. dr. Eveline van Leeuwen	Wageningen University
Prof. dr. Jos van Ommeren	Vrije Universiteit Amsterdam
Dr. Patricia Melo	University of Lisbon

The research in this thesis was done at the Faculty of Architecture and the Built Environment, Delft University of Technology. The project was funded by the Netherlands Organization for Scientific Research (NWO), grant number 452-14-004. Data was gathered from The Netherlands Institute for Social Research (SCP), CentERdata (Tilburg University), the Swedish Telecom Authority (PTS), the ASTRID register database from Umeå University in Sweden, Netherlands Statistics (CBS), HorecaDNA, Iens.nl, and the Netherlands Institute for Transport Policy Analysis (KiM).

Printed in the Netherlands

Contents

Preface

Before you, you have a PhD thesis that is by and large the product of luck. I was lucky to be offered a researcher position at VU Amsterdam after finishing my master. And lucky to find my PhD position at TU Delft under the supervision of Evert Meijers and Maarten van Ham. Lucky to have been able to choose my own path within this PhD. And lucky to have enjoyed the trips to conferences, my 6 month stay in Italy, and the relaxed atmosphere that surrounds academic work in general (for some).

I want to thank Evert for offering me this position, and for being there over the years not only as a boss but also as a kind of work-father and friend. I thank Maarten for the supervision, the help, and keeping in mind ‘the bigger picture’. It was nice to have you both as a team of promotors.

I am grateful to Jos for bringing up the possibility of a PhD during my master, and for helping me in my first steps in academia. I thank Antoine for being a friend from the start of our PhDs. I thank the people from the former SRO section, Arie, Erik, Kees, Rob, Jan-Jacob, Marjolein, Dena, Jaume, Rodrigo and Paul and André, for really making me feel at home when I came to OTB. I enjoyed that we always had lunch together with our group. And I thank the colleagues from the new Urban Studies section for a pleasurable last couple of months. Good luck!

Summary

Introduction

It is increasingly recognized that urbanization economies – the benefits of living in cities – can be generated by proximity to large cities (OECD 2015). Several scholars have put forward that places near other large cities are increasingly able to 'borrow size' of their neighbours to generate these economies, and that this may explain recent patterns of (economic) growth across European cities, whereby the largest cities have not necessarily had the highest growth rates (Dijkstra et al. 2013, Burger and Meijers 2016). These studies suggest that Europe's unique polycentric urban structure increasingly allows urbanization benefits to be generated by proximity to large agglomerations, due to improvements in physical and digital infrastructure.

Indeed, there is plenty of evidence that increasing the effective density of regions by improving physical transportation infrastructure leads to higher levels of urbanization economies (Graham 2019). For improvements in digital infrastructure however, such evidence is missing. In this thesis I attempt to fill this gap, and contribute to the discussion of whether information technology enables places in proximity of large cities to 'borrow' urbanization economies?

To understand the relation between IT and borrowed size it is important to have a plausible theoretical mechanism. In the introduction of this thesis I have put forward such a theoretical link, that is based on the relation between ubiquitous online information and travel behaviour. In short, I expect that in some cases IT may complement longer distance travel for jobs and local products, which means that in these markets urban scale economies (including better matching and wider product variety) are generated and enjoyed across a greater geographical scale. Based on this theoretical link, I devised two research questions.

1. To what extent does information technology increase the geographical extent of local labour and product markets?

2. To what extent has the advent of information technology led to better local (labour or product) market outcomes in places in proximity of large cities?

This thesis is divided into two parts, whereby the Chapters in Part 1 answer the research questions for local labour markets, and the Chapters in Part 2 answer the research questions for local product markets.

Results for local labour markets

Chapters 2 and 3 aimed to answer whether information technology has extended the geographical range of local labour markets (research question 1) by analysing the relationship between home-based teleworking and commuting distance. Chapter 2 named *Working from home and the willingness to accept a longer commute* addressed the empirical issues that have plagued research on the relationship between teleworking and the length of commutes. These issues relate to (1) omitted variables, (2) reverse causality, and (3) selection based on unobserved preferences. In this Chapter I use a panel data set that spans more than a decade, and I estimate the effect of telework on commute lengths with fixed effects models. This approach primarily addresses the issue of selection based on preferences, but also makes headway in addressing omitted variables and reverse causality by including detailed job characteristics in the estimation, and by analysing the effect of changes in telework on changes in commuting.

The first result of Chapter 2 is that it is important to control for unobserved preferences. Using an existing approach to measure commuting preferences by the Marginal Cost of Commuting (Van Ommeren and Fosgerau 2009) I find that teleworkers have a stronger dislike for commuting. But whereas this should theoretically bias observational results downwards, I find that adding fixed effects leads to a lower estimate. This suggests that another type of selection, based on residential preferences, may induce a stronger bias. The second and main result of this Chapter is that information technology has indeed increased the geographical scale of labour markets, as teleworking has a significant positive effect on commute lengths. Working from home allows people to accept 5% longer commuting times on average, and every 8 weekly hours working from home are associated with 3.5% longer commuting times.

The paper in Chapter 3 named Working from home and commuting: Heterogeneity over time, space, and occupations delved deeper into the relationship between telework and commuting. Specifically this Chapter reproduced the research of Chapter 2 with a different, more recent data set, and brought nuance to the results by showing the heterogeneity of the effect of teleworking. The key finding of this Chapter is that I can reproduce the positive relation between working from home and commuting found in Chapter 2 with another data set. I found a stronger effect than in Chapter 2: respondents who start teleworking increase their commutes by 12 percent. It should however be noted that the definition of teleworkers slightly differs between the two chapters. Across the study period of 2008-2018 we find that the effect remained remarkably similar.

Chapter 4 consists of a paper called *Does broadband internet allow cities to 'borrow size'? Evidence from the Swedish labour market*. This Chapter started from a discussion of recent literature that suggests that recent growth dynamics of European cities – whereby the largest cities did not necessarily have the highest economic growth – may be explained by increased possibilities for 'borrowing size'. In other words, that proximity to large cities may have become a better substitute for agglomeration in recent years. In this literature it is often noted that this may have to do with the increased availability of broadband internet (Dijkstra et al. 2013, Hesse 2016). I made use of a unique data set from Sweden that traces the geographical availability of different internet technologies at the level of 250m grid cells, between 2007 and 2015. I connected these data to the Swedish population register to estimate the effects of broadband availability on labour market outcomes.

Chapter 4 at best provides suggestive evidence that broadband internet has changed urbanization economies (i.e. matching) in local labour markets. Outside of cities I find no relationship between residential broadband and the prospects of employment. There is however a weak but significant relation between place-level broadband penetration and employment, in places between 20 and 50 kilometres from cities. While this result suggests that broadband penetration makes places in proximity of cities more attractive for firms and (employed) people to locate, the finding that it is not being connected, but rather local connectivity that counts poses a puzzle.

Results for local product markets

Part 2 of this thesis dealt with the relation between IT and borrowed urbanization economies in local product markets, as opposed to local labour markets. The paper in Chapter 5 titled Information technology and local product variety: Substitution, complementarity and spillovers used cuisine variety in local restaurant markets as indicator of consumption side agglomeration economies, following a longer tradition (Glaeser et al. 2001, Berry and Waldfogel 2010, Schiff 2015). I measured the usage and penetration of internet in local restaurant markets by the share of restaurants reviewed on Iens.nl, which at the time was the most popular culinary review website in the Netherlands. I found a strong and significant relationship between the share of reviewed restaurants and cuisine variety in Dutch places. The results confirm the finding of earlier studies that the effects of IT on product variety increase with city size. It is furthermore shown that the relationship between IT and variety decays with distance from large population centres. This suggests that information technology allows places near larger places to better capitalize on the size of their neighbours, and sustain higher levels of cuisine variety. Information technology may increase the willingness to travel for restaurants, and make these places more accessible for residents from neighbouring places.

Chapter 6 titled Information technology and the geographical extent of local product markets aimed to answer the first research question for local product markets. In this Chapter I first identified a lacuna in current research about the effect of information technology on travel behaviour: The literature has primarily dealt with trip purposes with a clear online counterpart, and it has ignored trips for purposes without a clear online pendant. At the same time, these activities make up an increasingly important share of the economic landscape of cities. I used trips related to services and personal care as a proxy for such activities, and I first showed that there has been a significant increase in travel distance for these purposes between 2000 and 2015, with about 17 percent. I then proceeded to test whether the rise of information technology can explain this trend.

Making use of a Dutch panel survey specifically designed to assess the impact of information technology on travel behaviour, I used fixed-effects models to estimate the effect of changes in engaging with local online content on changes in travel for services and personal care purposes. With slight variability depending on the measure of engaging with online content, I found a 40 percent increase in travel distance for service trips. The effect seems mainly driven by non-urban residents. I concluded that the rise of information technology is a credible explanation for the observed increase in travel distance for service trips over the last decades.

Conclusion

The aim of this thesis was to contribute to the discussion of whether information technology enables places in proximity of large cities to 'borrow' urbanization economies*?* Taken together, contributions have been made to this discussion by (1) proposing a theoretical link based on travel behaviour, (2) assessing the link between information technology and travel behaviour using panel data, (3) assessing the link between internet infrastructure and labour market outcomes across space, and (4) investigating the effects of information technology on local product variety using a national dataset.

The results pertaining to research question 1 suggest that there may be significant effects, because information technology has important impacts on travel behaviour in local markets. I only found suggestive evidence concerning the link between information technology and market outcomes. This evidence is indicative of a small but positive and significant effect of information technology on borrowed urbanization economies. Considering the strong relation between information technology and travel behaviour, it should not be ruled out that in the long run, information technology may have stronger effects on the geographical scale of urbanization economies.

Policy implications

The results of this thesis have important general implications for spatial policies. First, the results confirm the findings of existing studies, that information technology complements urbanization economies (Goldfarb and Tucker 2017). But policymakers should be aware that information technology likely complements 'proximity economies' as well. Second, when considering to subsidize broadband initiatives, governments should not overstate the expected local economic effects. And finally, information technology allows people to cover more ground rather than less. This should be considered when implementing policies that aim to substitute travel for online activities. The analyses in this thesis do not shed definitive light on the effects of teleworking on transport emissions. Still, the results give ample reason to not overstate the effects of teleworking as a strategy to lower harmful emissions related to commuting.

Summary in Dutch

Inleiding

Steeds vaker wordt erkend dat urbanisatievoordelen – de voordelen van wonen in steden – kunnen worden gegenereerd door de nabijheid van grote steden (OESO 2015). Verschillende wetenschappers hebben aangevoerd dat plaatsen in de buurt van andere grote steden in toenemende mate in staat zijn om 'grootte' van hun buren te 'lenen' om deze voordelen te genereren. Dit zou recente patronen van (economische) groei onder Europese steden kunnen verklaren, waarbij de grootste steden niet noodzakelijkerwijs de hoogste groeicijfers hadden (Dijkstra et al. 2013, Burger en Meijers 2016). Verder wordt gesuggereerd dat Europa's unieke polycentrische stedelijke structuur het in toenemende mate mogelijk maakt om urbanisatievoordelen te genereren door de nabijheid van grote agglomeraties, dankzij verbeteringen in fysieke en digitale infrastructuur.

Er is inderdaad ruim bewijs dat het verhogen van de effectieve dichtheid van regio's door verbetering van de fysieke transportinfrastructuur leidt tot meer urbanisatievoordelen (Graham 2019). Voor verbeteringen in de digitale infrastructuur ontbreekt echter dergelijk bewijs. In dit proefschrift probeer ik deze leemte te vullen en bij te dragen aan de vraag of informatietechnologie plaatsen in de nabijheid van grote steden in staat stelt urbanisatievoordelen te 'lenen'?

Om de relatie tussen IT en *borrowed size* (geleende omvang) te begrijpen, is het belangrijk om een plausibel theoretisch mechanisme te hebben. In de inleiding van dit proefschrift heb ik een dergelijke theoretische link geïntroduceerd, die gebaseerd is op de relatie tussen alomtegenwoordige online informatie en reisgedrag. In het kort verwacht ik dat IT in sommige gevallen een langere reisafstand voor banen en lokale producten mogelijk maakt. Dit betekent dat urbanisatievoordelen die zich voordoen in deze markten (zoals betere *matching* en een grote variatie in aanbod) kunnen worden gegenereerd en genoten in een groter gebied, met een grotere schaal. Op basis van deze theoretische link bedacht ik twee onderzoeksvragen.

1. In hoeverre vergroot informatietechnologie de geografische schaal van lokale arbeidsmarkten en productmarkten?

2. In hoeverre heeft de komst van informatietechnologie geleid tot betere lokale (arbeids- of product) marktresultaten in plaatsen in de nabijheid van grote steden?

Dit proefschrift is verdeeld in twee delen, waarbij de hoofdstukken in deel 1 de onderzoeksvragen voor lokale arbeidsmarkten beantwoorden, en de hoofdstukken in deel 2 de onderzoeksvragen voor lokale productmarkten beantwoorden.

Resultaten voor lokale arbeidsmarkten

De hoofdstukken 2 en 3 waren bedoeld om te beantwoorden of informatietechnologie de geografische schaal van lokale arbeidsmarkten heeft vergroot (onderzoeksvraag 1) door de relatie tussen telewerken en woon-werkverkeer te analyseren. Hoofdstuk 2 genaamd *Working from home and the willingness to accept a longer commute*, besprak de empirische kwesties die spelen in onderzoek naar de relatie tussen telewerken en de duur van het woon-werkverkeer. Deze kwesties hebben betrekking op (1) weggelaten variabelen, (2) omgekeerde causaliteit en (3) selectie op basis van niet-waargenomen voorkeuren. In dit hoofdstuk gebruik ik panel data die meer dan een decennium beslaat, en ik schat het effect van telewerken op woon-werkverkeer met zgn. *fixed effects* modellen. Deze benadering richt zich in de eerste plaats op het probleem van selectie op basis van voorkeuren, maar maakt ook vooruitgang bij het aanpakken van weggelaten variabelen en omgekeerde causaliteit door te corrigeren voor gedetailleerde baan kenmerken en door het effect van veranderingen in telewerken op veranderingen in het woon-werkverkeer te analyseren.

Het eerste resultaat van hoofdstuk 2 is dat het belangrijk is om te controleren voor niet-waargenomen voorkeuren. Met behulp van een bestaande benadering om woon-werkvoorkeuren te meten aan de hand van de marginale kosten van woon-werkverkeer (Van Ommeren en Fosgerau 2009) vind ik dat telewerkers een grotere afkeer hebben van woon-werkverkeer. Maar hoewel dit theoretisch de geschatte effecten naar beneden toe zou moeten vertekenen, vind ik dat het controleren voor niet-waargenomen voorkeuren tot een lagere schatting leidt. Dit suggereert dat een ander type selectie, gebaseerd op woonvoorkeuren, een sterkere vertekening veroorzaakt. Het tweede en belangrijkste resultaat van dit hoofdstuk is dat informatietechnologie inderdaad de geografische schaal van arbeidsmarkten heeft vergroot, aangezien telewerken een aanzienlijk positief effect heeft op reistijden. Werken vanuit huis stelt mensen in staat gemiddeld 5% langere reistijden te accepteren, en elke 8 wekelijkse thuiswerk uren gaan gepaard met gemiddeld 3,5% langere reistijden.

Het artikel in hoofdstuk 3 genaamd *Working from home and commuting: Heterogeneity over time, space, and occupations* ging dieper in op de relatie tussen telewerken en woon-werkverkeer. In dit hoofdstuk reproduceerde ik het onderzoek uit hoofdstuk 2 met een andere, recentere dataset. Ik nuanceer eerdere bevindingen door aan te tonen dat het effect van telewerken heterogeen is. De belangrijkste bevinding van dit hoofdstuk is dat de positieve relatie tussen thuiswerken en woon-werkverkeer uit hoofdstuk 2 kan worden gereproduceerd met een andere dataset. Ik vond een sterker effect dan in hoofdstuk 2: respondenten die beginnen met telewerken verlengen hun woon-werkverkeer met gemiddeld 12 procent. Er moet echter worden opgemerkt dat de definitie van telewerkers enigszins verschilt tussen de twee hoofdstukken. Gedurende de studieperiode van 2008-2018 veranderde het effect nauwelijks.

Hoofdstuk 4 bestaat uit een artikel met de titel *Does broadband internet allow cities to 'borrow size'? Evidence from the Swedish labour market*. Dit hoofdstuk startte met een bespreking van recente literatuur die suggereert dat de recente groeidynamiek van Europese steden - waarbij de grootste steden niet noodzakelijkerwijs de hoogste economische groei hadden - kan worden verklaard door de toegenomen mogelijkheden voor '*borrowed size*'. Met andere woorden, die nabijheid van grote steden is de afgelopen jaren wellicht een beter alternatief voor agglomeratie geworden. In deze literatuur wordt vaak opgemerkt dat dit mogelijk te maken heeft met de toegenomen beschikbaarheid van breedbandinternet (Dijkstra et al. 2013, Hessen 2016). Ik heb gebruik gemaakt van een unieke dataset uit Zweden die de geografische beschikbaarheid van verschillende internettechnologieën op het niveau van 250m rastercellen tussen 2007 en 2015 traceert. Ik heb deze gegevens verbonden met het Zweedse bevolkingsregister om de effecten van de beschikbaarheid van breedband op arbeidsmarkt uitkomsten te schatten.

Hoofdstuk 4 biedt op zijn best suggestief bewijs dat breedbandinternet urbanisatievoordelen (d.w.z. matching) op lokale arbeidsmarkten heeft veranderd. Buiten steden vind ik geen verband tussen breedband aansluitingen en de vooruitzichten op werk. Er is echter een zwakke maar significante relatie tussen breedbandpenetratie op het niveau van de woonplaats en werkgelegenheid, in regio's tussen 20 en 50 kilometer van steden. Hoewel dit resultaat suggereert dat breedbandpenetratie plaatsen in de nabijheid van steden aantrekkelijker maakt voor bedrijven en (werkende) mensen, is het onduidelijk waarom het effect wordt gedreven door 'lokale connectiviteit' en niet door individuele breedbandaansluitingen.

Resultaten voor lokale productmarkten

Deel 2 van dit proefschrift ging over de relatie tussen IT en 'geleende' urbanisatievoordelen op lokale productmarkten (tegenover lokale arbeidsmarkten in Deel 1). In het artikel in hoofdstuk 5 getiteld *Information technology and local product variety: Substitution, complementarity and spillovers* gebruikte ik variatie van keukentypes in lokale restaurantmarkten als indicator van *consumptieve urbanisatievoordelen*, in het spoor van een langere traditie (Glaeser et al. 2001, Berry en Waldfogel 2010, Schiff 2015). Ik heb het gebruik en de penetratie van internet in lokale restaurantmarkten gemeten aan de hand van het aantal restaurants dat op Iens.nl is beoordeeld, destijds de populairste culinaire review website in Nederland. Ik vond een sterke en significante relatie tussen het aandeel beoordeelde restaurants en variatie in keukens op plaatsniveau. De resultaten bevestigen eerdere bevindingen die aantonen dat de effecten van IT op productvariëteit toenemen met de grootte van een plaats. Verder wordt aangetoond dat de relatie tussen IT en variëteit vervalt met de afstand tot grote bevolkingscentra. Dit suggereert dat informatietechnologie plaatsen in de buurt van grotere plaatsen in staat stelt om beter te profiteren van de grootte van hun buren en om een hoger niveau van keukenvariëteit in stand te houden. Informatietechnologie kan de reisbereidheid voor restaurants vergroten en deze plaatsen toegankelijker maken voor bewoners uit naburige plaatsen.

Hoofdstuk 6 getiteld *Information technology and the geographical extent of local product markets* had als doel om de eerste onderzoeksvraag voor lokale productmarkten te beantwoorden. In dit hoofdstuk heb ik allereerst een lacune geïdentificeerd in huidig onderzoek naar het effect van informatietechnologie op reisgedrag: de literatuur heeft voornamelijk betrekking gehad op reisdoelen met een duidelijke online tegenhanger, en doeleinden zonder een duidelijke online tegenhanger zijn genegeerd. Tegelijkertijd vormen deze activiteiten een steeds belangrijker onderdeel van het economische landschap van steden. Ik gebruikte reizen met betrekking tot diensten en persoonlijke verzorging als een proxy voor dergelijke activiteiten, en ik toonde eerst aan dat de reisafstand voor deze doeleinden tussen 2000 en 2015 aanzienlijk was toegenomen, met ongeveer 17 procent. Vervolgens onderzocht ik of de opkomst van informatietechnologie deze trend kan verklaren.

In hoofdstuk 6 heb ik gebruik gemaakt van een Nederlands panelonderzoek dat specifiek is ontworpen om de impact van informatietechnologie op reisgedrag te beoordelen. Ik gebruikte modellen met zgn. *fixed effects* om het effect te schatten van veranderingen in het gebruik van internet voor lokale doeleinden op veranderingen in reizen voor diensten en persoonlijke verzorging. Met een kleine variabiliteit, afhankelijk van de mate van interactie met online inhoud, vond ik dat het gebruik van internet voor lokale doeleinden leidt tot een 40 procent grotere reisafstand voor de onderzochte doelen. Het effect lijkt vooral te worden veroorzaakt door niet-stedelijke bewoners. Ik concludeerde dat de opkomst van informatietechnologie een geloofwaardige verklaring is voor de waargenomen toename van de reisafstand voor dienstreizen in de afgelopen decennia.

Conclusie

Het doel van dit proefschrift was om bij te dragen aan de discussie of informatietechnologie plaatsen in de nabijheid van grote steden in staat stelt urbanisatievoordelen te 'lenen'? Deze bijdragen zijn geleverd door: (1) het introduceren van een theoretisch mechanisme gebaseerd op reisgedrag; (2) het onderzoeken van de relatie tussen informatietechnologie en reisgedrag met behulp van panel data; (3) het onderzoeken van de relatie tussen internet infrastructuur en uitkomsten op de arbeidsmarkt met behulp van panel data; en (4) het onderzoeken van de effecten van informatietechnologie op lokale productverscheidenheid.

De resultaten met betrekking tot onderzoeksvraag 1 suggereren dat er aanzienlijke effecten kunnen zijn, omdat informatietechnologie belangrijke gevolgen heeft voor reisgedrag in lokale markten. Ik vond echter enkel suggestief bewijs aangaande het verband tussen informatietechnologie en marktresultaten. Dit bewijs wijst op een klein maar positief effect van informatietechnologie op geleende urbanisatievoordelen. Gezien de sterke relatie tussen informatietechnologie en reisgedrag kan niet worden uitgesloten dat informatietechnologie op de lange termijn sterkere effecten zal hebben op de geografische schaal van urbanisatievoordelen.

Beleidsimplicaties

De resultaten van dit proefschrift hebben belangrijke implicaties voor ruimtelijk beleid. Ten eerste bevestigen de resultaten de bevindingen van bestaande studies, dat informatietechnologie een complement vormt voor urbanisatievoordelen (Goldfarb ant Tucker 2017). Maar beleidsmakers moeten zich ervan bewust zijn dat informatietechnologie waarschijnlijk ook 'nabijheidsvoordelen' complementeert. Ten tweede moeten overheden, wanneer zij overwegen om breedbandinitiatieven te subsidiëren, de verwachte lokale economische effecten niet overschatten. Ten slotte stelt informatietechnologie mensen in staat om een groter geografisch gebied te bestrijken. Hiermee moet rekening worden gehouden bij het implementeren van beleid dat erop gericht is reizen te vervangen door online activiteiten. De analyses in dit proefschrift geven geen definitief antwoord op de vraag in hoeverre telewerken leidt tot minder schadelijke uitstoot door woon-werkverkeer. Desondanks geven de resultaten genoeg reden om de effecten van telewerken als strategie om deze uitstoot te verminderen niet te overschatten.

1. Introduction

1.1 Background of study

The economic performance of places depends on the number – and density – of inhabitants. Urban economic research of the last three decades has convincingly shown that larger cities are generally more productive (Melo et al. 2009), and they offer a wider variety of products and local amenities (Handbury and Weinstein 2015, Schiff 2015). These improved local economic outcomes that are due to population size are a type of scale economies known as urbanization economies. These economies arise because of improved sharing, matching, and learning (Duranton and Puga 2004): Larger cities allow for more efficient sharing of goods and services that rely on scale economies; larger markets facilitate more efficient and better matching between demand and supply (employers and employees, buyers and sellers, etc.); and a large concentration of (different types of) people is conducive to knowledge sharing and learning, that may lead to innovation and higher productivity. Urbanization economies provide a plausible (partial) explanation of why large cities are generally characterized by higher levels of productivity (indicated by high wages), and higher land prices.

A recent report by the OECD (2015) draws attention to the notion that, especially in European countries, proximity to large cities may generate urbanization economies in places "beyond city limits" (p.99). In the report this idea is supported by the finding that regions that include agglomerations with more than half a million people have had higher per capita GDP growth rates between 1995 and 2010, and the finding that halving travel times to urban agglomerations is associated with 0.2-0.4 percentage points higher growth rates (Ahrend and Schumann 2014). The notion that population – or size – can be 'borrowed' from nearby agglomerations is also present in a number of other studies that try to explain recent patterns of (economic) growth across European cities, whereby the largest cities have not necessarily had the highest growth rates (Dijkstra et al. 2013, Camagni et al. 2014, Burger and Meijers 2016, Ahrend et al. 2017). These studies suggest that Europe's unique polycentric urban structure increasingly allows urbanization benefits to be generated by proximity to large agglomerations.

The literature that highlights the *increased* substitutability between agglomeration and proximity is however unclear about the exact causes of this trend. One relatively straightforward cause may be improvements in transportation technologies and transport infrastructure, that decrease travel times between cities and thereby increase the effective density of an urban area or region. A host of studies indeed confirms that there is a strong relationship between *effective urban density* (based on travel times) and labour productivity (see for instance Graham 2008, Melo et al. 2016, Rice et al 2006).

Increased digital connectivity is also often mentioned as a cause of increased possibilities for *borrowed size*. For instance Dijkstra et al. (2013) argue that “improvements in the access to services, including broadband, outside large cities may have facilitated the higher growth rates of smaller centres and rural regions” (p. 334) and Hesse (2016) notes that “borrowing size or significance no longer relies on physical proximity between the cities, but on embeddedness in overarching networks between and within polycentric city-regions, via corporate relations, market pervasion and, last but not least, information and communication networks” (p. 617). Compelling as they may seem, the suggestions that information technology enables places in proximity of urban agglomerations to capitalize on the size of their neighbours still lack clear theoretical underpinnings, and support by empirical findings.

To understand the relation between IT and borrowed size it is important to have a plausible theoretical mechanism that links information technology to spatial differences in performance of places, backed up by empirical evidence. The central aim of this thesis is therefore to contribute to the discussion of whether information technology enables places in proximity of large cities to ‘borrow’ urbanization economies.

The remainder of this introductory chapter is structured as follows. In Section 1.2 I introduce a theoretical link between information technology and borrowed urbanization economies, based on the interplay between IT and travel behaviour. This section ends with research questions that contribute to fulfilling the aim of this thesis. In Section 1.3 I discuss the empirical methods I will use to answer the research questions, and the types of data that will be used. Section 1.4 gives an overview of the remaining chapters and their contents, and it links each chapter to the research questions that it aims to answer.

1.2 Theory and research questions

As mentioned before, a straightforward cause of increased substitutability between proximity and agglomeration may be improvements in infrastructure that decrease transportation costs between places. Better and faster road or rail connections between places decrease travel times, and thus decrease the costs associated with travelling between these places. The transportation literature offers a considerable body of evidence that in some respects, information technology affects transportation costs between places in a similar way. In this subsection I review the literature on IT and travel behaviour and highlight the most important ways in which IT affects the costs of travel, and as a consequence the geographical extent of local markets. I then propose that this mechanism may link information technology to the increased possibilities for ‘borrowed size’ suggested in the literature, and I delve deeper in the expected effects for local labour markets and local product markets. Using readily available data I will show that the growth in information technology in two specific local markets, over the last two decades, has coincided with an increased geographical extent of these markets as measured by travel distance. This section ends with the research questions that will be addressed in this thesis.

IT and travel behaviour – Salomon (1985) was one of the first to give an account of the type of relations between information technology and travel behaviour. He distinguished 4 types of interactions: IT may complement, substitute, modify, or neutrally affect travel. Complementarity occurs if information technology leads to more travel. As an example of this, Salomon (1985) mentions that increased social and economic relations between places due to information technology may go hand in hand with increased travel. Substitution occurs if information technology decreases the need for travel. Salomon (1985) mentions telecommuting (working from home) decreasing the need for commuting, *teletext* decreasing the need to travel for information, and teleconferencing decreasing the need for business travel. Modification occurs if certain aspects of travel that would have occurred anyway are affected. Examples include improved coordination, real time information, and the ability to use travel time productively (Moktharian and Tal 2013). Finally there are numerous types of trips that may not be affected by information technology at all. Salomon (1985) notes that this categorization ignores long-run land-use effects of altered travel behaviour.

With all these potential interactions between IT and travel, the focus in transportation research shifted to estimating the net effect of IT on travel, for specific types of travel and for travel as a whole. In a conceptual review, Mokhtarian et al. (2004) suggests that the general opinion among transport researchers is that overall, IT complements travel. A literature review by Andreev et al. (2010) summarizes results from about 100 articles, and notes that substitution has been the most prevalent impact of telework, and complementarity has been the most prevalent impact of online shopping and -leisure.

Given that the focus of transportation research has traditionally been on the effects of IT on transport emissions, this literature has focused on the effects of IT on total travel. Substitution effects found in this literature may either mean that trip distances have decreased, or that the number of trips has decreased as a result of IT, and vice versa for complementarity. The relationship between IT and agglomeration economies does not depend on IT’s effects on total travel, but rather on IT’s effects on the demand for proximity (Gaspar and Glaeser 1999). In assessing the effects of IT on agglomeration economies it is thus vital to distinguish between the number of trips (the extensive margin of travel) and trip length or distance (the intensive margin). A positive effect on the number of trips would imply an increase in demand for physical travel and thus for proximity. But a positive effect on trip distance would imply decreased costs of overcoming distance, and thus lower demand for proximity.

Especially positive effects of information technology on travel distances may provide a viable explanation for some of the 'borrowed size' patterns that have been documented. A higher willingness to travel for local goods (including jobs) increases the geographical scale of local markets. The introduction of information technology leads to a better incorporation of places on or near the edges of large local markets (i.e. cities). As these places then converge – in terms of market outcomes – with the larger local market, one would expect the higher growth rates that have been witnessed for small and medium sized cities in proximity of larger cities.

Implications for local markets - In this thesis I focus on the effects of information technology on local labour- and local product markets. Outcomes in these markets are key indicators of urbanization economies from the perspective of people (workers, consumers), as opposed to total factor productivity which is a key indicator in the literature on *agglomeration economies* (i.e. benefits of concentration for firms). For both labour and product markets I will assess whether information technology leads to a greater geographical extent, and whether the penetration of information technology is associated with better market outcomes.

In local labour markets commuting and commuting tolerance determines the geographical extent of the market. One important way through which information technology affects commuting is through the increased possibilities for telework. Findings from the transportation literature suggest that telecommuting may lead to longer commutes, although it has been difficult to establish a causal link (De Graaff 2004, Andreev et al. 2010, Zhu 2012).

A greater geographical extent of labour markets could theoretically lead to improved labour market outcomes because of more efficient matching. It has been established that the quality of employer-employee matches has a strong link with the size of local labour markets. For instance Andersson et al. (2007) show that next to sorting of high quality workers, better quality matches are an important source of the urban wage premium in Florida and California, and Dauth et al. (2018) find similar evidence for Germany, and they show that matching externalities have become about 75% stronger between 1985 and 2014. By increasing the geographical extent of labour markets, telework may allow the matching externalities associated with labour market size to be generated across greater geographical areas, thereby fostering borrowed size in places surrounding large cities.

In markets for local products the geographical extent is determined by the willingness to travel for goods. Anenberg and Kung (2015) note that information technology, and the internet specifically, likely decreases 'spatial information frictions' associated with locally produced and consumed goods. This echoes suggestions from the transportation literature that through the dissemination of information about destinations and reducing uncertainty, online local information may decrease generalised costs of travel and lead to more trips and longer travel distances for specific destinations (Mokhtarian and Tal 2013).

Through market size externalities, a greater geographical extent of local product markets is generally associated with improved market outcomes. Berry and Waldfogel (2010) have shown that, consistent with theories from the industrial organization literature (Shaked and Sutton 1987), the type of improved outcomes due to market size depend crucially on whether quality is produced with fixed or variable costs. Markets in which quality is produced with fixed costs, including for instance newspaper markets, tend to remain concentrated as they grow large, offering higher quality goods. Markets in which quality is produced with variable costs, such as restaurant markets, tend to show an increase in the range of qualities available, as they grow larger. Through online information, places near large cities may become more accessible, and may be increasingly able to draw upon the population of their neighbours, sustaining higher quality local products, and greater product variety.

Stylized facts – Above I have argued that the recent growth patterns of Europe's small and medium sized cities, whereby not necessarily the largest cities have grown fastest (Dijkstra et al. 2013; Camagni et al. 2014; Ahrend et al. 2017), may be explained by the rise of information technology that has better enabled cities in proximity of larger cities to borrow size. I highlighted two mechanisms. First telecommuting may have resulted in a greater geographical extent of local labour markets, allowing matching externalities across greater geographical areas. Second, the increased availability and consumption of online information on local products may have increased the willingness to travel for these products, which may disproportionately affect market access in small places near large cities. In this subsection I show that readily available data on IT usage and travel behaviour is in line with these hypotheses, and further research is warranted.

In February 2013 Yahoo's CEO Marissa Mayer banned working from home.[1] This fed rumours that teleworking did not live up to its promises, and was on its way out. Data from Statistics Netherlands (CBS) displayed in Figure 1 shows that while there was a period of stagnation, during the economic recession, teleworking has been on the rise again since 2013. Between 2003 and 2015 growth has been significant, and more than 70% of firms with more than 10 employees employ teleworkers nowadays. Figure 2 shows that similar growth has taken place in the consumption of online information on local products, evidenced by the sharp increase of the usage of popular restaurant- and tourism review websites Iens.nl and Tripadvisor.

---

[1] https://www.nytimes.com/2013/02/26/technology/yahoo-orders-home-workers-back-to-the-office.html

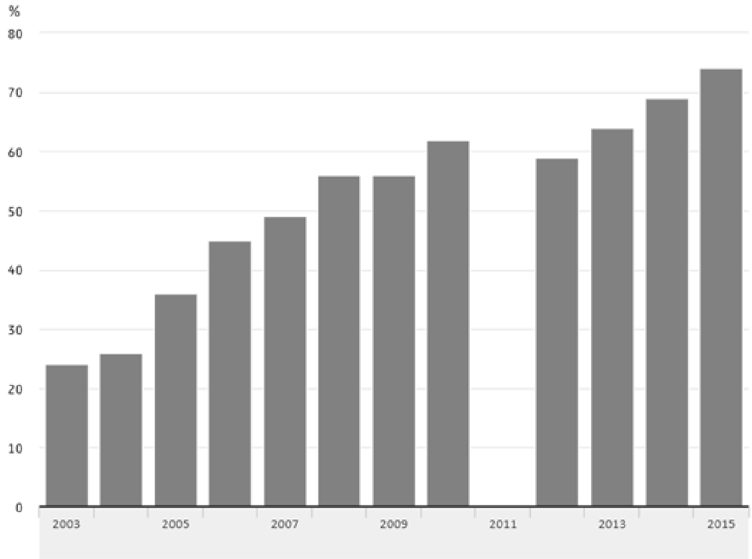


Figure 1: Firms (10+ employees) with teleworkers 2003-2015. Source: https://www.cbs.nl/nl-nl/nieuws/2015/51/telewerken-weer-in-de-lift

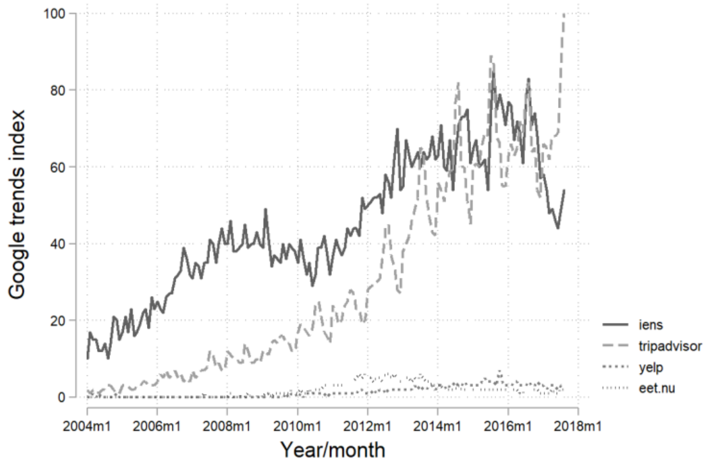


Figure 2: Google Trends of restaurant review websites.

Travel behaviour data from Dutch mobility surveys in 2000 (MON) and 2015 (OViN) show that during the rise of information technology usage, travel distances for both commuting and local products such as services have increased. The left panel in Figure 3 shows that the distribution of commuting trips contained relatively less short commutes in 2000. Between 2000 and 2015 the average distance of commuting trips has increased significantly with 12.4%, from 14.12 to 15.87 kilometres on average. The distribution of travel distances for service trips, depicted in the right panel of Figure 3, shifted to the right during the same period. For these trips the average increase was even larger with 21.5%, from 7.63 to 8.94 kilometres on average.

Overall the data show that over the last two decades there was a significant increase in the usage of information technology in local labour markets and local product markets. In local labour markets, this is evidenced by a sharp increase in telecommuting, up to the point that more than 70% of firms (10+ employees) employ telecommuters. In markets for local products, this growth is evidenced by the increased consumption of local online content. Travel behaviour data shows that over the same period the geographical extent of these local markets increased significantly. These facts warrant further research into the relation between information technology, the geographical extent of local markets, and the relation with recent growth patterns of places in proximity of larger cities.

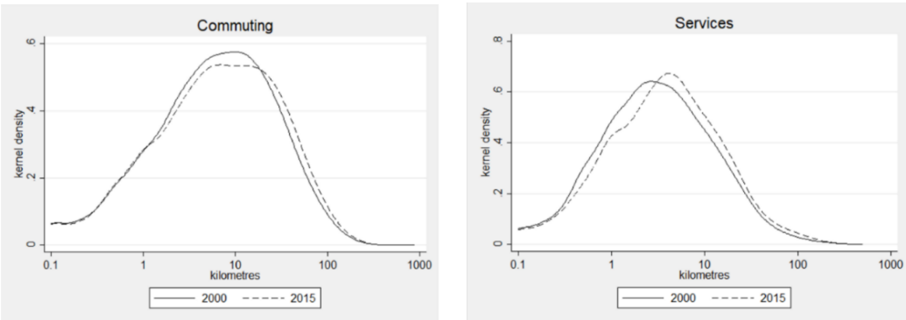


Figure 3: Distribution of travel distance for commuting (l) and services (r) trips. Source: MON 2000 and OViN 2015

Research questions - Several scholars have suggested that information technology or fast internet is a driver of ‘borrowed size’ (Hesse 2016), and a potential explanation of the remarkable growth patterns of Europe’s small- and medium sized cities (Dijkstra et al 2013). However these suggestions have been made without theoretical reasoning or empirical validation. In the current section I have already argued that there is a plausible theoretical relation between information technology and the geographical extent of local markets, through travel behaviour, that may form the missing link in this story. The objective of this thesis is to empirically test this theory. To accomplish this, this thesis addresses the following research questions.

1. *To what extent does information technology increase the geographical extent of local labour and product markets?*

2. *To what extent has the advent of information technology led to better local (labour or product) market outcomes in places in proximity of large cities?*

The first question will be answered by measuring individual IT usage and the geographical range of people in distinct markets. Individual IT usage will be measured using indicators such as teleworking (local labour market), and the number of days using the internet (local product market). The geographical range of people will be measured using travel behaviour data (commuting and travel for service purposes). Finally the relationship between IT usage and travel behaviour will be estimated using econometric models. Chapters 2, 3, and 6 will contribute to answering research question 1.

The second question focuses not on individual outcomes, but on local outcomes. This requires local measures of IT penetration or usage. These measures will be related to local measures of urbanization economies with the use of econometric models. For the local labour market, I use (changes in) the employment rate as a measure of matching on the labour market,[2] which is in turn one of the main economies of urbanization in this market (Duranton and Puga 2004). For local product markets I use local product variety as a measure of urbanization economies (Tabuchi and Yoshida 2000). Chapters 4 and 5 relate to research question 2.

1.3 Methods and data

The research questions of this thesis will be addressed by estimating the relation between information technology and outcome variables with the use of econometric models. Econometrics is a powerful tool, but one needs to take into account different sources of bias in order not to mistake mere correlations for true causal relationships. These sources include omitted variables, sorting based on preferences, reverse causality, and measurement error. Various techniques exist to overcome these biases, and in this thesis I mainly aim to come closer to causal estimations by (1) using time variation within units (individuals or places) to estimate relationships, (2) controlling as much as possible for relevant characteristics of individuals or places, and (3) by subjecting the estimation results to different sensitivity checks. Using time-variation to estimate effects is straightforward, but it is data intensive, requiring the same units to be observed over multiple years. In this section I delve deeper into the techniques I employ in this thesis to go beyond mere correlations and I discuss the types of data used in this study, and the advantages and drawbacks of each of these types.

Time variation, fixed effects, and lagged dependent variables – To measure the effect of information technology on travel behaviour and urbanization economies, I

[2] It should be noted that employment is a rather coarse measure of matching that only captures the extensive margin. Data issues prevented me from measuring match quality at a more detailed level.

will make use of models that use time-variation in both dependent and independent variables. Rather than looking at correlations between information technology usage and outcomes, I will measure correlations between changes in IT usage and changes in outcomes. The econometric model I employ most frequently in this thesis is the *fixed-effects model*. This model requires *panel data* with observations of the same units (individuals or places), over multiple years, and for every unit a dummy variable is included. The general model is as follows:

$$Y_{it} = \boldsymbol{\beta X}_{it} + \eta_i + \theta_t + \epsilon_{it} \quad (1)$$

Here Y_{it} is the outcome variable that differs across units i and time periods t, $\boldsymbol{X}_{it}$ is a vector of independent variables, $\boldsymbol{\beta}$ are the associated coefficients to be estimated, η_i are unit fixed effects, θ_t are time fixed effects, and ϵ_{it} is the error term.

The fixed effects η_i control for all time-invariant characteristics of units. This makes the fixed-effects model an excellent tool to (partially) overcome biases due to unobserved characteristics of places or individuals. The fixed effects model relies on the standard assumptions of linearity, random sampling and no multicollinearity. The key identifying assumption of fixed-effects models is that variation in the independent variable of interest is (almost) random, in the sense that it is not related to time-varying unobservable variables that are correlated with the outcome. Across this thesis, I aim to minimize the bias of these time-varying unobservables by including as many relevant control variables as possible. The final important assumption of FE estimation is the absence of serial correlation. This means that conditional on the included variables, the outcomes of individuals should not be correlated over time. In practice this assumption may be rarely met, but as with the problem of heteroscedasticity, the issue can be addressed by estimating robust standard errors (Wooldridge 2002).

One important drawback of the fixed-effects model is that it does not estimate effects of variables that do not vary over time. A standard alternative that does provide estimates of these effects is the *random-effects* model, that treats the unit effects η_i as random variables with an error distribution instead of parameters to be estimated. However, this model requires the assumption that the unobserved effects are not correlated with the independent variables (Wooldridge 2002). The fixed-effects model does not require this assumption, and is therefore often preferred in applied work (Angrist and Pischke 2008).

The random effects model uses both cross-sectional and time-variation to estimate effects. Whereas this can be more efficient, the policy relevant outcomes of this thesis are primarily related to time-varying factors (What will be the effects if more people take up teleworking? What will be the labour market effects of bringing broadband to peripheral areas? etc.). A promising option that combines the benefits of random effects estimation, and fixed effects estimation is the Correlated random effects model. This model is essentially a random effects model, but it includes the unit-level mean of all independent variables.The CRE model is asymptotically equivalent to the fixed-effects model (Dieleman and Templin 2014), but it still requires the assumption that the unobserved effects are uncorrelated with the independent variables. This is the second reason that I predominantly estimate fixed effects models in this thesis. I estimate fixed effects models in Chapters 2, 3, 4, and 6.

While panel data sources have become increasingly common over the last decades, the proper panel data for some of the research questions in this thesis are missing or unavailable. Particularly, it was hard to find panel data on geographical differences in usage of information technology for consumption purposes, in Chapter 5. We were only able to gather cross-sectional data on these geographical differences, but for the outcome variable we gathered data over multiple years. The main model in this Chapter is a cross-sectional OLS model that includes a (long) lagged dependent variable (LDV). This model has the following form:

$$Y_{it} = \boldsymbol{\beta X}_{it} + \delta Y_{i,t-h} + \epsilon_{it} \tag{2}$$

Here h denotes the number of periods time-lag that is used. LDV models can be used to control for a different type of selection than FE models, based on historical values of the outcome variable, rather than unobserved time-invariant characteristics (Angrist and Pischke 2008). The key identifying assumption of the LDV model is that variation in the independent variable of interest is random, conditional on $Y_{i,t-h}$. In Chapter 5, we include a long (16-year) lagged dependent variable (cuisine variety in this case), and we use it as a proxy for time-invariant culinary culture. In Chapters 2 and 3 we also estimate LDV models as part of a sensitivity analysis. Here we take the 2-year lag, and we assess the sensitivity of our results to controlling for a different type of selection (Angrist and Pischke 2008).

Data – As mentioned before, using time variation to estimate the effects of information technology on commuting and urbanization economies is a data-intensive strategy. Especially when considering that I want to control for as many relevant control variables as possible, to limit the bias of remaining time-varying omitted variables. Luckily, an increasing number of data sources provides the longitudinal data required. In this thesis I use combinations of data from panel surveys, population and firm registers, geographical data provided by statistical agencies, and web-sourced indicators.

In panel surveys, individuals are asked the same questions over multiple years. This allows to estimate the effects of changes in IT usage within individuals using fixed effects models. Usually, these surveys involve an extensive list of questions, that provide the necessary control variables for regression models. One important drawback of survey data is that it is based on a sample of the population. Even if care is taken to ensure random selection and representativeness, spurious relations may emerge, and results have only limited external validity. The possibility to 'difference out' time-invariant characteristics of individuals, and the great level of detail, makes panel survey data particularly useful to analyse human behaviour. Therefore in this thesis I use panel survey data to analyse relations between IT usage and travel behaviour.

Given the limited external validity of results from panel surveys, to measure the effects of IT on macro outcomes such as urbanization economies it is more logical to look at compiled data from statistical agencies, or at register data itself. The advantage of this type of data is that it involves the entire population of people or firms (in a country). However, even when combined with geographical background variables, register data still often lacks the level of detail that is present in survey data. One important detail for this study, often missing in these types of data, is the availability or usage of information technology at the local level. For instance, the Swedish population register that I use in Chapter 4 has detailed information on earnings, education, and occupation of individuals, but no information on whether households have access to internet. Therefore I combine the register data with geographical data from the Swedish telecom authority on the availability of high-speed internet. In Chapter 5 there is a similar issue. The firm register that I use has detailed data on the availability of a variety of local products in all Dutch places and municipalities, but it lacks information about the usage of IT. Here I measure IT usage with web-sourced data from a restaurant review website. Web sourced data is promising in many respects, especially to measure (historical) penetration of IT usage. But there are also drawbacks, including the continuous rise and fall of popular websites.

1.4 Research outcomes

Table 1 shows the structure of this thesis. The research questions are answered separately for local labour markets, in Chapters 2 to 4, and for local product markets, in Chapters 5 and 6. These chapters consist of distinct research articles. Although these articles are readable on their own, they all relate to the aim of this thesis. Chapter 7 concludes this thesis.

Table 1: Thesis structure

	Part 1 *Local labour markets*	**Part 2** *Local product markets*
RQ1: IT and travel behaviour	Chapters 2 and 3	Chapter 6
RQ2: IT and urbanization economies	Chapter 4	Chapter 5
Synthesis	Chapter 7	

In Chapter 2 I tackle the question whether information technology leads to a greater geographical scale of labour markets. To this end, I estimate the effect of teleworking on the length of commutes. This is not the first paper in the literature about working from home and commuting, but it addresses an important gap, namely the absence of panel data studies that allow to control for unobserved characteristics of individuals. Earlier studies that relied on cross sectional data reported difficulties with establishing to what extent there is a causal relation. By controlling for a number of biases due to unobserved characteristics, I aim to come closer to a credible causal estimate. The conclusion of this Chapter is that teleworking allows workers to accept longer commutes. I find a 5% increase in commuting time for workers who start teleworking.

In Chapter 3 I delve deeper into the relation between telework and the length of commutes. I use a more recent panel data set from the Netherlands to reproduce the results of Chapter 2. Importantly, I assess whether the effects have grown stronger over time, and whether the effects I find are driven by certain occupations, or by people residing in certain locations. The conclusion of this Chapter is that the relationships found in Chapter 2 are reproducible with different and more recent data. I find that workers who start teleworking increase their commute by 11% on average, and these effects are mainly driven by occupations that require 'academic' skills, and by administrative jobs. The effects are present across different levels of urbanization, but the strongest relationship is found in 'moderately urban' areas, largely situated between larger cities.

Where Chapters 2 and 3 focus on whether IT increases the geographical scale of labour markets, Chapter 4 addresses whether the advent of information technology also leads to longer commutes, and to better labour market outcomes, especially in places in proximity of large cities. In other words, here I investigate whether IT leads to 'borrowed size' patterns in the labour market. The data used in this Chapter is from Sweden, and I measure IT by the availability of broadband internet, at the level of 250*250 metre grid cells, over time between 2007 and 2015. With the use of fixed effects models I investigate whether changes in broadband availability are related to commuting distance and to matching on the labour market, measured by the propensity to be employed. The main result is that changes in the penetration of broadband in the place of residence are correlated with the probability to be employed, especially in places between 20 and 50 kilometres from large cities. But changes in availability of broadband at the residential location do not show this relation. I conclude that the effects of broadband on labour market outcomes appear rather small, and the suggestive evidence of borrowed size due to broadband requires further investigation.

Chapters 5 and 6 answer the research questions for a distinct local product market, namely the market for restaurants. Variety on the restaurant market has served as a prime example to measure urbanization economies on the consumption side of the economy (Glaeser et al. 2001, Couture 2013, Schiff 2015). In Chapter 5 I assess to what extent the penetration of information technology is related to better outcomes in the restaurant market, measured by cuisine variety. This Chapter uses data from the Netherlands, but whereas the Dutch registry for companies has detailed data on the type of cuisine of all registered restaurants, there are no readily available measures of IT penetration in the Netherlands. I therefore constructed a measure of penetration of information technology in local restaurant markets by comparing the number of restaurants reviewed online to the actual registered number of restaurants, for all named places in the Netherlands. I estimate the relationship between this share of reviewed restaurants and cuisine variety with a model relying on cross sectional data, but I use historical cuisine variety to control for bias due to unobserved culinary culture. The main finding of this Chapter is that the penetration of IT and cuisine variety have a strong relationship, that increases with city size, consistent with earlier findings by Anenberg and Kung (2015), but also decreases with distance to large population centres. We conclude that next to a complementary relation between IT and urbanization, there may exist a spillover mechanism whereby IT allows places in the vicinity of large cities to capitalize on the size of their neighbours and sustain more cuisine variety.

In Chapter 6 I investigate whether information technology is also associated with a greater geographical scale of local product markets. This is the hypothesized mechanism behind the findings of Chapter 5. I use data from a recent panel data source from the Netherlands, specifically designed to analyse the relation between IT and travel behaviour. I operationalize the geographical scale of local product markets using travel behaviour for service and personal care purposes. IT usage is measured by daily internet usage. This is one of the first papers to estimate the effect of IT usage on travel for purposes that do not have a clear online counterpart. The results suggests that changes in IT usage strongly affect travel distance for service purposes, and this relationship can explain the observed rise travel distance for services over the last decades. In contrast with the other Chapters that all consist of published or submitted papers, the work in Chapter 6 should be regarded as an empirical chapter that aims to address research question 1 for local product markets.

Chapter seven summarizes the results, assesses which contributions have been made to the literature, reiterates limitations, and provides a research agenda for the spatial economic effects of information technology.

References

Ahrend, R., Farchy, E., Kaplanis, I., & Lembcke, A. C. (2017). What makes cities more productive? Evidence from five OECD countries on the role of urban governance. Journal of Regional Science, 57(3), 385–410. https://doi.org/10.1111/jors.12334

Ahrend, R., & Schumann, A. (2014). Does regional economic growth depend on proximity to urban centres? OECD Regional Development Working Papers, 0–23. https://doi.org/10.1787/5jz0t7fxh7wc-en

Andersson, F., Burgess, S., & Lane, J. I. (2007). Cities, matching and the productivity gains of agglomeration. Journal of Urban Economics, 61(1), 112–128. https://doi.org/10.1016/j.jue.2006.06.005

Andreev, P., Salomon, I., & Pliskin, N. (2010). Review: State of teleactivities. Transportation Research Part C: Emerging Technologies, 18(1), 3–20. https://doi.org/10.1016/j.trc.2009.04.017

Anenberg, E., & Kung, E. (2015). Information technology and product variety in the city : The case of food trucks. Journal of Urban Economics, 90, 60–78. https://doi.org/10.1016/j.jue.2015.09.006

Angrist, J. D., & Pischke, J. S. (2008). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press.

Berry, S., & Waldfogel, J. (2010). Product Quality and Market Size. The Journal of Industrial Economics, 58(1), 1–31. https://doi.org/10.1111/j.1467-6451.2010.00404.x

Burger, M. J., & Meijers, E. J. (2016). Agglomerations and the rise of urban network externalities. Papers in Regional Science, 95(1), 5–15. https://doi.org/10.1111/pirs.12223

Camagni, R., Capello, R., & Caragliu, A. (2014). The Rise of Second-Rank Cities: What Role for Agglomeration Economies? European Planning Studies, 23(6), 1069–1089. https://doi.org/10.1080/09654313.2014.904999

Couture, V. (2013). Valuing the Consumption Benefits of Urban Density. Retrieved from http://faculty.haas.berkeley.edu/couture/download/jmp.pdf

Dauth, W., Findeisen, S., Moretti, E., & Suedekum, J. (2018). Matching in cities. In NBER Working Paper Series (No. 25227). Retrieved from https://www.nber.org/papers/w25227.pdf

De Graaff, T. (2004). On the Substitution and Complementarity between Telework and Travel: A Review and Application (Vrije Universiteit Amsterdam). Retrieved from http://degree.ubvu.vu.nl/repec/vua/wpaper/pdf/20040016.pdf

Dieleman, J. L., & Templin, T. (2014). Random-effects, fixed-effects and the within-between specification for clustered data in observational health studies: A simulation study. PLoS ONE, 9(10). http://doi.org/10.1371/journal.pone.0110257

Dijkstra, L., Garcilazo, E., & McCann, P. (2013). The Economic Performance of European Cities and City Regions: Myths and Realities. European Planning Studies, 21(3), 334–354. https://doi.org/10.1080/09654313.2012.716245

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. Handbook of Regional and Urban Economics Vol. 4, 4(04), 2063–2117. https://doi.org/10.1016/S0169-7218(04)07048-0

Gaspar, J., & Glaeser, E. L. (1998). Information Technology and the Future of Cities. Journal of Urban Economics, 43, 136–156. https://doi.org/10.1006/juec.1996.2031

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. Journal of Economic Geography, 1(1), 27–50. https://doi.org/10.1093/jeg/1.1.27

Graham, D. J. (2008). Identifying urbanisation and localisation externalities in manufacturing and service industries *. 88(1). https://doi.org/10.1111/j.1435-5957.2008.00166.x

Handbury, J., & Weinstein, D. E. (2015). Goods Prices and Availability in Cities. Review of Economic Studies, 82(1), 258–296. https://doi.org/10.1093/restud/rdu033

Hesse, M. (2016). On borrowed size , flawed urbanisation and emerging enclave spaces : The exceptional urbanism of Luxembourg , Luxembourg. European Urban and Regional Studies, 23(4), 612–627. https://doi.org/10.1177/0969776414528723

Melo, P. C., Graham, D. J., Levinson, D., & Aarabi, S. (2016). Agglomeration, accessibility and productivity: Evidence for large metropolitan areas in the US. Urban Studies, 54(1), 0042098015624850. http://doi.org/10.1177/0042098015624850

Melo, Patricia C., Graham, D. J., & Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration economies. Regional Science and Urban Economics, 39(3), 332–342. https://doi.org/10.1016/j.regsciurbeco.2008.12.002

Mokhtarian, P. L. (2004). Telecommunication and Travel: The Case for Complementarity. In Auswirkungen der virtuellen Mobilität (pp. 217–243). Retrieved from https://link.springer.com/chapter/10.1007/978-3-540-76793-0_18

Mokhtarian, P. L., & Tal, G. (2013). Impacts of ICT on Travel Behavior: a tapestry of Relationships. In J.-P. Rodrigue, T. Notteboom, & J. Shaw (Eds.), The SAGE handbook of transport studies. https://doi.org/10.4135/9781446247655.n14

OECD. (2015). The Metropolitan Century. https://doi.org/10.1787/9789264228733-en

Rice, P., Venables, A. J., & Patacchini, E. (2006). Spatial determinants of productivity: Analysis for the regions of Great Britain. Regional Science and Urban Economics, 36(6), 727–752. https://doi.org/10.1016/j.regsciurbeco.2006.03.006

Salomon, I. (1985). Telecommunications and travel: substitution or modified mobility? Journal of Transport Economics and Policy, Vol. 19, 219–235. https://doi.org/10.1080/09595238100185051

Schiff, N. (2015). Cities and product variety: Evidence from restaurants. Journal of Economic Geography, 15(6), 1085–1123. https://doi.org/10.1093/jeg/lbu040

Shaked, A., & Sutton, J. (1987). Product Differentiation and Industrial Structure. The Journal of Industrial Economics, 36(2), 131–146. Retrieved from https://www.jstor.org/stable/pdf/2098408.pdf

Tabuchi, T., & Yoshida, A. (2000). Separating Agglomeration Economies in Consumption and Production. Journal of Urban Economics, 48, 70–84. https://doi.org/10.1006/juec.1999.2157

Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data. The MIT Press, Cambridge, MA. Zellner, A, 58, 348–368.

Zhu, P. (2012). Are telecommuting and personal travel complements or substitutes? Annals of Regional Science, 48(2), 619–639. https://doi.org/10.1007/s00168-011-0460-6

Part 1: Local labour markets

2. Working from home and the willingness to accept a longer commute

Published in The Annals of Regional Science 61.2 (2018): 375-398.
Written by Duco de Vos, Evert Meijers, and Maarten van Ham.

Abstract
It is generally found that workers are more inclined to accept a job that is located farther away from home if they have the ability to work from home one day a week or more (telecommuting). Such findings inform us about the effectiveness of telecommuting policies that try to alleviate congestion and transport related emissions, but they also stress that the geography of labour markets is changing due to information technology. We argue that estimates of the effect of working from home on commuting time may be biased because of sorting based on residential- and commuting preferences. In this paper we investigate the relationship between telecommuting and commuting time, controlling for preference based sorting. We use 7 waves of data from the Dutch Labour Supply Panel and show that on average telecommuters have higher marginal cost of one-way commuting time, compared to non-telecommuters. We estimate the effect of telecommuting on commuting time using a fixed-effects approach and we show that preference based sorting biases cross-sectional results upwards. This suggests that the bias due to sorting based on residential preferences is strongest. Working from home allows people to accept 5 percent longer commuting times on average, and every additional 8 hours of working from home are associated with 3.5 percent longer commuting times.

2.1 Introduction

There is an ongoing debate about the extent to which working from home (also called telecommuting) affects the length of the commute people are willing to accept. Early interest in the effect of telecommuting on commuting distance and household travel was mainly aimed at establishing whether telecommuting could be an effective policy instrument to alleviate congestion and emissions associated with car use (Salomon 1985; Nilles 1991; Lund and Mokhtarian 1994). Increasingly, attention is being given to the notion that telecommuting also affects the geography of labour markets, for example, by having a positive effect on job accessibility (Muhammad et al 2008; Van Wee et al 2013). Understanding the relationship between telecommuting and the length of the commute may thus both inform policies aimed at alleviating congestion and transport related emissions, and policies that aim to improve the economic performance of cities and regions.

Most empirical work on the effects of working from home on commuting tends to corroborate the intuitive notion that being able to avoid the commute one day in the week makes workers more willing to accept a longer commute on the other days of the week (Jiang 2008; Zhu 2012; Kim et al 2015). However, estimates for the size of this effect vary across the literature, the set of control variables included differs between studies, and there is little attention for the intensity of telecommuting (the number of days per week/month). Moreover, there is no consensus on a strategy to deal with sources of bias stemming from the fact that commute length and telecommuting are often decided upon simultaneously. While some studies aim to eliminate the positive bias that arises if long commutes influence the decision to telecommute (Jiang 2008; Zhu 2012), there is a lack of attention for preference-based sorting. OLS estimates will be biased downward if workers who dislike commuting, and hence have shorter commutes, might also be more likely to work from home. On the other hand, those who have long commutes may be the ones that value residing in more rural areas, where housing quality is cheaper, and working from home may also be more attractive. The latter type of sorting would bias OLS estimates upward.

The objective of this study is to find out to what extent controlling for preference-based sorting affects the relationship between telecommuting and the length of the commute. Where earlier research on this subject is largely based on either panel data from specific experiments, or cross sectional data from large scale surveys, we use Dutch data from a panel survey, representative of the Dutch working age population, spanning 12 years. In the first part of our analysis we provide evidence that preferences for commuting differ between telecommuters and non-telecommuters by comparing the marginal costs of one-way commuting time (MCC) of both groups. To estimate the MCC we use job search and job mobility models, following the approach of Van Ommeren and Fosgerau (2009). The panel structure of the data then allows us to model commuting time and examine to what extent such individual preferences bias cross sectional results, through preference-based sorting. We do this by comparing OLS estimates of commuting time to the results of a fixed effects model that controls for unobservable time invariant characteristics of respondents. Finally in the sensitivity analysis we apply an even stricter identification method based on the timing and intensity of telecommuting, we employ two alternative identification methods, and we allow for a non-linear effect of weekly hours spent working from home.

2.2 Telecommuting and the length of the commute

Theoretical implications of telecommuting - The potential spatial effects of telecommuting, and other ICT activities have been theorized upon for at least 50 years. According to Webber (1963), the observed spatial expansion of market areas during the 1960s due to, inter alia, information flows was indicative of a looming "demise of the city" (Webber 1963, p. 1099). Such visions were generally based on the idea that information and communications technology would eventually substitute face-to-face contact, and have been a recurrent theme in futurist writings on the death of cities, and the death of distance (Toffler 1980; Naisbitt 1994; Cairncross 1997).

In much of the literature, telecommuting is seen as a potential policy instrument to decrease car-travel, of which the effectiveness is dependent on the overall effect on travel. In transportation research it is often stressed that telecommuting, and ICT activities in general, may substitute, complement, modify, or neutrally affect travel (Salomon 1985). The notion of complementary travel is based on the idea that telecommuting may induce people to accept jobs over longer distances, making the net travel effects of telecommuting not necessarily negative. Furthermore, it is argued that households have a rather fixed mobility budget, and a decrease in trips for commuting would be substituted by leisure trips, and trips of other household members (De Graaff 2004).

However, the welfare effects of telecommuting may stretch further, because workers that are able to telecommute can expand the geographical areas in which they look for jobs (Van Wee et al 2013). Basic urban economic models support the intuition that if telecommuters have less commuting trips than non-telecommuters, they bid less for homes closer to the Central Business District (the location of employment), and more for suburban homes (Alonso 1974; Lund and Mokhtarian 1994; Jiang 2008). Rhee (2008) shows that in theory, similar results could be obtained in cities with dispersed employment. In situations with little building restrictions, telecommuting may thus in theory promote residential sprawl in a similar way as the automobile did (Glaeser and Kahn 2004). In settings with strict urban containment policies, and a low elasticity of housing supply, such as the Netherlands (Vermeulen and Rouwendal 2007), possibilities for telecommuting may increasingly enable workers to live in one city and reap the benefits of access to labour in other cities (Muhammad et al 2008; Van Wee et al 2013).

In the current work we are predominantly interested in the effect of telecommuting on the geographical scale of labour market areas. Therefore we focus on the relatively uncontested mechanism by which telecommuting potentially increases the length of one-way commutes, because it allows workers to commute less frequently. We do not take into account the effects of telecommuting on non-commute trips, and travel behaviour of other household members.

Empirical issues – Empirical research on the effects of telecommuting on the length of the commute started in the early 1990s, when personal computers started to become a household commodity. In a seminal publication, Nilles (1991) investigates the potential effects of telecommuting on urban sprawl and household travel, using data from a telecommuting experiment with California State workers that spanned two years. He concludes that at the time, telecommuting did not (yet) exacerbate urban sprawl, and that it resulted in decreased household travel. He did however find that telecommuting was associated with moves farther away from the work location, so his findings did not rule out future *telesprawl* as a consequence.

Later evidence on the relationship between telecommuting and the length of the commute is somewhat scattered, in part because of different definitions of telecommuting.[1] In a review of evidence by De Graaff (2004) it is concluded that most studies show a negative relationship between telecommuting and the number of commuting trips, and studies that do investigate the length of the commute find mixed evidence, but do not rule out a positive relationship. Andreev et al (2010) conclude similarly, and stress that the majority of the literature suffers from problems such as the lack of a universal definition of telecommuting, the external validity of the results, and absence of theoretical substantiation of the results.

Recent endeavours increasingly pay attention to potential sources of bias that influence the results from observational studies. These sources can be divided into (1) omitted variables, (2) reverse causality, and (3) preference-based sorting. With respect to omitted variables, the advent of large scale surveys in which questions about telecommuting were asked, made it possible to control for a variety of respondent characteristics, and also made it possible to assess telecommuting across different industries. A notable work in this respect is Kim et al (2012), who estimated the effect of telecommuting on peripheral living, controlling extensively for household characteristics including income, and job locations. Accounting for wage seems particularly relevant in telecommuting research, because earnings and telecommuting status tend to be correlated (Muhammad et al 2008).

Jiang (2008, p. 10) provides a clear-cut definition of two other types of bias involved in the relationship between telecommuting and commuting distance, and the direction of these biases: "If [a] longer commute encourages an individual to work from home when allowed, a regression of commute length on telecommuting status will overestimate the effect of telecommuting. On the contrary, telecommuters could be those who feel more pressures from traffic. They would have shorter commutes in the absence of telecommuting opportunities. This unobserved selection will lead to a downward [bias] in the regression estimates." We refer to the first bias he addresses as reverse causality, and to the second as sorting based on commuting preferences.

Commuting time and telecommuting may not only be jointly influenced by commuting preferences, but also by residential preferences. Individuals that prefer rural living, and generally have longer commutes, may have larger or more comfortable houses because housing quality tends to be cheaper away from central business areas (Muth 1969). Assuming that spending time in higher quality housing is more pleasurable than spending time in houses of lower quality (Gubins and

[1] Mokhtarian et al (2005) illustrate that definitions, measurement instruments, sampling, and vested interests affect the quality and utility of data, using telecommuting as a case study.

Verhoef 2014), we may expect that working from home is more attractive for rural dwellers. With this type of sorting based on residential preferences, an OLS regression of commuting time on telecommuting status would overestimate the real relationship.

A study in which an attempt is made to overcome the bias from potential reverse causality is done by Zhu (2012). He employs an instrumental variables approach, using the number of phones in a household, and the usage of the internet at home as instruments, argued to influence commuting distance only through the effects on telecommuting. Although the reverse causality bias he refers to should lead to overestimation of the effect of telecommuting, he finds that his IV approach leads to *higher* estimates, compared to OLS. According to his IV results for the year 2009, telecommuters that work from home at least once a week have a 1576 percent longer commuting distance, and a 160 percent longer commuting duration on average.[2] While these estimates are large, the results suggest that the bias not accounted for in OLS models is positive rather than negative.

Jiang (2008) uses a similar IV approach, but the instruments in this study are based on the penetration of home-based teleworking across combinations of occupations and city size classes. The results of this study show that OLS tends to underestimate the real effect of telecommuting. While the OLS estimates in this study show that, at least for married women, telecommuting increases commuting time by 3 minutes, the IV estimates suggest an effect of 9 to 11 minutes. No significant results are found for men, and single women.

The current study addresses several gaps that emerge from the literature. First, next to household characteristics we include detailed job characteristics as control variables, including monthly wage, the type of industry, the type of employment, and the usual number of work days per week. Especially the latter control is a novelty in this type of research. Second, we make use of the time dimension of our data, and we focus on the effect of *changes* in telecommuting status, on *changes* in commuting time.[3] Arguably, this makes the potential bias of reverse causality less pressing. While exogenous changes in commuting time (for instance due to firm relocations) may influence the decision to telecommute, this still indicates that telecommuting increases the willingness to accept a longer commute. Finally, the time dimension of the data also allows us to control for all time-invariant characteristics of respondents through the use of fixed effects models. Such time invariant characteristics include unobserved commuting- and residential preferences, so this approach allows us to address the bias

[2] Given his log-linear model, these are the marginal effects of the telecommuting dummy for which the point estimate is 2.819 for the distance model, and 0.993 for the duration model. The marginal effects are calculated as $(e^{\beta} - 1) * 100\%$. The corresponding marginal effects of his OLS estimates are more realistic: 23 and 19 percent respectively.

[3] The fixed-effects models use variation in telecommuting *within* individuals, over time, to explain time-variation in commuting time within individuals.

due to preference-based sorting. This is one of the first studies to address the relationship between telecommuting and commuting distance with a fixed effects approach.[4]

2.3 Data and methods

Data description – Our empirical analyses are based on data from the Netherlands. The urban landscape of the Netherlands is characterized by a polycentric urban structure with many small- and medium sized cities. Labour- and housing markets stretch far beyond cities, and it is relatively common to live in or near one city, and work in another urban area (Burger and Meijers 2016). Another notable characteristic if the concentration of employment in the Randstad area, in the west of the country. This area is also characterized by a higher wage level (Groot et al 2014), and better matching between workers and employers indicated by lower levels of overeducation (Büchel and Van Ham 2003). In the Netherlands telecommuting is a relatively widespread phenomenon, due to the mass adoption of ICT, the high share of the tertiary sector in economic activities, and the high population density and the associated congestion problems (Muhammad et al 2007, 2008).

The data we use comes from the Labour Supply Panel (SCP 2016), and it consists of the 7 latest biannual waves (between 2002 and 2014). While the panel has been running since 1985, 2002 is the first year in which questions were asked about the degree to which people work from home. We have 18,730 observations for 7,497 individuals.[5]

Telecommuting definition – We are interested in the relationship between telecommuting and the (accepted) commuting time. As mentioned earlier, a variety of definitions of telecommuting is used in related literature. Mokhtarian et al (2005) give an overview, starting from the idea that telecommuting involves "salaried employees of an organization [that] replace or modify the commute by working at home or a location closer to home than the regular workplace, generally using ICT […]" (Mokhtarian et al 2005, p. 427). Studies from their literature review have threshold values for the intensity of telecommuting that range from at least once per month, to at least once per week.

Two questions from the survey we use relate to the intensity of telecommuting. The first asks to state how many days per month respondents work from home usually. The answers to this question are measured on an ordinal scale with 5 possibilities (0,

---

[4] Two notable studies that apply a fixed effects to approach to telecommuting research are De Graaff (2004), who looks at the relationship between telecommuting and total travel, and Kolko (2008), who aims to uncover how broadband availability affects the adoption of telecommuting.

[5] We excluded extreme observations with commuting times longer than 500 minutes, monthly wages lower than €500, and daily wages higher than €1,000, and following other studies we excluded self-employed workers (Mokhtarian et al 2005; Kim et al 2012).

1, 2, and 3 days, and more than 3 days). The second question asks to state the average number of weekly hours spent from home in the 4 weeks before the survey date, and the resulting variable is measured on a continuous scale. Both questions contain some ambiguities, the first one about what is meant by `usually' and the second one about whether these hours working from home substitute commuting activities. Therefore our definition of telecommuters will be based on workers who usually work at home at least once per month, and have worked from home at least 4 hours on average during the 4 weeks before the survey date.[6]

Another question asks whether respondents use email or internet to work from home. Limiting our sample to salaried employees, our data thus allow us to define telecommuting as workers who work regularly *from home*, substituting commuting activities, and using ICT. The only discrepancy with the definition of Mokhtarian et al (2005) involves workers who telecommute from a location closer to home than the regular workplace. To the extent that we still measure telecommuting with an error, our estimates will be biased towards zero.

We will perform our analyses using (1) a dummy that indicates whether or not a respondent telecommutes, (2) an ordinal factor variable that denotes the usual number of telecommuting days per month, and (3) a continuous measure of the average weekly hours working from home (for telecommuters). Figure 1 shows the time patterns of the share of telecommuters, and the average weekly hours working from home. Both graphs show an increase in telecommuting (intensity) between 2002 and 2014, with a dip in the years 2008 and 2010.[7] The overall increasing pattern stresses that telecommuting is still a dynamic and upcoming trend.

[6] The 4 hours on average per week are chosen to come close to a threshold value of 2 telecommuting days per month, which holds the middle ground between the threshold values mentioned in Mokhtarian et al (2005).

[7] Data from the US Community Survey shows a similar dip in working from home around 2008, and in the popular press this dip is often attributed to the Great Recession (see http://www.bbc.com/news/business-19594518).

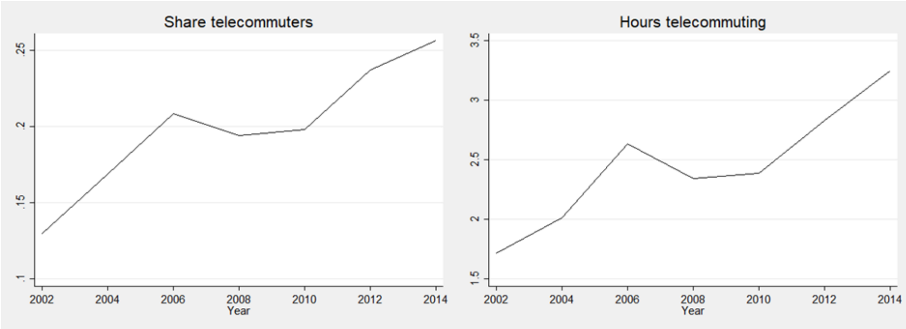


Figure 1: Time patterns of working from home

Other variables and summary statistics – Commuting time is measured as the usual time it takes to get to work from the residential location. The data does not contain information on commuting distance. However, modelling commuting distance is generally plagued by assumptions about mode choice and commuting speed (Isacsson et al 2013), while the use of commuting time can be justified by the assumption that commuting speed is optimally chosen (Van Ommeren and Fosgerau 2009).

Figure 2 shows the (kernel) distribution of one way commuting time for non-telecommuters, occasional telecommuters (up to 3 days per month), and regular telecommuters (more than 3 days per month). The figure shows that the distribution of commuting times for non-telecommuters has its bulk between 0 and 25 minutes, while the distributions of the other categories are more spread, including relatively longer commutes. The average commuting time for non-telecommuters is 24 minutes, versus 32 minutes for occasional telecommuters, and 31 minutes for regular telecommuters. So far, it seems that non-telecommuters have considerably shorter commutes on average, while regular telecommuters do not have longer commuting times than occasional telecommuters.

In Table 1 we make use of the time dimension of our data, and a similar pattern emerges. Commuting times of respondents who started to telecommute between two consecutive waves increased by 1.5 minutes on average, while the average increase is 0.31 for respondents for whom nothing changed, and average commuting times decreased for respondents that quit telecommuting.[8]

[8] T-tests confirm that differences between these groups are significant at the 5 percent confidence level.

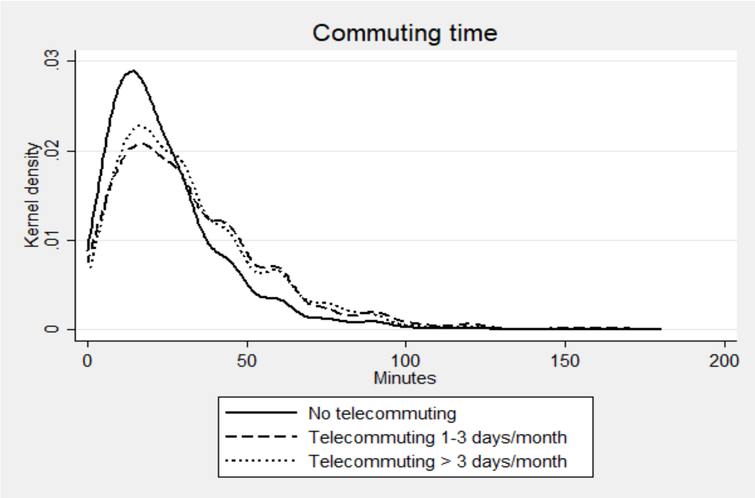


Figure 2: Distribution of commuting time according to telecommuting status. The density functions are estimated with a Gaussian kernel (bandwidth of 5 minutes).

Table 1: Telecommuting and changes in commuting time

	N	Mean change	Std. dev
Started telecommuting between t-2 and t	975	1.494	17.358
No change	9,554	0.307	14.377
Quit telecommuting between t-2 and t	704	-0.652	19.885

Other essential variables are all deduced from answers to questions in the survey: We calculate the daily wage of respondents based on the (stated) net wage per month and the usual number of working days per week, assuming 6 weeks of vacation on average; job search is measured as a dummy indicating the respondent is searching for a job at the moment the survey was conducted; job mobility is measured as a dummy that indicates whether or not the respondent changed jobs between two consecutive survey waves, and in our analysis we use the 2-year lead of this variable.

Table 2 shows the summary statistics of variables used in our analyses, including the average number of working days per week, firm sector and size, age, sex, the presence of children and a partner, and the wage of the partner.

Table 2: Summary statistics full sample

Variables	N	Mean	Std. dev.	Min.	Max.
Telecommuting	18,730	0.2	0.4	0	1
Telecommuting 0 days/month	18,730	0.681	0.466	0	1
Telecommuting 1 day/month	18,730	0.0984	0.298	0	1
Telecommuting 2 days/month	18,730	0.0758	0.265	0	1
Telecommuting 3 days/month	18,730	0.0753	0.264	0	1
Telecommuting >3 days/month	18,730	0.0698	0.255	0	1
Telecommuting weekly hours	18,730	2.467	6.803	0	97
Commuting time	18,730	26.81	21.07	0	180
Job search at t	18,730	0.114	0.317	0	1
Job move between t-2 and t	18,730	0.154	0.361	0	1
Job move between t and t+2	11,320	0.104	0.306	0	1
Daily wage	18,730	106.5	44.76	21.06	808.7
Monthly wage	18,730	1,742	791.9	501	16,667
Working days/week	18,730	4.323	0.965	1	7
Firm size	18,730	613.2	2,446	0	70,000
Age	18,730	43.11	11.04	16	66
Female	18,730	0.468	0.499	0	1
Partner	18,730	0.798	0.402	0	1
Partner wage	18,730	1,148	2,567	0	75,000
Children at home	18,730	0.571	0.495	0	1
Primary education	18,730	0.0153	0.123	0	1
Basic education	18,730	0.189	0.392	0	1
Higher education	18,730	0.38	0.485	0	1
Vocational education	18,730	0.29	0.454	0	1
Bachelor degree	18,730	0.125	0.331	0	1
Sector: Agriculture	18,730	0.00769	0.0873	0	1
Sector: Industry	18,730	0.122	0.327	0	1
Sector: Construction	18,730	0.041	0.198	0	1
Sector: Trade	18,730	0.122	0.328	0	1
Sector: Transport	18,730	0.0569	0.232	0	1
Sector: Business services	18,730	0.178	0.382	0	1
Sector: Healthcare	18,730	0.204	0.403	0	1
Sector: Other	18,730	0.0456	0.209	0	1
Sector: Government	18,730	0.107	0.309	0	1
Sector: Education	18,730	0.117	0.321	0	1
Jobtype: Civil servant	18,730	0.191	0.393	0	1
Jobtype: Employee	18,730	0.8	0.4	0	1
Jobtype: Director	18,730	0.00892	0.094	0	1

2.4 Methods

We introduced two types of preference-based sorting that may bias OLS results: sorting based on commuting preferences and sorting based on residential preferences. Our data allow us to investigate whether commuting preferences significantly differ between telecommuters and non-telecommuters. In the first part of our empirical analysis therefore, we use a model based on job search theory to calculate the marginal monetary value of one-way commuting time (MCC) both for both groups separately. If the MCC is significantly higher for telecommuters, we interpret this as evidence for sorting based on commuting preferences.

Using the job search approach we relate commuting time and wage levels with each other through their effects on (1) on-the-job search, and (2) job mobility (Van Ommeren and Fosgerau 2009). The intuition behind this approach is that workers are not in their preferred job per se, and are able to improve upon their situation by searching for jobs, and moving jobs if they find a better fit that improves their *lifetime* utility (Van Ommeren et al 2000). By calculating the effect of commuting time on job search and job moving, we get an indication of the willingness to accept longer commuting times. Moreover, by calculating the ratio of the effect of commuting time and the effect of wages, we can put a monetary value on this willingness to accept (Van Ommeren and Fosgerau 2009). The advantage of this approach is that we do not need to assume that labour markets are in equilibrium (Gronberg and Reed 1994). In other words, we do not need to assume that observed situations in the labour market reflect the best choices among all available alternatives.

Conceptually, our regression model distinguishes between telecommuting as a job asset, and telecommuting as a substitute for commuting physically. We investigate the effect of telecommuting on the acceptability of one-way commuting times by examining the interactions between a telecommuting dummy and commuting distance. The main effect of this telecommuting dummy tells us something about the intrinsic value of telecommuting. Several studies suggest for instance that working from home on designated, individual tasks may increase worker productivity (Bernardino 2017). Furthermore, the possibility to telecommute may increase the chance of matching between workers and employers if labour markets for telecommuting jobs indeed have a larger geographical scale.

In the second part of the analysis, we estimate the extent and direction of sorting bias. We compare the outcomes of an OLS commuting time regression to the results from an individual fixed effects approach that makes use of the panel dimension of the data. The latter method corrects for all time-invariant characteristics of people, including time-constant preferences. We control for time-variant confounders as much as possible by accounting for (changes in) monthly wage, the industry in which people work, the type of employment, whether or not individuals have a partner or kids at home, and the wage of the partner.

2.5 Results

Evidence for preference-based sorting – In this subsection we examine the difference in commuting preferences between telecommuters and non-telecommuters. We do this by estimating the effect of commuting time on job search and the propensity to change jobs, for both groups. We standardize this effect by the effects of wage on job search and mobility to obtain the marginal costs of one-way commuting time (MCC), measured as the average amount of daily wage people are willing to give up to shorten their (one-way) commute with 1 minute (Van Ommeren and Fosgerau 2009). First, in Figure 3 we show the bivariate relationship between commuting time and job search and mobility for the whole sample. It is clear that both the share of people looking for a job, and the share of people changing jobs within two years is positively related with commuting time. This confirms the intuitive notion that longer commutes are seen as a negative aspect of jobs.[9]

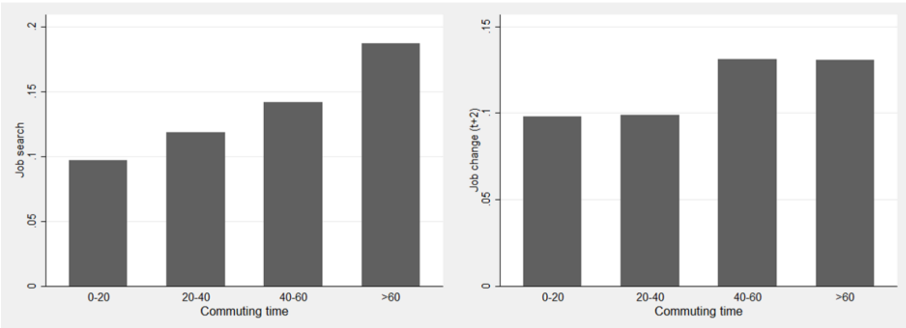


Figure 3: Bivariate relationships between commuting time and job search (l), and job mobility (r)

In Table 3 we estimate the daily MCC using the two distinct approaches. We follow the literature and use a random effects probit model to deal with potential heterogeneity among different individuals.[10] According to the job search model in column (1) commuting time has a greater effect on job search for telecommuters than for non-telecommuters. In monetary terms, non-telecommuters are willing to accept a 1 minute longer one-way commute for €2.63 per work day, while telecommuters are willing to accept a 1 minute longer commute for €3.80.[11] Note that this is in spite of the fact that, by definition, telecommuters commute less frequently, compared to non-

[9] We found no such clear bivariate patterns between telecommuting and job- search and mobility.

[10] This allows the error terms of the same individuals to be correlated over time (Van Ommeren and Fosgerau 2009).

[11] The MCC is derived from the ratio of the effects of commuting time and daily wage.

telecommuters, so the MCC per commuting trip may be even higher for telecommuters. Furthermore, according to this model age has a positive but marginally decreasing effect on the propensity to search, and higher educated people search more. The effect of telecommuting itself is insignificant.

In column (2) we estimate the same model with job mobility (changing jobs within two years) as the dependent variable. According to this model the MCC is €1.91 for non-telecommuters, and €2.15 for telecommuters. These values are lower than the estimates in the previous model. The ratio between these values is also lower (1.13 vs. 1.44). According to this model the effect of age on mobility is predominantly negative, and higher educated people seem more mobile, but not significantly so. The effect of telecommuting itself on job moving is not significant.

Table 3: Willingness to pay for commuting regressions

	(1) Job search	(2) Job mobility
Commute * No telecommuting	0.00494***	0.00330***
	(0.000927)	(0.00117)
Commute * Telecommuting	0.00712***	0.00372**
	(0.00128)	(0.00155)
Telecommuting	-0.0597	0.112
	(0.0671)	(0.0824)
Daily wage	-0.00187***	-0.00172**
	(0.000545)	(0.000787)
Firm size	4.75e-06	-2.89e-05
	(6.12e-06)	(1.77e-05)
Age	0.0755***	-0.0320*
	(0.0131)	(0.0170)
Age^2	-0.00114***	3.01e-05
	(0.000157)	(0.000207)
Female	0.0263	-0.0954*
	(0.0441)	(0.0570)
Partner	-0.217***	-0.177***
	(0.0475)	(0.0610)
Partner wage	-5.59e-06	1.58e-05**
	(6.71e-06)	(6.96e-06)
Children at home	0.0130	0.0329
	(0.0424)	(0.0533)
Basic education	-0.0764	-0.0290
	(0.152)	(0.160)
Higher education	0.0201	0.00369
	(0.151)	(0.158)
Vocational education	0.279*	0.166
	(0.154)	(0.162)
Bachelor degree	0.449***	0.210
	(0.160)	(0.172)
Constant	-1.841***	0.703
	(0.389)	(0.567)
Individual random effects	Yes	Yes
Year dummies	Yes	Yes
Control dummies	Yes	Yes
Observations	18,730	11,320
Individuals	7,497	4,481
Rho	0.335	0.243
Log likelihood	-6136	-3493
MCC non-telecommuters	€2.63	€1.91
MCC telecommuters	€3.80	€2.15
Relative difference	1.44	1.13

Notes: Robust std. errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Control dummies include 7 working days-, 10 industry-, and 5 job type dummies. MCC stands for Marginal cost of commuting, and should be interpreted as the daily willingness-to-pay for a one minute reduction in one-way commuting time.

Commuting time – In this subsection we estimate the effect of telecommuting on commuting time, controlling for preference-based sorting by employing individual fixed effects. We start with an OLS model and we compare the resulting estimates with the results of a fixed effects model. Because the dependent variable is in logs, 187 observations with 0 commuting time are excluded from the analysis, so we are left with 18,543 observations.

Table 4 shows the OLS results. In column (1) we use a telecommuting dummy that corresponds to our telecommuting definition. According to this model telecommuting results in a 11.7 percent longer commute on average.[12] Furthermore, a 10 percent increase in daily wage is associated with a 4.8 percent increase in commuting time, the level of education has a positive effect on commuting time, commuting patterns are gendered (women have about 7.6 percent shorter commutes), and individuals with children at home have about 7 percent shorter commutes. Except for the insignificant effect of age, these findings are in line with earlier results on Dutch commuting behaviour, which showed that females, and people with children, have shorter commutes on average, and people of higher socio-economic status commute longer (Van Ham 2002; Burger et al 2014). Employees of larger firms commute longer according to this model.

In column (2) we distinguish between 1, 2, 3, and more than 3 days of telecommuting per month. The results show that the positive effect found in the previous column is mainly driven by telecommuters that telecommute 2-4 days per month, as the effect of telecommuting 1 day per month is small and insignificant. The coefficients of the other variables are virtually unaffected by this alternative measure of telecommuting. In column (3) we measure telecommuting by the usual number of hours per week spent telecommuting. Arguably this is the most precise measure of telecommuting intensity. According to the model every 8 additional hours of telecommuting lead to a 5.2 percent increase in commuting time. The other coefficients are again similar to those in previous models.

[12] The coefficients in these log-linear models should be interpreted as an $(e^{\beta}-1)*100\%$ increase for every unit increase. For logged independent variables the coefficients can be interpreted as an elasticity.

Table 4: OLS commuting time regressions. Dependent variable: Commuting time (log)

	(1)	(2)	(3)
Telecommuting	0.111***		
	(0.0164)		
Telecommuting 1 day/month		0.0308	
		(0.0229)	
Telecommuting 2 days/month		0.163***	
		(0.0226)	
Telecommuting 3 days/month		0.0982***	
		(0.0217)	
Telecommuting > 3 days/month		0.0568**	
		(0.0231)	
Telecommuting weekly hours			0.00636***
			(0.00126)
Daily wage (log)	0.481***	0.479***	0.487***
	(0.0213)	(0.0214)	(0.0211)
Firm size	1.81e-05***	1.82e-05***	1.79e-05***
	(2.93e-06)	(2.94e-06)	(2.93e-06)
Age	-0.00395	-0.00441	-0.00376
	(0.00455)	(0.00455)	(0.00455)
Age^2	2.47e-05	3.10e-05	2.28e-05
	(5.34e-05)	(5.35e-05)	(5.34e-05)
Female	-0.0728***	-0.0725***	-0.0738***
	(0.0150)	(0.0150)	(0.0150)
Partner	-0.00568	-0.00634	-0.00502
	(0.0162)	(0.0162)	(0.0162)
Partner wage	-3.24e-06	-3.39e-06	-3.24e-06
	(2.47e-06)	(2.49e-06)	(2.48e-06)
Children at home	-0.0674***	-0.0669***	-0.0660***
	(0.0139)	(0.0139)	(0.0140)
Basic education	0.0223	0.0215	0.0217
	(0.0477)	(0.0477)	(0.0477)
Higher education	0.115**	0.113**	0.117**
	(0.0471)	(0.0472)	(0.0471)
Vocational education	0.172***	0.169***	0.178***
	(0.0483)	(0.0484)	(0.0483)
Bachelor degree	0.259***	0.253***	0.268***
	(0.0508)	(0.0509)	(0.0508)
Constant	-0.0207	0.00857	-0.0597
	(0.179)	(0.180)	(0.178)
Year dummies	Yes	Yes	Yes
Control dummies	Yes	Yes	Yes
Observations	18,543	18,543	18,543
R-squared	0.132	0.132	0.132

Notes: Robust std. errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In columns (1) and (2) No Telecommuting is the reference category. Control dummies include 7 working days-, 10 industry-, and 5 job type dummies.

In Table 5 we estimate the same models including individual specific fixed effects that correct for all time-invariant attributes of individuals, including preferences. Coefficients are estimated based on variation *within* individuals over time. The results from column (1) indicate that telecommuting leads to 5 percent longer commutes, rather than the 11.7 percent estimated in column one. Thus the extent of the bias due to sorting is positive (+128 percent) according to this specification. The fixed effects model results in several different coefficients compared to the OLS estimates. First, the effect of daily wage on commuting time is lower when accounting for time-invariant unobservables. This may for instance be driven by correlations between capability and labour mobility. Second, it seems that ageing does not significantly influence commuting time. Third, changes in firm size and having children at home have significant but smaller effects on commuting time, compared to the OLS model. Finally, while we see an increasing pattern in the effects of education on commuting, the estimates are not significant.

Column (2) is the fixed effects equivalent of Table 4, column (2). The results from this column show that compared to non-telecommuters, individuals that telecommute 1 day per month accept a 6.1 percent longer commute, those telecommuting 2 days per month a similar but lower 5.1 percent, those that telecommute 3 days do not have significantly longer commutes. This result is somewhat counter-intuitive as it suggests positive but decreasing effect of telecommuting on commuting time. It should however be noted that the only significant difference in coefficients between *consecutive* categories is the one between no telecommuting and telecommuting 1 day per month. Other coefficients in this model are similar to those in the previous column. Finally, in column (3) we estimate the effect of (changes in) the usual weekly hours spent telecommuting on (changes in) commuting time. The effect is estimated at a 3.5 percent increase in commuting time for every 8 additional weekly hours spent working at home, indicating a 50 percent upward bias due to preference-based sorting in the OLS estimate in column (3), Table 4.

From the analyses in this subsection we conclude that telecommuting significantly affects commuting time and overall, the bias induced by preference-based sorting of individuals into telecommuting is positive rather than negative, between 50 and 128 percent.[13] An explanation for this may be that overall, the (negative) bias induced by residential preferences is stronger than the (positive) bias due to commuting preferences. According to our results telecommuting allows people to accept 5 percent longer commutes on average, and for every 8 additional weekly hours spent working from home, people accept a 3.5 percent longer commute.

[13] The models with a telecommuting dummy suggest a 128 percent bias (0.0486 (FE) versus 0.111 (OLS)), the models with a continuous measure of weekly hours spent working from home suggest a 50 percent bias (0.00425 (FE) vs 0.00636 (OLS)).

Table 5: FE commuting time regressions. Dep. variable: Commuting time (log)

	(1)	(2)	(3)
Telecommuting	0.0486***		
	(0.0141)		
Telecommuting 1 day/month		0.0590***	
		(0.0187)	
Telecommuting 2 days/month		0.0494***	
		(0.0189)	
Telecommuting 3 days/month		0.0290	
		(0.0186)	
Telecommuting > 3 days/month		0.000951	
		(0.0185)	
Telecommuting weekly hours			0.00425***
			(0.000827)
Daily wage (log)	0.176***	0.176***	0.176***
	(0.0270)	(0.0270)	(0.0269)
Firm size	5.36e-06**	5.37e-06**	5.28e-06**
	(2.13e-06)	(2.13e-06)	(2.13e-06)
Age	0.00571	0.00582	0.00558
	(0.00745)	(0.00747)	(0.00745)
Age^2	2.83e-05	2.36e-05	2.94e-05
	(7.98e-05)	(8.00e-05)	(7.98e-05)
Partner	0.0942***	0.0956***	0.0952***
	(0.0276)	(0.0276)	(0.0276)
Partner wage	3.17e-07	3.04e-07	2.99e-07
	(1.82e-06)	(1.82e-06)	(1.82e-06)
Children at home	-0.0381**	-0.0382**	-0.0389**
	(0.0175)	(0.0175)	(0.0175)
Basic education	0.0644	0.0652	0.0644
	(0.0576)	(0.0577)	(0.0576)
Higher education	0.0683	0.0696	0.0686
	(0.0600)	(0.0600)	(0.0600)
Vocational education	0.0807	0.0826	0.0810
	(0.0641)	(0.0641)	(0.0640)
Bachelor degree	0.112	0.112	0.114
	(0.0716)	(0.0716)	(0.0715)
Indiv. fixed effects	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Control dummies	Yes	Yes	Yes
Observations	15,505	15,505	15,505
R-squared	0.797	0.797	0.797

Notes: Robust std. errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In columns (1) and (2) No Telecommuting is the reference category. Control dummies include 7 working days-, 10 industry-, and 5 job type dummies.

Sensitivity analysis – In this subsection we subject our results to several sensitivity checks. We employ a stricter identification approach based on the timing and intensity of telecommuting, and two alternative identification approaches using a Lagged Dependent Variable Model and a Long Difference model.

First, we analyse individuals that telecommuted *at some point* during the study period. For these individuals we know that they are able to telecommute, so the decision of whether or not to telecommute, and for how many days and hours, suffers less from potentially omitted variables and self-selection. The drawback of this approach is that the external validity of the results is limited, because the effects we obtain in principle only apply to those able to telecommute. The results of these *timing* regressions, presented in Table 6, Appendix A, are comparable to the estimates from Table 5.

Second, we use an identification method based on a lagged dependent variable, proposed by (Angrist and Pischke 2008) as a robustness check for fixed effects models. Specifically, instead of assuming that telecommuting is randomly assigned across respondents conditional on unobserved time-invariant characteristics, this method assumes random assignment conditional on the 1 year lag of commuting distance. Checking the robustness of our results to this assumption makes sense because commuting time is time-varying, and those who start to telecommute may do so because over time, they have become tired of their long commute. This method thus corrects for a different type of selection bias (based on commuting history), and as Angrist and Pischke (2008) note, the results of fixed effects and lagged dependent variable models can be regarded as bounding the effect of interest, depending on the type of selection bias that is controlled for. In Table 7 we show the results of the models based on this identification strategy, and it is reassuring that the outcomes of this analysis are remarkably similar to the estimates from Table 5.

Third, we estimate a 'long-differences' model in which we only include the first and last year of our data, controlling for time-invariant characteristics of respondents. This approach is only based on 516 respondents for whom we have data for both years. The idea behind this robustness check is that it takes time to get used to new technologies and situations, and to adjust behaviour in housing and labour markets. The results, presented in Table 8, suggest that our estimates based on short-run behaviour may be somewhat conservative. Respondents that have picked up telecommuting between 2002 and 2014 have on average 32 percent longer commuting times, and every 8 hours increase in weekly telecommuting hours during this period resulted in 20.5 percent longer commuting times. It should however not be ruled out that these high estimates are the result of sample selection effect: respondents with high residential mobility are less likely to be contacted over multiple years, but they are more prone to shorten their commutes by moving residence (Van Ommeren 1998).

Finally, we investigate whether there are nonlinearities in the effect of hours working from home on commuting time. We do this by estimating a dummy specification, in which the variable denoting weekly hours spent working from home is divided up into 7 categories (0, 0-8, 8-16, 16-24, 42-32, 32-40, and 40+). The model, presented in the column (1) of Table 9 in Appendix A, is an alternative version of Table 5 column (3), and the marginal effects of the dummies are depicted in Figure 4. While the graph does not show significant effects of telecommuting categories 16-24 and 32-40, the overall pattern of point estimates follows a somewhat linear pattern, at least up until the 24-32 hours mark. Considering the observed pattern, and the significance of the other dummies, we may conclude that the parametric approach in Table 5 column (3) is a reasonable approximation of the non-parametrically estimated shape of the relationship, and as it is more efficient it has our preference. In Table 9 columns (2-4) we show the results of this dummy specification using the other identification strategies, and in all specifications the pattern is roughly linear until 24-32 hours.

In conclusion, the main result – that the effect of telecommuting on commuting time remains positive after controlling for sorting – is robust to identification based on the timing and intensity of telecommuting, and to an identification strategy based on a lagged dependent variable. A 'long-difference' model suggests our estimates are somewhat conservative, and a linear specification of average weekly hours working from home is not problematic.

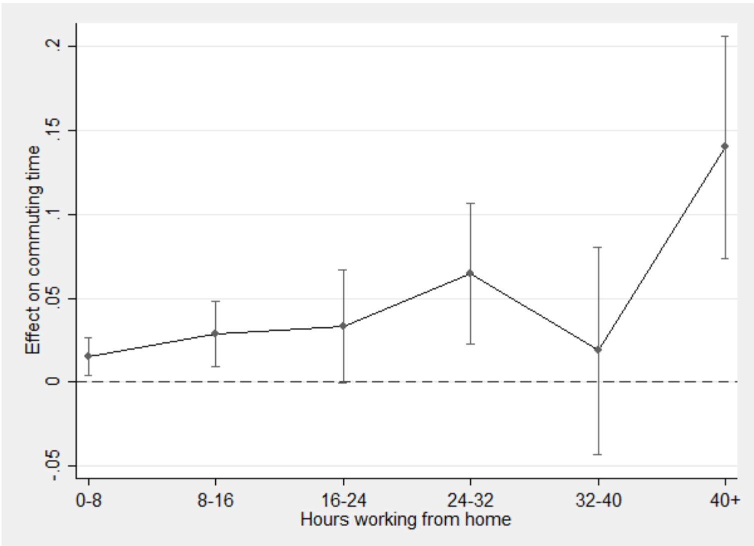


Figure 4: Non-linear effect of hours working from home. No telecommuting is the reference category.

2.6 Conclusion

This paper shows that the relationship between telecommuting and commuting time suffers from a bias due sorting based on residential- and commuting preferences, that should be accounted for. On the one hand, the effects of commuting time on labour search and labour mobility suggest that telecommuters have a higher value of one-way commuting time, despite their lower commuting frequency, which would lead to a downward bias in OLS estimates of the effect of telecommuting on commuting time. However, using fixed effects to control for (stationary) preferences of individuals, our analysis shows that OLS estimates are biased upward in the range of 50 to 128 percent. This suggests that both residential- and commuting preferences distort OLS findings, and that the bias due to residential sorting is stronger.

Our preferred estimates suggest that moving from a situation with no telecommuting, telecommuting allows people to accept 5 percent longer commuting times on average, and every additional 8 weekly hours of working from home are associated with 3.5 percent longer commuting times. The main result of this study is that the effect of telecommuting on commuting time remains positive and significant, after controlling for individual preferences. This result is robust to a number of sensitivity checks in which we apply alternative identification methods, and allow for a non-linear effect of weekly hours working from home.

There are some limitations to the research approach in this paper that could inspire further research. First, we only analyse commuting time, and not commuting distance, because we lack the proper data. Several studies do investigate the effect of telecommuting on both commuting time and distance, and generally find greater elasticities for distance (Andreev et al 2010; Zhu 2012). Theoretically, focusing on commuting time, and ignoring commuting distance, is justified by assuming that commuting speed is optimally chosen (Van Ommeren and Fosgerau 2009). Second, while we take into account the effects of self-selection, we ignore the possibility that long commutes trigger telecommuting. We argue that whether or not increases in commuting time trigger telecommuting, or telecommuting triggers longer commutes is irrelevant, because it both entails that telecommuters are willing to accept a longer commute, all else equal. However, future research may be directed at finding instruments for *changes* in telecommuting, unrelated to changes in commuting time to assess the one-way causal effect. Finally, our fixed effects model corrects for the bias induced by preference-based sorting only to the extent that these preferences are *time-invariant*. We are hopeful that our extensive list of control variables captures remaining changes in these preferences, and we note that including fixed effects may at least capture more of the sorting bias than OLS models.

In line with earlier work, our results suggest that the travel savings made by working one or several days at home are not fully offset by the positive effects on commuting distance alone (Jiang 2008; Andreev et al 2010; Zhu 2012). Beyond that, this paper stresses the effects of telecommuting on the geographical territory of labour markets. Next to reducing the labour accessibility gap between central and remote areas, telecommuting may also allow the externalities associated with the size of local labour markets, including improved searching and matching and less unfilled vacancies (Moretti 2011), to be increasingly generated across greater geographical areas, and through wider infrastructure networks (Burger and Meijers 2016). Further research may focus on the welfare effects associated with a wider geographical extent of labour markets.

References

Alonso, W. (1974). Location and land use: Toward a general theory ofland rent. Cambridge, Massachusettes: Harvard University Press.

Andreev, P., Salomon, I., & Pliskin, N. (2010). Review: State of teleactivities. Transportation Research Part C: Emerging Technologies, 18(1), 3–20. https://doi.org/10.1016/j.trc.2009.04.017

Angrist, J. D., & Pischke, J. S. (2008). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press.

Bernardino, A. (2017). Telecommuting: Modelling the Employer's and the Employee's Decision-Making Process. London: Routledge.

Büchel, F., & Van Ham, M. (2003). Overeducation, regional labor markets, and spatial flexibility. Journal of Urban Economics, 53(3), 482–493. https://doi.org/10.1016/S0094-1190(03)00008-1

Burger, M. J., & Meijers, E. J. (2016). Agglomerations and the Rise of Urban Network Externalities. Papers in Regional Science, 95(1), 5–15. https://doi.org/10.1111/pirs.12223

Burger, M. J., Meijers, E. J., & Van Oort, F. G. (2014). Multiple perspectives on functional coherence: Heterogeneity and multiplexity in the randstad. Tijdschrift Voor Economische En Sociale Geografie, 105(4), 444–464. https://doi.org/10.1111/tesg.12061

Cairncross, F. (1997). The death of distance: how the communication revolution is changing our lives. Boston, Massachusettes: Harvard Business School Press.

De Graaff, T. (2004). On the Substitution and Complementarity between Telework and Travel: A Review and Application (Vrije Universiteit Amsterdam). Retrieved from http://degree.ubvu.vu.nl/repec/vua/wpaper/pdf/20040016.pdf

Glaeser, E. L., & Kahn, M. E. (2004). Sprawl and urban growth. In Handbook of Regional and Urban Economics (Vol. 4, pp. 2481–2527). https://doi.org/10.1016/S1574-0080(04)80013-0

Gronberg, T., & Reed, W. (1994). Estimating workers' marginal willingness to pay for job attributes using duration data. Journal of Human Resources, 29(3), 911–931. https://doi.org/10.2307/146258

Groot, S. P. T., de Groot, H. L. F., & Smit, M. J. (2014). Regional wage differences in the Netherlands: Micro evidence on agglomeration externalities. Journal of Regional Science, 54(3), 503–523. https://doi.org/10.1111/jors.12070

Gubins, S., & Verhoef, E. T. (2014). Dynamic bottleneck congestion and residential land use in the monocentric city. Journal of Urban Economics, 80, 51–61. https://doi.org/10.1016/j.jue.2013.09.001

Isacsson, G., Karlström, A., & Swärdh, J.-E. (2013). The value of commuting time in an empirical on-the-job search model – an application based on moments from two samples. Applied Economics, 45(19), 2827–2837. https://doi.org/10.1080/00036846.2012.678981

Jiang, Y. (2008). The Impact of Telecommuting on the Journey to Work : A Two-Sample Instrumental Variables Approach (University of Maryland). Retrieved from http://drum.lib.umd.edu/bitstream/handle/1903/8506/umi-umd-5557.pdf

Kim, S.N., Mokhtarian, P. L., & Ahn, K.-H. (2012). The Seoul of alonso: New perspectives on telecommuting and residential location from South Korea. Urban Geography, 33(8), 1163–1191. https://doi.org/10.2747/0272-3638.33.8.1163

Kim, Seung Nam, Choo, S., & Mokhtarian, P. L. (2015). Home-based telecommuting and intra-household interactions in work and non-work travel: A seemingly unrelated censored regression approach. Transportation Research Part A: Policy and Practice, 80, 197–214. https://doi.org/10.1016/j.tra.2015.07.018

Kolko, J. (2012). Broadband and local growth. Journal of Urban Economics, 71(1), 100–113. https://doi.org/10.1016/j.jue.2011.07.004

Lund, J. R., & Mokhtarian, P. L. (1994). Telecommuting and residential location: theory and implications for commute travel in monocentric metropolis. Transportation Research Record, (1463), 10–14. Retrieved from http://onlinepubs.trb.org/Onlinepubs/trr/1994/1463/1463-002.pdf

Mokhtarian, P. L., Salomon, I., & Choo, S. (2005). Measuring the measurable: Why can't we agree on the number of telecommuters in the U.S.? Quality and Quantity, 39(4), 423–452. https://doi.org/10.1007/s11135-004-6790-z

Moretti, E. (2011). Local labor markets. In Handbook of Labor Economics (Vol. 4, pp. 1237–1313). https://doi.org/10.1016/S0169-7218(11)02412-9

Muhammad, S., de Jong, T., & Ottens, H. F. L. (2008). Job accessibility under the influence of information and communication technologies, in the Netherlands. Journal of Transport Geography, 16(3), 203–216. https://doi.org/10.1016/j.jtrangeo.2007.05.005

Muhammad, S., Ottens, H. F. L., Ettema, D., & Jong, T. (2007). Telecommuting and residential locational preferences: A case study of the Netherlands. Journal of Housing and the Built Environment, 22(4), 339–358. https://doi.org/10.1007/s10901-007-9088-3

Muth, R. F. (1969). Cities and Housing. Chicago: University of Chicago Press.

Naisbitt, J. (1994). The global paradox. New York: William Morrow & Co.

Nilles, J. M. (1991). Telecommuting and urban sprawl: mitigator or inciter? Transportation, 18(4), 411–432. https://doi.org/10.1007/BF00186567

Rhee, H. J. (2008). Home-based telecommuting and commuting behavior. Journal of Urban Economics, 63(1), 198–216. https://doi.org/10.1016/j.jue.2007.01.007

Salomon, I. (1985). Telecommunications and travel: substitution or modified mobility? Journal of Transport Economics and Policy, Vol. 19, 219–235. https://doi.org/10.1080/09595238100185051

SCP. (2016). Arbeidsaanbodpanel 1985 t/m 2014. https://doi.org/10.17026/dans-x8q-46p7

Toffler, A. (1980). The third wave. New York: Bantam Books.

Van Ham, M. (2002). Job access, workplace mobility, and occupational achievement. Delft: Eburon.

Van Ommeren, J. (1998). A test for randomly sampling using duration observations. Applied Economics Letters, 5(4), 243–246. https://doi.org/10.1080/135048598354906

Van Ommeren, Jos, & Fosgerau, M. (2009). Workers' marginal costs of commuting. Journal of Urban Economics, 65(1), 38–47. https://doi.org/10.1016/j.jue.2008.08.001

Van Ommeren, Jos, Van Den Berg, G. J., & Gorter, C. (2000). Estimating the Marginal Willingness To Pay for Commuting. Journal of Regional Science, 40(3), 541–563. https://doi.org/10.1111/0022-4146.00187

Van Wee, B., Geurs, K. T., & Chorus, C. (2013). Information, communication, travel behavior and accessibility. Journal of Transport and Land Use, 6(3), 1–16. https://doi.org/10.5198/jtlu.v6i3.282

Vermeulen, W., & Rouwendal, J. (2007). Housing Supply and Land Use Regulation in the Netherlands. In Tinbergen Institute Discussion Paper (No. 3; Vol. 58). https://doi.org/10.2139/ssrn.1003955

Webber, M. M. (1963). The Post-City Age. Daedalus, 97(4), 1091–1110. Retrieved from http://www.jstor.org/stable/20013413

Zhu, P. (2012). Are telecommuting and personal travel complements or substitutes? Annals of Regional Science, 48(2), 619–639. https://doi.org/10.1007/s00168-011-0460-6

Appendix A: Supplementary tables

This appendix contains tables with regression results from the sensitivity analyisis. Table 6 presents models based on a sample that only consists of telecommuters. Table 7 shows the lagged dependent variable models. Table 8 shows the `long differences' models, and Table 9 shows models in which we allow for a nonlinear specification of weekly telecommuting hours.

Table 6: Sensitivity regressions I: Telecommuting sample

	(1)	(2)	(3)
Telecommuting	0.0453***		
	(0.0148)		
Telecommuting 1 day/month		0.0590***	
		(0.0197)	
Telecommuting 2 days/month		0.0482**	
		(0.0200)	
Telecommuting 3 days/month		0.0311	
		(0.0197)	
Telecommuting > 3 days/month		0.0125	
		(0.0199)	
Telecommuting weekly hours			0.00407***
			(0.000868)
Indiv. fixed effects	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Control variables	Yes	Yes	Yes
Observations	8,309	8,309	8,309
R-squared	0.773	0.773	0.773

Notes: Robust std. errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In columns (1-2) No Telecommuting is the reference category. Control variables are the same as in Tables (4-5).

Table 7: Sensitivity regressions II: Lagged dependent variable

	(1)	(2)	(3)
Telecommuting	0.0526***		
	(0.0128)		
Telecommuting 1 day/month		0.0334*	
		(0.0180)	
Telecommuting 2 days/month		0.0783***	
		(0.0180)	
Telecommuting 3 days/month		0.0411**	
		(0.0165)	
Telecommuting > 3 days/month		0.0268	
		(0.0197)	
Telecommuting weekly hours			0.00398***
			(0.00101)
Commuting time at t-1 (log)	0.726***	0.726***	0.725***
	(0.0101)	(0.0101)	(0.0102)
Indiv. fixed effects	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Control Variables	Yes	Yes	Yes
Observations	11,090	11,090	11,090
R-squared	0.617	0.618	0.618

Notes: Robust std. errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In columns (1-2) and (4) No Telecommuting is the reference category. Control variables are the same as in Tables (4-5).

Table 8: Sensitivity regressions III: Long differences

	(1)	(2)	(3)
Telecommuting	0.277***		
	(0.0777)		
Telecommuting 1 day/month		0.224**	
		(0.0911)	
Telecommuting 2 days/month		0.187	
		(0.121)	
Telecommuting 3 days/month		0.234*	
		(0.132)	
Telecommuting > 3 days/month		0.0442	
		(0.0987)	
Telecommuting weekly hours			0.0233***
			(0.00416)
Indiv. fixed effects	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Control variables	Yes	Yes	Yes
Observations	1,032	1,032	1,032
R-squared	0.759	0.756	0.767

Notes: Robust std. errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In columns (1-2) No Telecommuting is the reference category. Control variables are the same as in Tables (4-5).

Table 9: Sensitivity regressions IV: Nonlinear telecommuting hours

	(1)	(2)	(3)	(4)
Telecommuting weekly hours 0-8	0.0348***	0.0349**	0.0375**	0.115
	(0.0133)	(0.0150)	(0.0176)	(0.0743)
Telecommuting weekly hours 8-16	0.0659***	0.0609**	0.0841***	0.317***
	(0.0230)	(0.0260)	(0.0310)	(0.119)
Telecommuting weekly hours 16-24	0.0767*	0.0696	0.0773	0.627***
	(0.0393)	(0.0426)	(0.0534)	(0.225)
Telecommuting weekly hours 24-32	0.149***	0.142***	0.159***	0.685*
	(0.0491)	(0.0518)	(0.0502)	(0.386)
Telecommuting weekly hours 32-40	0.0431	0.0409	0.123	1.024**
	(0.0722)	(0.119)	(0.0784)	(0.438)
Telecommuting weekly hours 40+	0.323***	0.319***	0.0526	1.812***
	(0.0778)	(0.104)	(0.115)	(0.493)
Commuting time at t-1 (log)			0.101***	
			(0.0281)	
Model	FE	TC Sample	LDV	LD
Indiv. fixed effects	Yes	Yes	No	Yes
Year dummies	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes
Observations	15,505	8,309	6,110	1,032
R-squared	0.797	0.773	0.850	0.766

Notes: Robust std. errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No Telecommuting is the reference category. Control variables are the same as in Tables (4-5). FE stands for Fixed Effects, TC stands for Telecommuting, LDV stands for Lagged Dependent Variable, and LD stands for Long Differences.

3. Working from home and commuting: Heterogeneity over time, space, and occupations

Published as IZA Discussion Paper 12578 (2019)
Written by Duco de Vos, Maarten van Ham and Evert Meijers.

Abstract

Teleworking may increase the willingness to accept a longer commute. This paper presents new evidence of the effect of teleworking on the length of commutes. We use novel panel data from the Netherlands, for the years 2008-2018, and find stronger effects compared to studies that use older data. Between 2008 and 2018 however, the effect was remarkably stable: workers that started teleworking increased their commutes by 12 percent on average. We analyse heterogeneity in the effect of teleworking on commuting across different levels of urbanization and across occupations. This study stresses the effects of teleworking on the geographical scale of labour markets, and provides important inputs for policymakers that aim to promote teleworking.

3.1 Introduction

Advances in information and communication technology allow an increasingly large number of workers to work from home on or more days per week. This has caught the attention of policy-makers that aim to reduce CO2 emissions, in order to fulfil the reductions agreed upon in the Paris Agreement.[1] However, when people work from home one day in the week, they may be willing to travel to work for longer distances on the other days. This casts doubts on the net effect of working from home. Policies that promote teleworking may actually have adverse effects on commuting distance. At the same time, longer commuting distances imply an increase in the size of labour markets, and increasing levels of job-accessibility in remote areas. As a consequence, the externalities associated with large labour markets, such as better matching and higher productivity, may spread over larger areas.

Despite the substantial interest in policy circles for the possible effects of teleworking, only a handful of studies explore whether it leads to longer commutes. Most of these studies note that the proper data to identify the direction of causality between teleworking and commuting is lacking. Two recent studies that use (synthetic) panel data are exemptions to this, and show that those who start

[1] See for instance the proposed measures for the Dutch Climate Agreement (https://www.klimaatakkoord.nl/documenten/publicaties/2018/07/10/hoofdlijnen-compleet).

teleworking tend to accept longer commutes. De Vos et al. (2018) use data from 2002 until 2014 and show that workers tend to accept a 5 percent longer commute after adopting teleworking. Gubins et al. (2019) show that between 1996 and 2010, teleworkers increased their commutes 5-9 times more than non-teleworkers.

Despite such first findings, there are still important gaps in the literature on teleworking and commuting. First, there is a general lack of studies using longitudinal data, apart from the two studies cited above (De Vos et al. 2018, Gubins et al. 2019). The estimates from panel data studies should approach the causal effect of teleworking on commuting because they are based on changes in teleworking and commuting, and because time-invariant characteristics are controlled for by fixed effects (Angrist and Pischke, 2008). Additional evidence based on different panel data sources adds to the robustness of the findings of the few studies that used panel data. Second, it is still unclear whether the effects of teleworking on commuting are growing as telework technologies progress and working from home becomes a closer substitute to working at the job location (Pajevic and Shearmur, 2017). Third, there has been little attention for spatial heterogeneity in the effects of teleworking on commuting. One may for instance expect that the effects of teleworking are smaller in urban areas, with high-quality infrastructure and thick labour markets, compared to rural areas (Rhee, 2008). Finally, there is a lack of clear and specific estimates for different job-types, that can be used as inputs for teleworking policies. This is important in order to make realistic estimates of the effects that teleworking policies will have on CO2 emissions, and labour market dynamics.

In this paper we aim to fill these gaps by using the Longitudinal Internet Studies for the Social Sciences (LISS) data from the Netherlands, that covers the years 2008 until 2018. We use this data set to estimate the effect of teleworking on commuting, and we compare our findings with the handful of previous studies. Furthermore, we broaden the scope by analysing whether the effects differ over time, space, and different occupations. We present an overview of effect sizes that can be used by policymakers to make projections about the environmental, economic and social effects of schemes that promote teleworking.

We find that on average, workers that start teleworking increase their commute by 12 percent. This estimate is higher than previous estimates with Dutch data from 2002 until 2014. When we split our study period 2008-2018 in early and later years, we find that the magnitude of the effect of teleworking on commuting has remained remarkably stable. Regarding spatial heterogeneity, we find strong and significant effects across all levels of urbanization, and the effects are strongest in moderately urban areas. Our results further show that not all occupations show the same relationship between teleworking and commuting, and the effects are especially sizeable for jobs that require academic skills and for lower-order administrative jobs.

The rest of this paper is structured as follows. In Section 2 we discuss earlier literature on teleworking and commuting, and we position our work in this literature. In Section 3 we introduce the methods used in the estimations, and we discuss the data. Section 4 presents the results, and Section 5 concludes.

3.2 Literature review

Previous studies - Empirical research on the relationship between home-based teleworking and commuting started with Nilles (1991), who used data from the California Pilot Telecommuting Project that was done with State employees. One of the main findings of this study was that telework was associated with relocations to places farther from the workplace. Around this time, several other teleworking studies were conducted based on data from such experiments (e.g. Hamer et al. 1991, Henderson and Mokhtarian 1996, Koenig et al. 1996, see Andreev et al. 2010 for an overview), although few have analysed the effects on commuting distance. In a review paper by De Graaff (2004) it is noted that up until then, studies of teleworking and commuting did not rule out a positive relationship.

With the increased accessibility of large-scale surveys, research on teleworking and commuting moved beyond analysing the results from one-off experiments. Examples include Jiang (2008), Zhu (2012,2013), Kim et al. (2012), and Melo and De Abreu e Silva (2017). These studies pay explicit attention to the problem of estimating a causal relationship between teleworking and commuting. Next to omitted variables, these problems relate to reverse causality (long commutes leading to teleworking) and sorting (people who dislike commuting or who prefer rural living may telework more). These studies do not however resolve these empirical problems. For instance: Zhu (2012) uses 'internet use at home' as an instrument for teleworking, but does not consider that individual abilities may be correlated with the use of ICT (Gubins et al. 2019); Kim et al. (2012) concur that "causality cannot be established with these data" (p. 1163); and Melo and de Abreu Silva (2017) note that "data-related limitations did not allow us to address issues of selection and/or simultaneity bias" (p. 1).

Two more recent studies shed new light on the relation between teleworking and commuting. Using Dutch panel data, De Vos et al. (2018) find that workers that start teleworking increase their commute by 5%. Their model employs individual fixed effects, so the estimates are based on time variation, and this method should deal with biases due to sorting. Gubins et al. (2019) have a slightly different angle, and estimate the effect of the penetration of teleworking across 'professions' on average commuting times within these professions. They use survey data from 1996 and 2010, and employ propensity score matching to construct credible counterfactuals, essentially creating a synthetic panel. While they do not find differences in average commuting times between treated (teleworking) and untreated (non-teleworking) professions, their results show that teleworkers increased their commutes 5-9 times more than non-teleworkers (12.07 vs 1.32 or 2.3 km).

Contributions to the literature - We contribute to the literature in four ways. First, we reproduce the results of De Vos et al. (2018) using more recent data from a different panel survey. Over the last decades a literature on teleworking and commuting has emerged, in which it has been common to re-estimate earlier results with different data (see the literature reviews of De Graaff 2004 and Andreev et al. 2010). Given the increasingly stressed importance of reproduction of research findings, this study continues this tradition.

Second, we assess whether the effects of teleworking on commuting have grown stronger or weaker over the years. We aim to answer the question whether the increases in *workplace mobility* in recent years, due to inter alia advancements in technology and increasing flexibility and temporality of contracts (Pajevic and Shearmur 2017), have made working from home a better substitute for working at the job location, and teleworkers more footloose. As changes in the relationship between teleworking and commuting may equally be influenced by for instance housing market dynamics, this is mainly an exploratory exercise.

Third, we analyse geographic variation in the effect of teleworking on commuting by estimating the effect separately across different levels of urbanization. A modelling exercise by Rhee (2008) suggests that whether or not people telecommute, and to what extent this impacts commute lengths is dependent on the residential location (distance to CBD), together with the degree to which employment is centralized, the substitutability of home-work and on-site work, and congestion levels. In a setting such as the Netherlands, with a polycentric structure of proximate cities, telework may not only serve as a means to sustain a long commute to the nearby city for suburban or rural residents, but it may equally allow inhabitants of one city to sustain a longer commute to another city (Muhammad et al. 2008).

Finally, we pay attention to the differences in teleworking penetration across professions, by estimating the effects of teleworking on commuting separately for different (groups of) occupations. A number of studies has paid attention to heterogeneity in teleworking across occupations. For instance Vilhelmson and Thulin (2016) show that in Sweden teleworking is still concentrated in higher occupations in the advanced services sector. Studies such as Melo and De Abreu e Silva (2017) and Kim et al. (2015) focus on particular sets of occupations, such as non-agricultural or white-collar workers. Gubins et al. (2019) use variation in teleworking between professions to estimate the effects of teleworking on commuting distance. Our study is the first to include specific estimates for different occupations, that can be used as inputs for telework policies.

3.3 Empirical strategy

Methodology - We will start our analysis by reproducing the results from De Vos et al (2018), with our data set that contains more recent data. The main regression model is as follows:

$$Commuting\ time_{it} = \alpha Telework_{it} + \beta X_{it} + \eta_i + \theta_t + \epsilon_{it} \quad (1)$$

where X_{it} is a vector of control variables, η_i are individual fixed effects, θ_t are time fixed effects, and ϵ denotes the error term. The individual fixed effects control for all time-invariant characteristics of persons. The control variables include personal characteristics and education, occupation, and industry dummy variables.

In their sensitivity analysis, De Vos et al. (2018) estimate two different models. First, they estimate a lagged dependent variable (LDV) model. According to Angrist and Pischke (2008) the results of fixed effects models on the one hand, and LDV models on the other, provide boundaries of the effect of interest, depending on the type of selection bias at force. This model includes the 2-year lag of commuting time, and is as follows:

$$Commuting\ time_{it} = \alpha Telework_{it} + \beta X_{it} + \gamma Commuting\ time_{i,t-2} + \epsilon_{it} \quad (2)$$

Second they estimate a long-differences model using only the first and last year of the data, as it may take time for the effects of teleworking to take full effect. We follow De Vos et al. (2018) in these two robustness checks.

Next we examine the heterogeneity of the effect of teleworking on commuting. To this end we estimate fixed effects models that include an interaction between teleworking and time periods, (groups of) occupations and 5 levels of urbanization (*stedelijkheid* in Dutch) of the municipality of residence. The data source that we use has a measure of *professions* that is comparable to the ISCO-08 standard of occupations, but not the same. Table 4 in Appendix A shows the correspondence between the LISS measure of professions and the ISCO-08 standard. The LISS measure does not include a separate category for military workers, it does not distinguish between clerical, and sales workers, but it does distinguish between higher- and intermediate management. The measure of urbanization we use is the only spatial indication present in the data, and it corresponds to a standard measure used by Netherlands Statistics, based on the density of addresses in the municipality. It distinguishes 'Very strongly urban areas' (2500 addresses/km^2 or more), 'Strongly urban areas' (1500-2500/km^2), 'Moderately urban areas' (1000-1500 addresses/km^2), 'Slightly urban areas' (500-1000/km^2) and 'Non-urban areas'. Figure 5 in Appendix A shows a map of levels of urbanization across Dutch municipalities.

Data description - The data used in this study comes from the Dutch Longitudinal Internet Studies for the Social Sciences (LISS). This panel has been running since 2008, and we use the 11 available waves up until 2018. The results of this study should be interpreted keeping in mind the specifics of the context in which the data was gathered. First, the Netherlands is a country characterized by a polycentric urban structure, where it is relatively common to live in one city and work in another city. In the Netherlands teleworking is thus not necessarily most attractive for residents of suburban or rural areas. Second, teleworking is a widespread phenomenon due to a high penetration of fast internet among other things (Muhammad et al. 2008), and in this sense the results of this study may be used to forecast the effects of teleworking in countries that still have lower levels of connectivity and teleworking. Finally, house price dynamics have been extreme between 2008 and 2018. They have plummeted between 2008 and 2012, and risen fast beyond 2012, especially in urban areas. High urban house prices may have pushed workers to use telework as a substitute for living in urban areas close to the workplace.

LISS is a panel survey that is filled out online, but respondents are selected based on a representative sample from the population registers, and households that could not otherwise participate were offered a pc and internet connection (Scherpenzeel and Das, 2010). The panel consists of 6000-7500 respondents every year. Our estimation sample consists of employed respondents between 18 and 65 years old, with non-missing values for education, occupation, industry, and hours worked. This leads to a sample of 22,577 observations. Whereas in the literature it has been common to exclude certain occupations (for instance agrarian occupations), one of the purposes of this work is to examine the heterogeneity between different types of occupations. Therefore all occupations are included in the estimation sample.

The survey specifically asks respondents if they have a ‘working at home day’, and whether this occurs less than one, one, or more times a week. Our definition of teleworkers is based on this question, and we consider every respondent with a ‘working at home day’ as a teleworker. Our definition of teleworkers is comparable to an often used definition given by Mokhtarian et al. (2005), in that we consider employees that replace the commute by working at home. It differs in the sense that we do not include teleworkers who work from other locations than home, and that it is ambiguous about the usage of ICT. Figure 1 shows the share of teleworkers for every year in our sample. A stark increasing pattern is visible between 2008 and 2016, which shows that teleworking is still on the rise.

Figure 2 shows that the spatial variation of growth in teleworking, across levels of urbanization, is limited. That said, growth rates have been especially high in ‘Very strongly urban’ and ‘Moderately urban areas’. Whether teleworking is feasible seems much more dependent on the type of occupation. Figure 3 shows a clear distinction between occupations where teleworking grew, and other occupations. Teleworking grew in ‘Higher academic’, ‘higher supervisory’, ‘Intermediate academic’, ‘Intermediate supervisory and commercial’, and ‘Other mental work’. It decreased or remained stable in manual, unskilled, and agrarian work.

Due to data limitations the only feasible measure of commuting in the data is one-way commuting time. The survey asks to state how many minutes it usually takes to travel between home and the workplace. Figure 4 shows that the distribution of this variable differs across people with different telework intensities. The starkest difference is between those that telework less than 1 day a week or not at all, and those that telework 1 day per week or more. The lack of information on commuting distance should not be problematic, and the use of commuting time can be justified by the (reasonable) assumption that commuting speed is optimally chosen (Van Ommeren and Fosgerau 2009, De Vos et al. 2018).

Table 1 shows the summary statistics. On average, 16.7 percent of the sample consists of teleworkers, and most of these telework less than a day per week on average. The average commuting time is 27.72 minutes. We also show summary statistics of control variables. We include control variables that are customary in this literature, on age, sex, having children or a partner, work duration, education, occupation type and industry. We lack reliable data on income and firm size, but we are confident that our use of individual fixed effects and industry controls will limit biases due to these omitted variables. In Table 5 in the Appendix, summary statistics are separated out for the different telework categories. From this table it can be concluded that there is considerable heterogeneity between the groups in terms of commuting time, education level and occupation type, but not so much in demographic characteristics or hours worked.

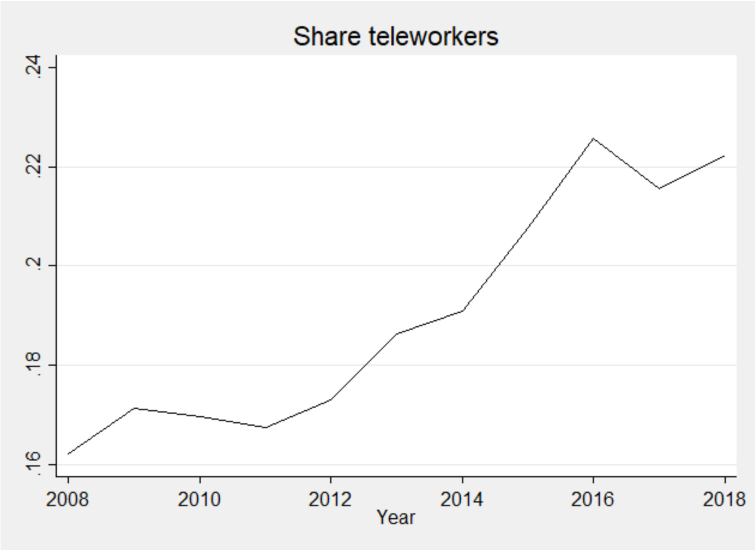


Figure 1: Share of teleworkers over time

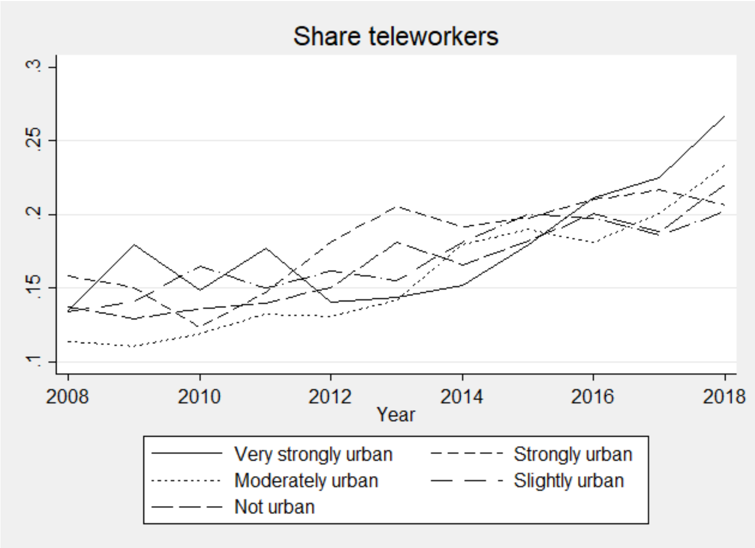


Figure 2: Share of teleworkers across urban levels over time

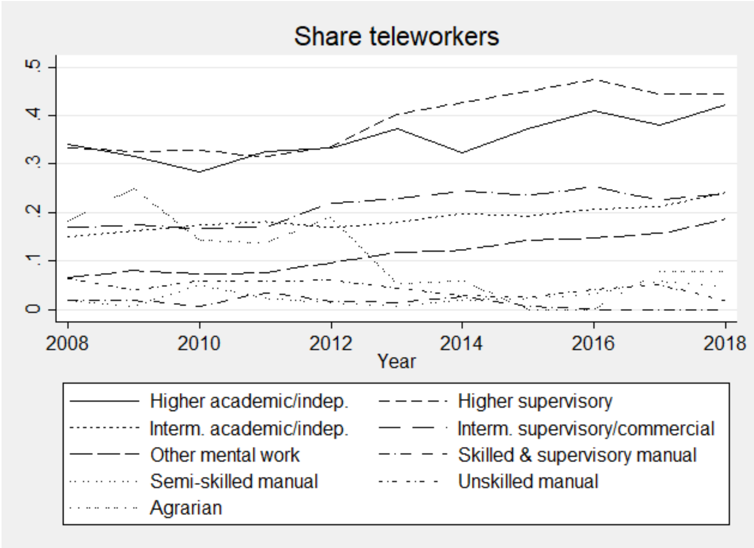


Figure 3: Share of teleworkers across occupations over time

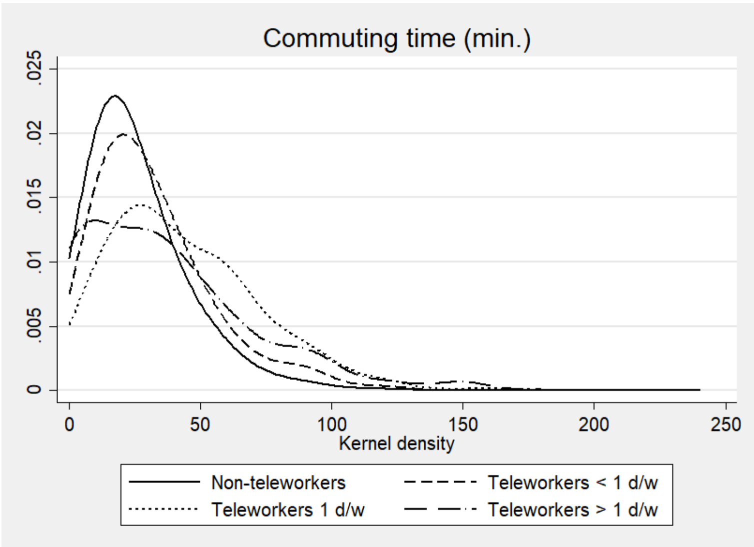


Figure 4: Distribution of commuting time

Table 1: Summary statistics

VARIABLES	(1) mean	(2) sd	(3) min	(4) max
Telework dummy	0.167	0.373	0	1
No telework	0.829	0.377	0	1
Telework < 1 d/w	0.0878	0.283	0	1
Telework 1 d/w	0.0569	0.232	0	1
Telework > 1 d/w	0.0263	0.160	0	1
Commuting time (min)	27.72	21.67	0	240
Age	44.55	11.21	18	65
Female	0.360	0.480	0	1
Children	0.462	0.499	0	1
Partner	0.758	0.428	0	1
Weekly hours worked	31.81	9.441	0	137
Basic education	0.236	0.425	0	1
Lower education	0.367	0.482	0	1
Medium education	0.257	0.437	0	1
Higher education	0.139	0.346	0	1
Occupation				
Higher academic/independent	0.0786	0.269	0	1
Higher supervisory	0.0794	0.270	0	1
Intermediate academic/independent	0.261	0.439	0	1
Intermediate supervisory/commercial	0.134	0.341	0	1
Other mental work	0.251	0.434	0	1
Skilled & supervisory manual	0.0719	0.258	0	1
Semi-skilled manual	0.0707	0.256	0	1
Unskilled manual	0.0434	0.204	0	1
Agrarian	0.00939	0.0964	0	1
Industry				
Agriculture	0.0137	0.116	0	1
Mining	0.000354	0.0188	0	1
Industrial	0.108	0.310	0	1
Utilities	0.0112	0.105	0	1
Construction	0.0435	0.204	0	1
Retail	0.0705	0.256	0	1
Catering	0.0222	0.147	0	1
Transport	0.0477	0.213	0	1
Financial	0.0510	0.220	0	1
Business services	0.0627	0.242	0	1
Government	0.107	0.309	0	1
Education	0.0963	0.295	0	1
Healthcare	0.207	0.405	0	1
Environmental/culture	0.0202	0.141	0	1
Other	0.139	0.346	0	1
N	22,577			

3.4 Results

Reproduction of earlier estimates - Table 2 shows the results of regressions of commuting time on teleworking and control variables. The dependent variable is the natural logarithm of commuting time, so coefficients can be interpreted as ($e^{\beta} - 1 * 100\%$) percentage change in the dependent variable due to a unit change in the independent variable. This also means that the analysis is only based on observations with a commuting time greater than zero. Column (1) shows the results of an OLS regression with year dummies. The coefficient of the telework dummy is 0.218, which means that those workers that start teleworking increase their commute by 24 percent. In column (2) we include individual fixed effects that should control for all time-invariant characteristics of people, including to a large extent (commuting) preferences (De Vos et al., 2018). The coefficient in this model is smaller than in the previous model, and the point estimate of 0.117 suggests a 12 percent increase in commuting times due to teleworking.

In column (3) we include a more flexible measure of teleworking, that distinguishes between less than one, one, and more than one day a week of teleworking. Consistent with expectations, the point estimate increases with more teleworking. Those that start to telework less than a day a week, shorten the commute by about 10 percent, compared to non-teleworkers. Those that telework one day a week see an increase of 16 percent, and those that start teleworking more increase their commute by about 17 percent.

In columns (4) and (5) we replicate two sensitivity checks from De Vos et al. (2019). Column (4) shows the results of a lagged dependent variable model, that does not include individual fixed-effects but instead includes the lag of commuting time (log). Consistent with De Vos et al. (2018) we use the 2 year lag. This model controls for selection bias due to commuting history. The results from this model show a smaller effect of teleworking on commuting. Teleworkers have 7 percent longer commutes. In column (5) we show the results of a long differences model that only includes the first and last year of the dataset, together with individual fixed effects. This model shows a greater effect of teleworking on commuting. Those workers that started to telework between 2008 and 2018 had on average 22 percent longer commutes in 2018.

The patterns we find in the data are overall consistent with De Vos et al. (2018). For instance, we find that sorting based on preferences biases OLS results upwards, as the FE estimate is smaller than the OLS estimate. There are important differences however. First, we tend to find greater effects in both the OLS and FE models (24 and 12 percent, vs. 12 and 5 percent respectively). Second, in contrast to De Vos et al. (2018) we do find that more teleworking leads to even longer commutes. Third, the spread between the FE and LDV model is greater in this study, although the effects are significant, and greater than earlier estimates. And finally, the long differences model in this study suggests a smaller effect of teleworking (22 vs. 32 percent).

Table 2: Regression results

	(1)	(2)	(3)	(4)	(5)
Telework dummy	0.218***	0.117***		0.0682***	0.199***
	(0.0143)	(0.0155)		(0.0110)	(0.0658)
Telework < 1 d/w			0.0948***		
			(0.0177)		
Telework 1 d/w			0.149***		
			(0.0229)		
Telework > 1 d/w			0.160***		
			(0.0476)		
Commuting time (log) (t-1)				0.834***	
				(0.00749)	
Age	-0.00948**	0.0124	0.0125	-0.00813**	0.0561**
	(0.00399)	(0.0119)	(0.0119)	(0.00373)	(0.0223)
Age^2	0.000118**	5.69e-05	5.74e-05	8.36e-05**	0.000122
	(4.59e-05)	(6.34e-05)	(6.34e-05)	(3.97e-05)	(0.000180)
Female	-0.0368***	-	-	-0.0191**	-
	(0.0122)			(0.00930)	
Children	-0.0374***	0.00483	0.00503	0.00272	-0.118**
	(0.0119)	(0.0115)	(0.0114)	(0.00845)	(0.0570)
Partner	0.0280**	-0.00244	-0.00193	-0.00442	0.121
	(0.0120)	(0.0200)	(0.0199)	(0.00850)	(0.0990)
Weekly hours worked	0.0124***	0.00604***	0.00598***	0.00174***	0.00443
	(0.000713)	(0.00112)	(0.00112)	(0.000604)	(0.00435)
Basic education	-0.170***	0.0294	0.0293	0.00914	-0.0808
	(0.0208)	(0.0290)	(0.0291)	(0.0162)	(0.177)
Lower education	-0.186***	0.0421	0.0425	-0.00600	-0.00962
	(0.0181)	(0.0262)	(0.0263)	(0.0146)	(0.139)
Medium education	-0.0515***	0.0163	0.0174	0.00820	-0.0499
	(0.0169)	(0.0204)	(0.0205)	(0.0132)	(0.0990)
Higher education	-	-	-	-	-
Constant	2.923***	2.275***	2.269***	0.680***	-0.164
	(0.106)	(0.455)	(0.456)	(0.0998)	(0.827)
Model	OLS	FE	FE	LDV	LD
Year dummies	Yes	Yes	Yes	Yes	Yes
Individual FE	No	Yes	Yes	Yes	Yes
Control dummies	Yes	Yes	Yes	Yes	Yes
Observations	21,859	20,519	20,519	11,796	1,228
R-squared	0.119	0.859	0.859	0.749	0.823
Within r-squared		0.0282	0.0288		0.1128

Dependent variable: Commuting time (log). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The FE models include individual and year fixed effects. The LDV model includes a lagged dependent variable. The LD model is a long-differences model based on 2008 and 2018 data alone. Control dummies include occupation and industry dummies.

Heterogeneity analysis - In this subsection we allow the effect of teleworking to differ across time, space, and occupations. The previous section has shown that our estimates with recent data, from 2008-2018, are higher than estimates from De Vos et al. (2018) who used data from the years 2002-2014. In Table 3 column (1) we interact the telework dummy with a period indicator that distinguishes between 2008-2013 and 2014-2018. The effects are remarkably similar: between 2008 and 2013 the effect is 12.2 percent, and between 2014 and 2018 it is 12.6 percent.

In column (2) we estimate the effect of teleworking separately for 5 levels of urbanization, as measured by address density of municipalities. The estimates show that the effects of teleworking on commuting are most pronounced (22 percent) in Moderately urban areas, with an address density between 1000 and 1500 per square kilometre. The effects are smallest (6 percent) in slightly urban areas, with an address density of 500-1000 per square kilometre. Across the other categories the effects are similar (between 10 and 13 percent).

Figure 3 has shown that there have been differential growth patterns of teleworking across occupations, and that there was a clear difference between occupations where teleworking remained stable or diminished between 2008 and 2018, and occupations where teleworking has grown. The latter occupations are 'Higher academic', 'Higher supervisory', 'Intermediate academic', 'Intermediate supervisory/commercial', and 'Other mental work'. In columns (3) we interact the telework dummy with and indicator of these two groups. For 'growth occupations', the effect we find is close to the estimate in Table 1 column (2) (13 vs. 12 percent). For 'other occupations' there is no significant effect. It should be noted that a large part of the (representative) sample is employed in growth occupations. Considering that the FE coefficients are estimated based on changes in teleworking, it is not surprising that the effects we find are driven by growth occupations where a lot of people have started teleworking.

In column (5) we focus on growth occupations, and see if there are any differences between these occupations. Among Higher academic, Intermediate supervisory/commercial, and Other mental work, the effects are high (between 19 and 23 percent) and do not significantly differ from each other. Higher supervisory workers show a commuting time increase of 6 percent due to teleworking that is only significant at the 10 percent confidence level, and intermediate academic occupations show no significant relationship.

Table 3: Heterogeneity regressions

	(1)	(2)	(3)	(4)
Telework * Period				
*2008-2013	0.115***			
	(0.0177)			
*2014-2018	0.119***			
	(0.0184)			
Telework * Urbanization				
*Very strongly urban		0.0923**		
		(0.0449)		
*Strongly urban		0.122***		
		(0.0324)		
*Moderately urban		0.198***		
		(0.0381)		
*Slightly urban		0.0575*		
		(0.0318)		
*Not urban		0.108***		
		(0.0373)		
Telework * Occupation				
*Growth occupations			0.120***	
			(0.0158)	
*Other occupations			0.0494	
			(0.0855)	
*Higher academic				0.201***
				(0.0462)
*Higher supervisory				0.0608*
				(0.0347)
*Intermediate academic				0.0303
				(0.0240)
*Intermediate supervisory/commercial				0.176***
				(0.0353)
*Other mental work				0.209***
				(0.0335)
Constant	2.276***	2.302***	2.285***	2.525***
	(0.455)	(0.475)	(0.454)	(0.469)
Control variables	Yes	Yes	Yes	Yes
Sample	Full sample	Full sample	Full sample	Growth occ.
Model	FE	FE	FE	FE
Observations	20,519	15,952	20,519	16,523
R-squared	0.859	0.847	0.859	0.861
Within r-squared	0.0282	0.0313	0.0283	0.0298

Dependent variable: Commuting time (log). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The 5 Growth occupations are included in the interactions. Other occupations are skilled and supervisory manual, semi-skilled manual, unskilled manual, and agrarian work.

The results from this subsection show that although our estimates are higher than estimates with older data, between 2008 and 2018 the effect of teleworking on commuting was stable. There is however considerable spatial heterogeneity, and what stands out is that the effect of teleworking on commuting is most pronounced in moderately urban areas. Figure 5 in Appendix A shows that these Moderately urban municipalities include municipalities within the core area, between and around cities in the *Randstad*¸ and between cities along another urban corridor known as the *Brabantse stedenrij*. The analysis further shows that the effects of teleworking on commuting are concentrated in occupations where teleworking grew between 2008 and 2018. Even among these occupations, there is considerable heterogeneity in the relationship between teleworking and commuting. Specifically, 'Higher supervisory' and 'Intermediate academic' occupations show a strong significant relationship, as opposed to 'Higher academic', 'Intermediate supervisory/commercial', and 'Other mental work'.

3.5 Conclusion and policy implications

This paper gives updated estimates of the effect of teleworking on commuting time, and it is the first to examine heterogeneity in this effect – across time, space, and occupations – with the use of panel data. We use a novel longitudinal data source from the Netherlands, for the years 2008-2018, and find that workers who start teleworking increase their commute by 12 percent on average. This estimate is higher compared to studies that use older data, although we find that during our study period the effect has remained stable. A potential explanation for this stability is that between 2008 and 2018 teleworking technologies in the Netherlands have not improved a great deal, in the sense that teleworking has become a better substitute for working at the job location. This is supported by data from the OECD (2019) that shows that by 2009 already 90 percent of the Dutch population had access to internet, and growth rates of internet access were higher before 2009 than after. Given this stability in a situation with matured technology, we expect that this 12 percent will remain a good estimate of the relation between working from home and commuting.

We find considerable spatial variation in the effect of teleworking on commuting, and the effect is largest in Moderately urban municipalities, largely situated between cities belonging to the urban core. Given that teleworking increases the acceptance of a longer commuting distance, a tentative explanation for this spatial pattern is that these municipalities profit from a combination of proximity to larger cities, and the presence of good quality infrastructure between large cities. For Moderately urban municipalities situated between larger cities, an increase in spatial reach may disproportionately affect job accessibility as a consequence.

We finally show that the effect of teleworking on commuting is driven by occupations in which teleworking has grown. These are occupations that are characterized by intellectual rather than manual work, and even within these occupations there are differences. The notion that the ability to telework, and the relation between teleworking and commuting time differs between occupations is not new, but this is the first paper to give clear-cut estimates of the effects by occupation, based on time-variation in teleworking.

There are at least three important limitations to the research in this paper. First, our estimates are based on commuting time and not commuting distance. Ideally the robustness of our results should be checked with data on actual commuting distance. Second, the fixed effects model controls for time-invariant unobserved preferences, but changes in preferences over time are not fully captured. Finally, due to lacking data on residential locations and job locations we did not address whether changes in commuting time are due to changing jobs or changing residences. For the purpose of solely estimating the effect of telework in commute length this simultaneity is not relevant per se, but from a policy perspective it is crucial to know which mechanism drives the relationship.

We distil three policy implications from this research. First, policymakers should not underestimate the effects of teleworking on commuting distance, especially if teleworking is promoted as a means to reduce harmful emissions. Our results suggest that 1 day of teleworking per week increases commutes by 16 percent on average (Table 2 column (3)). For people who work 4 days a week (about the average 31 weekly hours), introducing a working at home day leads to a reduction in travelling for work, but only by 13 percent instead of the naively expected 25 percent. For those that work 5 days per week, there is a 7.2 percent reduction instead of 20%. It should be noted that on average, teleworking still reduces work travel, and we have not taken into account the effects of teleworking on mode choice.

Second, teleworking may improve the fortunes of remote areas, as it increases job-accessibility. Even in the least urban areas, we find that people who start teleworking increase their commutes by 11 percent on average. Teleworking makes it easier to live in remote places, and have a job at a longer distance. However, the effects are most pronounced for places situated in urban regions, between larger cities. This suggests that access to high-speed internet is not a complete substitute, but also a complement of proximity to thick labour markets. More generally, telecommuting increases the size of labour markets. Our average of 12% longer commuting times, increases the area of the labour market with 25%. Assuming that jobs, infrastructure and transport are equally spread and that commuting times translate neatly into commuting distances, this means a substantially larger labour market, which is even larger for certain professions and some areas.

Finally, in projecting the consequences of teleworking policies on commuting distance, policymakers should be aware that the effects differ between intellectual and *manual* occupations. This result may not only be driven by the nature of occupations, but also by the geography of job-types. Our data shows that the share of intellectual occupations increases with the level of urbanization, while the share of manual occupations increases. It should also be noted that there are stark differences even within intellectual occupations, and the effects are especially small for higher management (8 percent of workforce) and intermediate academic occupations (26 percent of workforce). This has a social impact, as the ability to work from home implies that some groups in society obtain more freedom in where they locate, whereas others do not, and are more bound to places and smaller labour markets.

References

Andreev, P., Salomon, I. and Pliskin, N. (2010) Review: State of teleactivities, Transportation Research Part C: Emerging Technologies, 18(1), 3–20. https://doi.org/10.1016/j.trc.2009.04.017.

Angrist, J. D. and Pischke, J. S. (2008) Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press.

De Graaff, T. (2004) On the Substitution and Complementarity between Telework and Travel: A Review and Application, Research Memorandum. Vrije Universiteit Amsterdam. Available at: http://degree.ubvu.vu.nl/repec/vua/wpaper/pdf/20040016.pdf.

Gubins, S., van Ommeren, J. and de Graaff, T. (2019) Does new information technology change commuting behavior?, Annals of Regional Science, 62(1), 187–210. https://doi.org/10.1007/s00168-018-0893-2.

Hamer, R., Kroes, E. and Van Ooststroom, H. (1991) Teleworking in the Netherlands: an evaluation of changes in travel behaviour, Transportation, 18(4), 365–382. https://doi.org/10.1007/BF00186565.

Henderson, D. K. and Mokhtarian, P. L. (1996) Impacts of center-based telecommuting on travel and emissions: analysis of the puget sound demonstration project, Transportation Research Part D: Transport and Environment, 1(1), 29–45. https://doi.org/10.1016/S1361-9209(96)00009-0.

Jiang, Y. (2008) The Impact of Telecommuting on the Journey to Work : A Two-Sample Instrumental Variables Approach. University of Maryland. Available at: http://drum.lib.umd.edu/bitstream/handle/1903/8506/umi-umd-5557.pdf.

Kim, S.-N., Mokhtarian, P. L. and Ahn, K.-H. (2012) The Seoul of alonso: New perspectives on telecommuting and residential location from South Korea, Urban Geography, 33(8), 1163–1191. https://doi.org/10.2747/0272-3638.33.8.1163.

Kim, S. N., Choo, S. and Mokhtarian, P. L. (2015) Home-based telecommuting and intra-household interactions in work and non-work travel: A seemingly unrelated censored regression approach, Transportation Research Part A: Policy and Practice. Elsevier Ltd, 80, 197–214. https://doi.org/10.1016/j.tra.2015.07.018.

Koenig, B. E., Henderson, D. K. and Mokhtarian, P. L. (1996) The travel and emissions impacts of telecommuting for the State of California Telecommuting Pilot Project, Transportation Research Part C: Emerging Technologies, 4(1), 13–32. https://doi.org/10.1016/0968-090X(95)00020-J.

Melo, P. C. and de Abreu e Silva, J. (2017) Home telework and household commuting patterns in Great Britain, Transportation Research Part A: Policy and Practice, 103, 1–24. https://doi.org/10.1016/j.tra.2017.05.011.

Mokhtarian, P. L., Salomon, I. and Choo, S. (2005) Measuring the measurable: Why can't we agree on the number of telecommuters in the U.S.?, Quality and Quantity, 39(4), 423–452. https://doi.org/10.1007/s11135-004-6790-z.

Muhammad, S., de Jong, T. and Ottens, H. F. L. (2008) Job accessibility under the influence of information and communication technologies, in the Netherlands, Journal of Transport Geography, 16(3), 203–216. https://doi.org/10.1016/j.jtrangeo.2007.05.005.

Nilles, J. M. (1991) Telecommuting and urban sprawl: mitigator or inciter?, Transportation, 18(4), 411–432. https://doi.org/10.1007/BF00186567.

OECD (2019) Internet access (indicator). https://doi.org/10.1787/69c2b997-en.

Van Ommeren, J. and Fosgerau, M. (2009) Workers' marginal costs of commuting, Journal of Urban Economics, 65(1), 38–47. https://doi.org/10.1016/j.jue.2008.08.001.

Pajević, F. and Shearmur, R. G. (2017) Catch Me if You Can: Workplace Mobility and Big Data, Journal of Urban Technology, 24(3), 99–115. https://doi.org/10.1080/10630732.2017.1334855.

Rhee, H. J. (2008) Home-based telecommuting and commuting behavior, Journal of Urban Economics, 63(1), 198–216. https://doi.org/10.1016/j.jue.2007.01.007.

Scherpenzeel, A. C. and Das, M. (2010) "True" longitudinal and probability-based internet panels: Evidence from the Netherlands. in Das, M., Ester, P., and Kaczmirek, L. (eds) Social and behavioral research and the internet: Advances in applied methods and research strategies. New York: Routledge, 77–104. https://doi.org/10.4324/9780203844922-4.

De Vos, D., Meijers, E. and van Ham, M. (2018) Working from home and the willingness to accept a longer commute, Annals of Regional Science, 61(2), 375–398. https://doi.org/10.1007/s00168-018-0873-6.

Zhu, P. (2012) Are telecommuting and personal travel complements or substitutes?, Annals of Regional Science, 48(2), 619–639. https://doi.org/10.1007/s00168-011-0460-6.

Zhu, P. (2013) Telecommuting, Household Commute and Location Choice, Urban Studies, 50, 2441–2459. https://doi.org/10.1177/0042098012474520.

Appendix

Table 4: Correspondence between LISS Professions and ISCO-08

LISS measure of profession	ISCO-08 codes
1 higher academic or independent profession (e.g. architect, physician, scholar, academic instructor, engineer)	2
2 higher supervisory profession (e.g. manager, director, owner of large company, supervisory civil servant)	1
3 intermediate academic or independent profession (e.g. teacher, artist, nurse, social worker, policy assistant)	3
4 intermediate supervisory or commercial profession (e.g. head representative, department manager, shopkeeper)	1
5 other mental work (e.g. administrative assistant, accountant, sales assistant, family carer)	4, 5
6 skilled and supervisory manual work (e.g. car mechanic, foreman, electrician)	7
7 semi-skilled manual work (e.g. driver, factory worker)	8
8 unskilled and trained manual work (e.g. cleaner, packer)	9
9 agrarian profession (e.g. farm worker, independent agriculturalist)	6

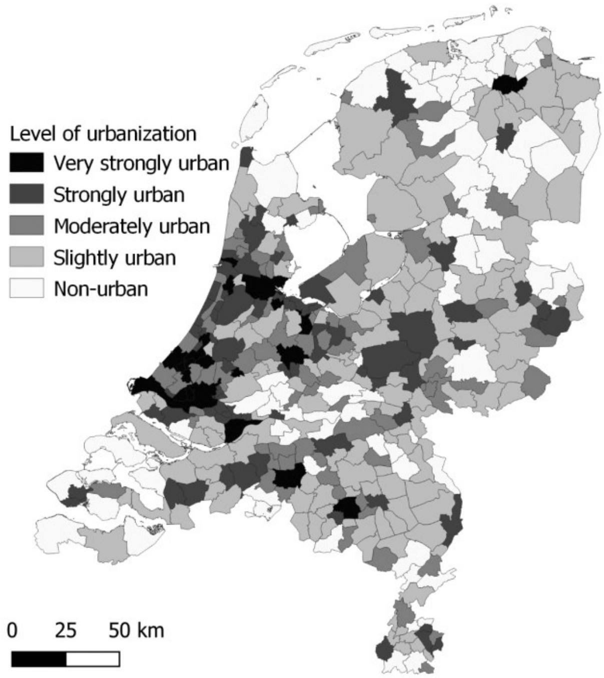


Figure 5: Map of levels of urbanization

Table 5: Summary statistics by telework intensity

VARIABLES	*No telework* mean	sd	*Telework <1 d/w* mean	sd	*Telework 1 d/w* mean	sd	*Telework > 1 d/w* mean	sd
Commuting time (min)	25.92	19.87	31.74	22.67	43.52	28.95	36.78	33.27
Age	44.24	11.23	45.50	10.74	45.38	10.43	47.05	12.21
Female	0.370	0.483	0.320	0.467	0.300	0.458	0.334	0.472
Children	0.457	0.498	0.512	0.500	0.517	0.500	0.400	0.490
Partner	0.755	0.430	0.760	0.427	0.798	0.401	0.763	0.425
Weekly hours worked	31.45	9.506	33.41	8.155	35.41	6.579	30.77	11.79
Basic education	0.257	0.437	0.0949	0.293	0.125	0.331	0.272	0.445
Lower education	0.395	0.489	0.245	0.430	0.197	0.398	0.325	0.468
Medium education	0.237	0.425	0.411	0.492	0.341	0.474	0.236	0.425
Higher education	0.112	0.315	0.249	0.433	0.337	0.473	0.167	0.373
Higher academic/indep.	0.0612	0.240	0.143	0.351	0.202	0.401	0.115	0.320
Higher supervisory	0.0595	0.237	0.190	0.393	0.178	0.382	0.103	0.305
Interm. academic/indep.	0.257	0.437	0.330	0.470	0.244	0.430	0.234	0.424
Interm. supervisory/commercial	0.127	0.333	0.167	0.373	0.164	0.371	0.153	0.360
Other mental work	0.268	0.443	0.149	0.356	0.173	0.378	0.235	0.424
Skilled & supervisory manual	0.0857	0.280	0.00722	0.0847	0.00478	0.0690	0.0314	0.174
Semi-skilled manual	0.0828	0.276	0.00516	0.0716	0.0143	0.119	0.0489	0.216
Unskilled manual	0.0485	0.215	0.00258	0.0507	0.0104	0.101	0.0692	0.254
Agrarian	0.00973	0.0981	0.00567	0.0751	0.0104	0.101	0.00923	0.0956
Agriculture	0.0139	0.117	0.0144	0.119	0.0135	0.116	0.0101	0.100
Mining	0.000328	0.0181	0	0	0.000797	0.0282	0.000923	0.0304
Industrial	0.113	0.317	0.0707	0.256	0.0916	0.289	0.106	0.308
Utilities	0.0113	0.106	0.00825	0.0905	0.0135	0.116	0.0120	0.109
Construction	0.0482	0.214	0.0227	0.149	0.00956	0.0974	0.0406	0.197
Retail	0.0716	0.258	0.0593	0.236	0.0661	0.249	0.0756	0.265
Catering	0.0232	0.151	0.00774	0.0876	0.0159	0.125	0.0378	0.191
Transport	0.0539	0.226	0.0150	0.121	0.0255	0.158	0.0268	0.161
Financial	0.0442	0.205	0.0712	0.257	0.102	0.303	0.0720	0.259
Business services	0.0524	0.223	0.0872	0.282	0.138	0.345	0.107	0.309
Government	0.0945	0.293	0.178	0.383	0.170	0.376	0.113	0.317
Education	0.0882	0.284	0.147	0.355	0.151	0.359	0.0784	0.269
Healthcare	0.224	0.417	0.188	0.391	0.0701	0.255	0.110	0.313
Environmental/culture	0.0211	0.144	0.0144	0.119	0.0183	0.134	0.0175	0.131
Other	0.140	0.347	0.116	0.320	0.114	0.318	0.192	0.394
	18,299		1,939		1,255		1,084	

4. Does Broadband Internet Allow Cities to 'Borrow Size'? Evidence from the Swedish Labour Market

Published in Regional Studies (2020)
Written by Duco de Vos, Urban Lindgren, Maarten van Ham, and Evert Meijers.

Abstract
Borrowed size refers to the idea that small cities nearby larger metropolitan centres are able to reap the advantages of large agglomerations, without the agglomeration costs. This study explores whether broadband internet helps such smaller cities to enjoy the labour market benefits of a larger city. Using Swedish micro-data from 2007-2015, together with unique data on broadband, our results suggest that broadband allows smaller cities to reap such benefits indeed. We find that borrowed size is primarily driven by the overall penetration of broadband in the place of residence, rather than by broadband availability at the residence.

4.1 Introduction

In 1973 Alonso coined the concept of 'borrowed size', referring to the situation "whereby a small city or metropolitan area exhibits some of the characteristics of a larger one if it is near other population concentrations" (Alonso 1973, p. 200). He argued that to reap the benefits of agglomeration in large cities, it is often not necessary to locate in these cities themselves, but it suffices to be within reasonable proximity. Recently, this concept has been used to explain why medium sized cities in Europe do not perform worse than the largest cities, contrary to the theory of agglomeration economies (Dijkstra et al 2013, Meijers and Burger 2017). Even while large cities are increasingly perceived as engines of economic growth, it is argued that Europe's medium-sized cities can continue to flourish, provided they are within reasonable proximity of larger urban areas (Dijkstra et al 2013).

In the original conceptualisation of borrowed size (Alonso 1973), not much attention is paid to the mechanisms that allow borrowed size to occur – just proximity is stressed. Yet, there is evidence for integration and interaction between cities more generally playing a key role (e.g. Phelps et al, 2001; Meijers et al, 2018) and particular emphasis has also been put on data infrastructure allowing the transfer of information. For instance, Phelps (1992, p. 44) notes that "information costs […] have played an increasing role in constraining the geographical availability of external economies over time"; Dijkstra et al (2013, p. 334) argue that in Europe "improvements in the access to services, including broadband, outside large cities may have facilitated the higher growth rates of smaller centres and rural regions"; and Hesse (2016, p. 617) even contends that "borrowing size or significance no longer relies on physical proximity between the cities, but on embeddedness in overarching networks between and within polycentric city-regions, via corporate relations, market pervasion and, last but not least, information and communication networks." Thus the idea that broadband internet and borrowed size are related is pervasive, but to date there is no empirical work that investigates this relationship. The main aim of our paper is to fill this gap.

We expect the effects of the local availability of broadband internet on borrowed size to be highly pronounced when it comes to labour market relations. High speed internet allows workers to live further from spatial concentrations of employment opportunities. This may arise because broadband internet enables people to work from home one or more days per week, which in turn makes them accept longer commutes on the other days (De Vos et al 2018). In addition, high speed internet allows for a more efficient (online) job search process (Autor 2001), for instance because 'word of mouth' on labour market opportunities is moving online, and because a specific search for jobs in a particular place is now often replaced with search for jobs within a certain radius. Likewise, this gives entrepreneurs more freedom in settling in smaller cities. Phelps et al (2001) show that firms can locate in small towns while still accessing specialised labour and informational external economies of larger nearby places. This all increases labour market accessibility overall and may lead to higher employment rates, especially in smaller cities at some distance from larger metropolitan areas.

In this paper we analyse the relation between broadband internet and labour market outcomes across space more in detail. We use a unique dataset that traces the spatial diffusion of broadband internet technologies in Sweden between 2007 and 2015, and we combine this with register data on all Swedish working age individuals. We first assess whether broadband internet improved labour accessibility outside of Sweden's large cities. Second, we assess whether the local introduction of broadband leads to higher employment rates outside of the largest cities.

In our empirical approach we distinguish between the effect of broadband availability at the residence location and at the place level. The former measure should pick up broadband effects on individual labour market decisions, whereas the latter should pick up effects on the local economic environment. We estimate the effects of *changes* in broadband on *changes* in labour market outcomes – commuting and employment – using fixed effects models. Because it takes time for people to adapt their labour market status to a new situation, we focus on the effect of broadband two years after its introduction. To uncover potential 'borrowed size' interactions, we allow the effects of broadband to differ between urban cores, urban regions and rural regions, and across distance from the nearest urban core.

The rest of this paper is structured as follows. In Section 2 we discuss the literatures that deal with interactions between IT, borrowed size, commuting, and local economic performance. Section 3 introduces the data. In Section 4 we discuss our empirical approach more in detail. Section 5 presents the results and in Section 6 we conclude.

4.2 Literature review

Borrowed size in labour markets – Alonso (1973) understood borrowed size as a process in which small places have some characteristics of larger nearby cities. Meijers and Burger (2017) distil five features of borrowed size from Alonso's original description. Borrowed size is (1) an analytical concept, (2) concerning small cities, (3) in geographical proximity of larger metropolitan areas, (4) that retain the advantages of smaller cities but enjoy the benefits of agglomeration nearby (5) through accessibility and network connectivity. Broadband connections are a clear example of infrastructure that allows for both accessibility and network connectivity. It appears relevant to distinguish an accessibility effect for individuals, which can be captured with broadband availability at the residential location, and a network connectivity effect that affects the local economy, which can be captured by the share of broadband connections in the place of residence.

As mentioned in the introduction, several authors suggest that borrowed size can be derived from fast internet (Phelps 1992, Dijkstra 2013, Hesse 2016). We expect the effect of broadband internet on borrowed size to be most pronounced when it comes to labour market outcomes. As Alonso (1973) put it, labour markets of small cities in large metropolitan regions "enjoy a wider and more flexible range of demand and supply" (p.200) than isolated cities, but in the literature on borrowed size labour market outcomes have received little attention, compared to outcomes for business services (Phelps 1998; 2001) and access to urban functions such as shopping and leisure amenities (Meijers and Burger 2017; Meijers et al 2015). Therefore in this paper we measure whether broadband enables borrowed size by looking at the differential effects of broadband on employment rates, across urban and non-urban areas.

With borrowed size likely being a result of interaction between cities, it makes sense to consider commuting as an enabler of borrowed size on the labour market, with different effects on the level of individuals and of places. On the individual level, we expect longer commutes as high quality internet is a precondition for working from home. Interestingly, recent evidence confirms the long standing hypothesis that home-based teleworking using the internet leads to longer commutes, and an increased geographical scale of labour markets for a number of different jobs. Using data from the Netherlands, De Vos et al (2018) find that working from home at least two days per month allows people to accept a 5 percent longer commute, and over time this can increase to 32 percent. This means that for those able to work from home, labour market accessibility increases significantly with the advent of broadband internet. On the place level, it could be that commuting distances actually decrease. On the one hand broadband internet allows the benefits of agglomeration to spread over a larger territory, enabling firms to locate in smaller places at some distance from larger cities, and perhaps closer to workers, while still being able to tap into the agglomeration benefits of the larger city. On the other hand, broadband internet may lead to a more efficient job search process that eliminates the extent of excess commuting. This may especially be relevant in large cities with a lot of employment opportunities, and possibilities to shorten the commute. By assessing the direct, reduced form relation between internet connections and commuting distance, the current paper also adds to the literature on teleworking and commuting.

Local performance and IT – While there are no studies on the relationship between internet and borrowed size, several studies investigated to what extent advances in digital technology complement or substitute agglomeration economies (Gaspar and Glaeser 1998, Sinai and Waldfogel 2004). These studies focused on the effects on product variety, and the importance of face-to-face contact, but the effects of information technology on labour market outcomes have been ignored. However, there are studies that go into the relationship between information technology and more efficient matching on the labour market, without reference to spatial issues. Autor (2001) discusses the potential consequences of improved worker-firm communications through the internet, stressing the effect on matching between employers and employees. He provides several arguments of why search costs in the labour market are lower when using the internet. One of the main findings of basic labour market search theory is that lowering the cost of finding a job increases productivity through higher match quality.

While there are several empirical studies on the link between the internet and (labour) productivity, there is only a limited number of studies that have taken the proposed relationship between the internet and better matching to the data, with inconclusive results. Kuhn and Skuterud (2004) use data from the 1998 and 2000 US Current Population Survey and find that among online job searchers, unemployment spells are not shorter, and possibly even longer than their offline counterparts. In contrast, a follow-up on this study that uses 2005-2008 data finds that online searchers find jobs about 25% more quickly (Kuhn and Mansour 2014). Kroft and Pope (2014) use the geographic penetration of the local advertising website Craigslist. They find evidence that Craigslist crowded out traditional job advertising outlets, but they find no effect on the unemployment rate. Czernich (2014) uses an instrumental variables strategy based on technical constraints of telephone connections to assess the effects of broadband penetration on unemployment in Germany, and finds no significant effect.

Research on the link between broadband provision and firm location behaviour does suggest that broadband is associated with new firm establishment. The strength of the relationship appears to vary over space (Tranos and Mack 2016) and over different industries (Mack and Grubesic 2009). Importantly, early research pointed out that the urban bias of IT may have hampered firm location outside core regions (Sohn 2003, Mack and Grubesic 2009). More recent research highlights the manner in which broadband technologies substitute the agglomeration benefits of large cities in rural and remote areas. For instance Mack (2014) finds that broadband is correlated with the presence of agricultural and rural firms, in line with the hypothesis that IT substitutes face-to-face contact .

A study by Atasoy (2013) is perhaps most similar to the current study. This study estimates the effect of broadband expansion on labour market outcomes in the US between 1999 and 2007. In this study broadband penetration is measured as the share of the county population living in an area where broadband internet is available. The main finding is that broadband is associated with a 1.8 percentage point higher employment rate, with larger effects in more remote areas.

We add to the literature on the relation between internet and employment by paying attention to potential heterogeneous effects across space. Specifically, we allow the effect of internet connections on employment to differ between urban, suburban, and rural areas, and we assess the intermediating effect of distance to urban cores. We analyse the effect of broadband availability on employment at a finer geographical scale than Atasoy (2013) and importantly, we disentangle the effects of broadband penetration at the level of places and broadband availability at the residential location.

Hypotheses – Here we define our hypotheses for the effects of broadband on commuting and employment, and what we interpret as evidence for borrowed size. First, based on the working-from-home mechanism (De Vos et al. 2018), we expect residential broadband availability to increase commuting distances. We also expect a positive effect on employment rates, because distant employment opportunities become more accessible. Second we expect that as place-level broadband rates increase, places become more attractive to firms, consistent with Phelps (2001) and Mack and Grubesic (2009). This should lead to an increase in local employment (Atasoy 2013), and it may correlate negatively with commuting because more people can find a job nearby. Third, we consider it evidence of borrowed size if the effects of broadband are largest at some distance from larger cities rather than close-by, namely in places that were at the edge of labour market areas, but that have become more closely incorporated through broadband internet. Note that borrowed size can either arise through workers being increasingly able to find a job in the nearby city, or through the place becoming more attractive for firms to locate.

4.3 Data

Astrid data – We use longitudinal register data from the Swedish ASTRID database at Umeå University. Our data includes all Swedish individuals aged 18-64, between 2007 and 2015, with information on earnings, education, demographics, the residential location, and workplace data including firm- and establishment identification, location, industry, sector, and size.

The earnings variable denotes annual income from working. Education levels are given according to the *Svensk Utbildningsnomenklatur* (SUN). This classification includes 7 levels that range from primary school to PhD level education. Demographic variables include sex and age. The residential location is available with great geographic detail, with an accuracy of 100 metres. We recode the data to 250*250m grid cells so it can easily be matched to the grid level data of broadband availability. For every worker we have ID variables for both the establishment and the firm at which the worker works. A unique feature of the ASTRID data is that there is information on the location of the exact workplace, which is especially relevant for multi-plant firms. Industry information is available according to the *Standard för Svensk näringsgrensindelning 2007* (SNI 2007), which corresponds with international standards. We take into account 2-digits so we end up with 21 industries.

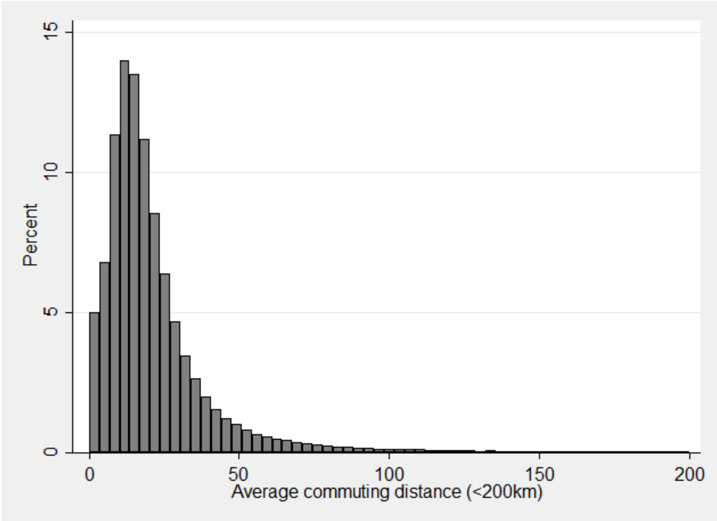
We use the register data to construct relevant control variables, and to measure commuting and employment. Having both information on residential- and work locations allows us to calculate the (Euclidean) commuting distance for every working individual. The Euclidean measure of distance may introduce a measurement error, but it is not clear in which direction, because short distances are generally associated with lower travel speeds (Duranton and Overman 2005). However, Combes and Lafourcade (2005) show that Euclidean distance correlates strongly (0.97) with generalized transportation costs, so we do expect this measurement error to be limited. For every grid cell we record the average commuting distance of all *working* individuals. The distribution of this variable is shown in Figure 1 panel (a).

We measure the employment rate at the level of grid cells, and following Andersson Joona et al (2013) we define every individual with an annual income higher than 160,000 SEK (about 15,000 EUR) as employed. Using this cut-off point we may measure the employment rate with an error, because for instance students or part-time workers may be considered unemployed. However, as the cut-off point is lower than the lowest collectively agreed income (Andersson Joona et al 2013), we do not expect the error to differ between years. Therefore we expect the measurement error related to changes in the employment rate to be limited. The distribution of this variable is shown in the panel (b) of Figure 1.

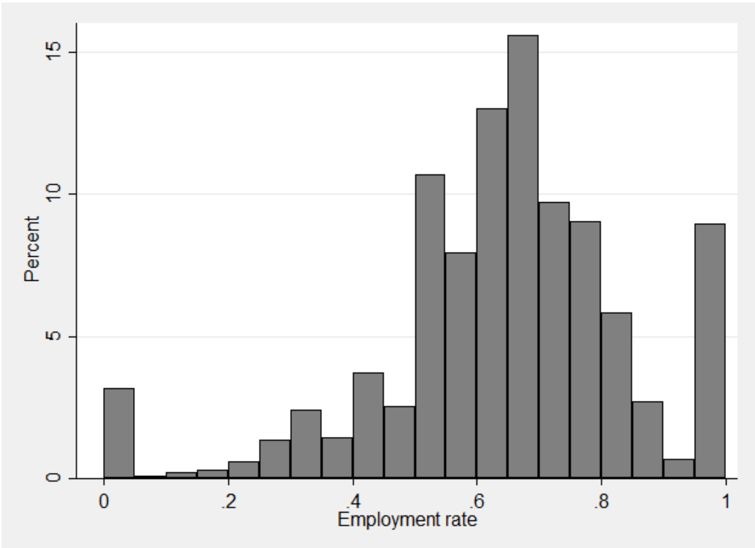
Using the ASTRID data we measure aggregate socio-economic indicators and industrial structure at the level of local labour market areas (*lokala arbetsmarknadsregioner* of FA-regions). This regional division was constructed by the Swedish Agency for Economic and Regional Growth (Tillväxtverket) and consists of 60 regions. Note that this regional division covers the whole of Sweden, and is different from the Functional Urban Area definition we use to delineate cities and their hinterlands. On average, these labour market regions have a working age population of 100,000.

Broadband data – Broadband penetration data come from annual surveys of internet suppliers, conducted by the Swedish Post and Telecom Authority (PTS) starting from 2007. The internet suppliers are requested to produce a list of all addresses that have a connection, by connection type. These data are used by PTS to construct a geographically referenced dataset that expresses the availability of broadband internet in 250*250 metre grid cells. Broadband internet is considered available if at least one address in a grid cell has a broadband connection. We use data on the number of addresses in each grid to construct a measure of broadband penetration for each grid. Figure 2 shows time variation in the penetration rate of different internet technologies across grid cells in Sweden. While DSL access (high-speed internet through telephone lines) is almost ubiquitous 3G mobile access has increased significantly between 2007 and 2009, and broadband access through Cable and Glass fibre has increased in the years after, with a significant bump after 2014.

Our main results are based on a dummy variable that indicates any *fixed* connection at the residential location (DSL, Cable, or Glass fibre). While according to the *Bredbandskartläggning* report of 2011 the average speed of DSL is slower than the speed of Cable and Fibre (14 vs. 100 mbit/s), in the sensitivity analysis we show that models based on the *fastest fixed connection* deliver similar results, and the effects of DSL and Cable/Fibre are comparable. In our models we also estimate the effect of the introduction of mobile internet. Because of the limited speed (6.5 mbit/s) we expect 3G internet to have a smaller effect on working from home and the attractiveness of places for business, and thus a smaller effect on commuting and employment.



(a)



(b)

Figure 1: Distributions of commuting distance (a) and employment rates (b) across grid cells

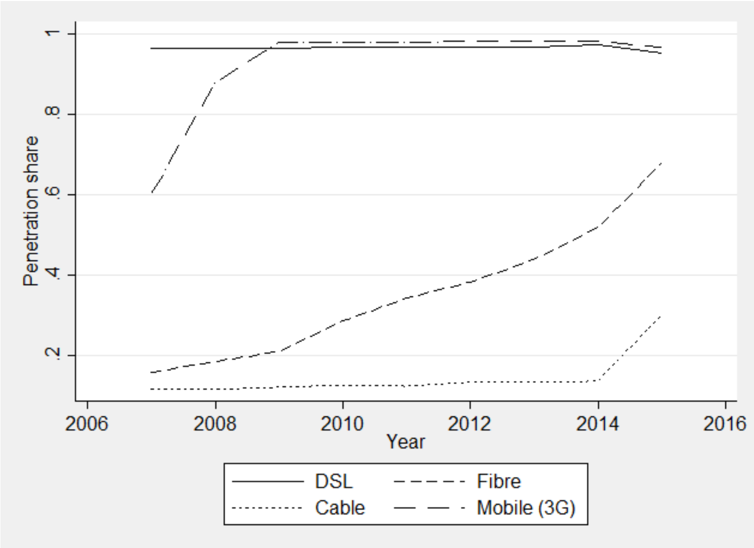


Figure 2: Penetration of different internet technologies

Summary statistics – Table 1 shows the summary statistics. The average employment rate in grid cells is about 64 percent. This is consistent with other sources, for instance the Swedish Labour Force Survey of 2017 reports an employment rate of 67.6 percent. The average commuting distance in grid cells is 23.6 kilometres, which is slightly higher than usual values reported in the literature because low-density grid cells (with longer commutes) have equal weights as high density grid cells. Broadband penetration levels are close to 100 percent on average, both in grid cells and in places, but especially at the grid cell level there is still a lot of variation. The table also reports a summary of control variables at the grid cell level and at the local labour market level. These control variables are similar to the ones used in Atasoy (2013). At the level of grid cells we include population, and in median income, to control for the effects of population dynamics that may coincide with, or even cause broadband adoption. At the level of local labour markets we include variables on demographic structure, education, and firm dynamics, to control for changes in demographic and economic structure that may coincide with broadband expansions. Since broadband internet has been more common in urban areas, we include the share of urban population as a control variable. Omitted from this table are 21 variables that indicate employment shares in different industries, that should control for further changes in economic structure.

Table 1: Summary statistics

VARIABLES	(1) mean	(2) sd	(3) min	(4) max
250m grid cell				
Employment rate	0.639	0.208	0	1
Avg. commuting distance*	23.60	29.06	0.0354	500.0
Broadband (res.)	0.975	0.158	0	1
Cable/Fibre (res.)	0.393	0.489	0	1
DSL (res.)	0.581	0.493	0	1
Mobile (3G)	0.926	0.262	0	1
Population	53.29	111.3	1	3,058
Median income	238,005	105,639	0	11,353,600
Place				
Broadband (place)	0.986	0.0826	0	1
DSL (place)	0.490	0.383	0	1
Cable/fibre (place)	0.496	0.383	0	1
Labour market area				
Share male	0.510	0.00681	0.501	0.548
Share low educated	0.0308	0.0101	0.00706	0.0591
Share medium educated	0.185	0.0257	0.115	0.282
Share high educated	0.520	0.0447	0.447	0.648
Share college educated	0.250	0.0526	0.121	0.344
Share young	0.297	0.0147	0.241	0.360
Share old	0.299	0.0295	0.257	0.421
Number of firms	39,647	49,448	190	147,185
Total payroll (in millions)	11,514	15,028	21.00	48,840
Employees/establishment (avg,)	8.555	1.040	4.190	11.14
Population	495,546	587,662	1,588	1,720,469
Urban population	242,294	327,414	0	955,805
Income	23,267	2,499	10,710	32,160
N	871,688			

* Due to grid cells without employment we have only 837.368 commuting distance observations.

4.4 Methodology

To investigate the effect of broadband availability on commuting and employment rates we aggregate the data to the level of 250m*250m grid cells, and we assess whether access to broadband is associated with the outcomes. To estimate the effect of *time-variation* in broadband availability on *time-variation* in outcomes, we use a fixed effects model. The models are specified as follows:

$$O_{gt} = \alpha B_{gt}^{res} + \beta B_{lt}^{loc} + \gamma X_{glrt} + \eta_g + \theta_t + \epsilon_{gt} \qquad (1)$$

where O_{gt} denotes the outcome (commuting distance or employment) in grid cell g at time t, B^r is a residential broadband dummy, B^p is the place-level broadband penetration rate, X_{glrt} are control variables at the grid cell, place, and regional level, η_g are grid cell fixed effects, θ_t are time fixed effects, and ϵ is the error term.

To assess whether the effects are larger in areas surrounding large cities, we estimate a model including interactions between broadband and an indicator for urban core, urban hinterland, and rural region. We define these areas using the OECD's Functional Urban Area (FUA) definition.[1] It is somewhat unfortunate that these FUAs are defined from a monocentric perspective, defining the commuting zone of a single city. These zones sometimes touch each other, while in reality they probably overlap to some extent (e.g. Linköping and Norrköping, see Appendix Figure A1). However, the OECD definition has the important advantage of being readily available and consistent, which allows reproduction of results. Finally, we interact the broadband variables with a more flexible measure of urbanization, based on distance to the nearest FUA.

There are several identification problems when analysing the relationship between broadband internet and commuting distance and employment rates. First, there may be omitted variables that correlate with both broadband adoption and the outcomes, without there being a direct causal relationship. For instance, the introduction of broadband in a place may be part of a larger development project, that may include new infrastructure or new housing. We tackle this issue to a large extent by controlling for a number of variables, including income and population at the level of grid cells, and a socio-economic indicators and industrial structure at the level of local labour market areas. Specifically, changes in broadband availability in grid cells may be correlated with broadband penetration in the place of residence, so the broadband indicator may pick up on effects of broadband penetration on the local economy. Therefore, we include broadband penetration at the level of the place[2] to which the

[1] http://www.oecd.org/cfe/regional-policy/Sweden.pdf

[2] We use Statistics Sweden's *tätorter* (localities) and *småorter* (smaller localities) to define places. Localities are places with at least 200 inhabitants, and less than 200 metres between

grid cell belongs, to isolate the effect on personal outcomes from the effects of broadband availability on the local structure of the economy. We report the effects of both variables, where the effects of residential broadband are more likely to be causal. The effects of the place level broadband rate should be interpreted with some caution, due to potential confounders such as new infrastructure or housing projects.

Second, commuting behaviour and employment levels may also drive broadband adoption, instead of the other way around. We tackle this issue to a large extent by taking the 2-year lag of broadband (penetration) as our independent variable of interest. It should be noted that we analyse the effects of *changes* in broadband penetration on *changes* in employment rates and commuting. Although current and future outcomes may be correlated, it is less straightforward that future changes in these outcomes drive current changes in broadband penetration.

Finally, the institutional setting in Sweden is such that broadband adoption is not dependent on market forces alone. Public funding supports broadband initiatives of rural municipalities that lack commercial incentives to invest.[3] Sweden's wish to provide broadband to all of its citizens is one of the main reasons for the existence of the broadband data that we use, and it is reflected in the recent broadband strategy titled "A Completely Connected Sweden by 2025" (Government Offices of Sweden 2017). While the effort of communities for broadband initiatives may not be random, the timing of these initiatives, and of the arrival of broadband in rural municipalities (in one year or the next) may have an element of randomness, which increases the likelihood that our estimates can be interpreted as causal effects. To the extent that local effort is endogenous, our results should be interpreted as upper bounds of the effect of broadband introduction.

4.5 Results

Broadband and labour market outcomes – Our first hypothesis wat that residential broadband availability increases commuting distances as it enables working from home, and increases employment rates because distant employment opportunities become more accessible. Table 2 shows the results of six regression models with commuting distance (columns (1-3)) and the employment rate (columns (4-6)) as dependent variables. In all models observations correspond to grid cells. As our control variables at the local labour market level have several multicollinearity problems that affect their coefficients, we only focus on the effects of our independent variables of interest, that have consistently low VIF values. Column (1) shows a regression of commuting distance in year t on the broadband indicator in the same

buildings. Smaller localities have between 50 and 200 inhabitants, and no more than 150 metres between buildings.

[3] https://ec.europa.eu/digital-single-market/en/country-information-sweden

year. According to this model, the introduction of broadband internet in a grid cell coincides with a 1.7 kilometre decrease in commuting distance on average. This model also shows that the introduction of 3G internet correlates positively with commuting distance. The high r-squared of 0.530 is mainly due to the fixed effects, as in Appendix Table A1 we show that a model without control variables has an r-squared of 0.525.

In column (2) we take the 2 year lag of the broadband indicator and the Mobile (3G) variable. As argued in the methodology section, changes in commuting distance and employment may cause broadband adoption rather than the other way around, but it is unlikely that current changes in these behaviours affect broadband adoption in the past. According to this model, the introduction of broadband in year t-2 coincides with a decrease in commuting distance of 1.15 kilometres on average in year t. The effect of 3G internet is not significant.

The findings of these models run counter to our first hypothesis. Most likely, broadband availability at the residence in grid cells is correlated with the overall broadband penetration in a place. Therefore, our broadband variable may pick up on the effect of broadband on local economic structure and business climate (Eriksson et al 2014). Especially outside of central cities, broadband availability may improve the business environment and attract more firms (which results in higher employment, our second hypothesis). This may explain why we find a negative effect of broadband on commuting, as a higher share of people may be able to find employment locally rather than further away. In column (3) therefore, we include the 2 year lag of place-level broadband penetration (share of connected individuals in places) as an additional independent variable. Consistent with our hypothesis, the negative effect of broadband internet on commuting is largely driven by place-level broadband penetration, rather than the residential availability of broadband for households. A 0.1 increase in the share of connected individuals at the place level (slightly more than a standard deviation) leads to a small decrease of commuting distance by 0.18 kilometres, or about 0.8% of average commuting distance.

In columns (4-6) we repeat the steps above with the employment rate as dependent variables. Column (4) shows that the introduction of broadband in grid cells coincides with a 0.8 percentage point employment increase on average. Taking the 2 year lag in column (5), this effect is only 0.3 percentage points. If we take into account the 2 year lag of the place-level broadband rate, in column (6) again we find that broadband penetration at the place level drives this effect. The effect is rather small, as a 0.1 increase in the share of connections corresponds with a 0.09 percentage point increase in employment. Moving from no connections to 100% connectivity increases employment by 0.9 percentage points, about 1.4% of the average employment rate. Across these three models, 3G introduction has a positive and significant but small effect on employment rates.

In summary we find that overall, changes in the availability of broadband internet affect commuting distance negatively and employment rates positively. These effects are primarily driven by the penetration of broadband in the place of residence, rather than by the availability of broadband at the residence. The effect on commuting distance is relatively small: every 10 percentage point increase in broadband connection is associated with an average decrease of 0.18 kilometres in commuting distance. Our estimate of the effect of residential broadband availability on employment is lower than the estimates by Atasoy (2013) (0.8 vs 1.8 percentage point). This may be because the estimates of Atasoy (2013) refer to the period 1999-2007, and regions that stand to gain most from broadband may be expected to connect earlier rather than later. When taking lags and controlling for the place-level broadband rate, Mobile (3G) internet does not affect commuting distance, but it has a small effect on the employment rate.

Table 2: Regression results

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Commuting distance			Employment rate		
Broadband (res.) (t)	-1.721***			0.00843***		
	(0.302)			(0.00141)		
Broadband (res.) (t-2)		-1.149***	-0.708		0.00344*	0.00117
		(0.377)	(0.431)		(0.00177)	(0.00201)
Broadband (place) (t-2)			-1.750**			0.00920**
			(0.826)			(0.00389)
Mobile (3G) (t)	0.480***			0.00122*		
	(0.140)			(0.000678)		
Mobile (3G) (t-2)		-0.0505	-0.0741		0.00206***	0.00219***
		(0.143)	(0.144)		(0.000685)	(0.000687)
Constant	-203.1	-263.0*	-260.0	-0.0541	1.702**	1.686**
	(137.7)	(158.6)	(158.6)	(0.669)	(0.763)	(0.763)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year and grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	825,353	634,033	634,033	844,449	646,345	646,345
R-squared	0.530	0.579	0.579	0.732	0.764	0.764

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Control variables include Population (log) and Income (log) at the grid cell level, and Population (log), Income (log), Share urban population, Share male, Share lower-, medium-, higher-, and college educated, Share young and old, Number of establishments, Number of employees per establishment, the aggregate income, and 21 industry share variables at the level of labour market areas.

Broadband and borrowed size – In the preceding section we have assumed that broadband affects commuting and employment similarly across the urban hierarchy and across space. However this is not straightforward, for instance if there are interaction effects between information technology, agglomeration, and borrowed size (Gaspar and Glaeser 1998, Sinai and Waldfogel, 2004, Hesse 2016). Our third hypothesis stated that the effects of broadband are largest at some distance from larger cities rather than close-by, namely in places that were at the edge of labour market areas, but that have become more closely incorporated through broadband internet. In Table 3 we estimate our models again, taking into account this spatial heterogeneity in two ways. First by allowing the effect of the availability of broadband to differ between urban cores, the remainder of their urban regions, and rural areas. Second, by allowing the effect of broadband to depend on distance to the nearest urban core.

Column (1) is similar to Table 2 column (3), except that the effects of both broadband measures are estimated separately for grid cells within FUA cores, outsides of cores within FUA's, and outside of FUAs. There seems to be a great deal of heterogeneity. First, we find no significant effects of broadband on commuting within urban cores. Second, in the urban regions surrounding urban cores, we find that residential broadband availability reduces commuting distances (by 2.2 kilometres), while it increases commuting distances (1.3 kilometres, only significant at the 10% confidence level) outside of urban regions, in rural areas. These effects may seem modest, but since the average commuting distance in our sample is 23.6 kilometres, a 2.2 (1.3) increase means that commutes are 9.3% (5.5%) longer. This is in line with estimates of the effect of working from home on commuting distance (De Vos et al, 2018). Third, the effects of place-level broadband are only present outside of urban regions. Here, moving from 0 to 100% connected individuals leads to a 4.3 kilometre decrease in commuting distance on average. The reason for this may again be that local employment growth due to broadband leads to a larger share of local workers compared to longer distance commuters. In this model Mobile (3G) internet has no significant effect.

Column (2) and Figure 3 panel (a) show the results of a regression in which the effects of broadband are allowed to differ over distance categories from the nearest FUA core. The closest category includes the urban core itself. Figure 3 offers two broad conclusions. First, the effect of residential broadband availability on commuting distance is small and negative up until 30 kilometres from urban cores, while in the more distant places, from 50 kilometres and onwards, it is positive and slightly larger. In these more distant areas, the availability of broadband at the residence (blue line) leads to a 4.3 kilometres increase in commuting distance on average. This effect may well be driven by working from home (De Vos et al 2018). Second, the overall penetration of broadband at the place level (red line) is associated with longer commutes at intermediate distance from urban cores, between 20 and 30 kilometres.

This suggests that high speed internet makes the places at this distance part of the labour market area of the urban core, while they were previously not or only weakly integrated with this labour market. This is a clear sign of 'borrowed size' as these places are now able to reap the advantages of being part of the larger urban labour market. For the places up until 20km that were probably already included in this labour market, the general availability of broadband in a place does not alter the situation much. This also holds for the places at greater distance (50 and more kilometres away), that are still too distant to really integrate in the labour market.

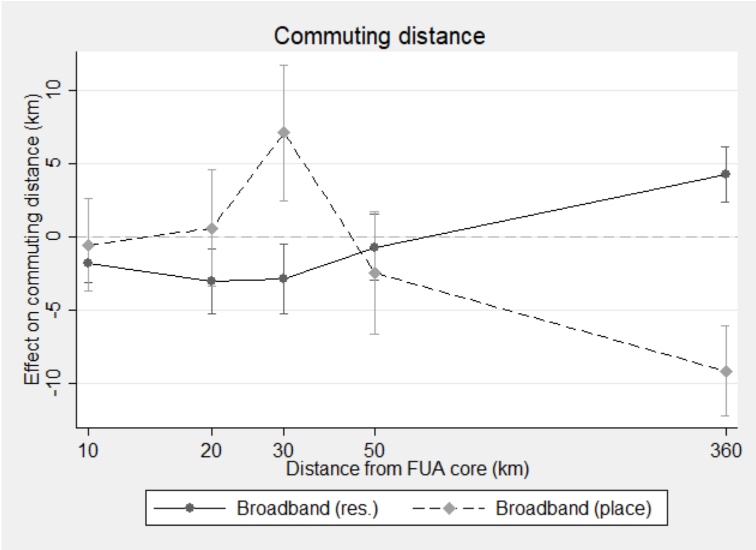
In column (3) we allow the effect of broadband on employment to differ across urban cores, urban regions, and rural regions. Here the only significant broadband effects come from place-level broadband penetration. In urban cores, increasing the broadband share with 0,1 results in 1.4% higher employment rates, while within FUAs outside of urban cores a 0.1 increase results in a much smaller 0.14% increase in employment. These results may indicate that broadband not only allows people to search and maintain jobs at longer distances, but also that it limits the extent of excess commuting at the local scale. In more remote places, such a borrowed size effect cannot be seen, suggesting that it is indeed available only in reasonable proximity to an urban core. The effect of Mobile (3G) internet is again small and similar to the result from Table 2 column (6).

Column (4) and Figure 3 panel (b) show the results of a regression in which the effects of broadband on employment are allowed to differ with distance from FUA cores. We can conclude from Figure 3 that broadband availability at the residential location (blue line) leads to increased employment rates only near and in the urban core; this is where broadband internet connection seems to lead to better matching. However, the effects of the overall penetration of broadband in places (red line) on employment rates are significant at intermediate distances from urban cores, between 20 and 50 kilometres. Again, this suggests that effect of broadband internet are largest in the 'expansion zone' of the labour market – the area that gets added to, or more integrated with, the urban labour market area through broadband.

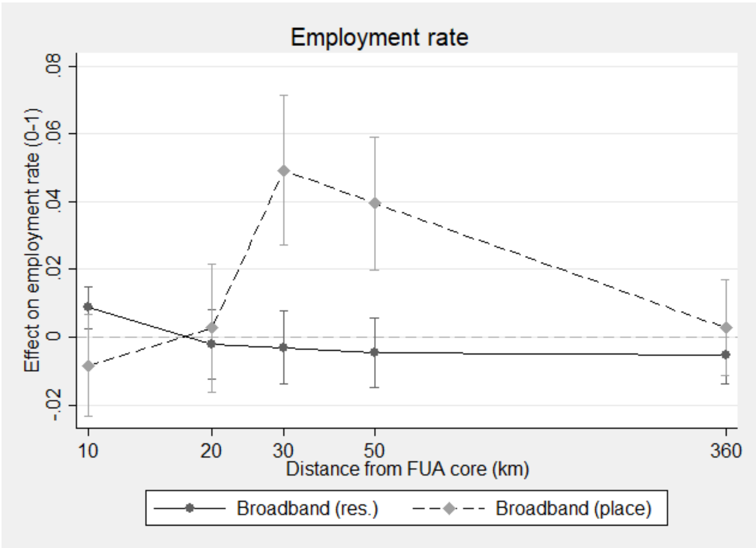
Table 3: Spatial heterogeneity

VARIABLES	(1) Commuting distance	(2)	(3) Employment rate	(4)
FUA Core				
Broadband (res.) (t-2)	-0.933		0.00365	
	(1.058)		(0.00506)	
Broadband (place) (t-2)	-1.605		0.143*	
	(15.33)		(0.0741)	
FUA outside core				
Broadband (res.) (t-2)	-2.175***		0.00223	
	(0.615)		(0.00288)	
Broadband (place) (t-2)	0.316		0.0140**	
	(1.191)		(0.00564)	
Outside FUA				
Broadband (res.) (t-2)	1.304*		-0.000732	
	(0.699)		(0.00323)	
Broadband (place) (t-2)	-4.325***		0.00529	
	(1.202)		(0.00562)	
Distance decay		See Fig. 3 (a)		See Fig 3 (b)
Mobile (3G) (t-2)	-0.0680	-0.0600	0.00205***	0.00215***
	(0.144)	(0.144)	(0.000690)	(0.000688)
Constant	-261.8*	-264.9*	1.693**	1.635**
	(158.6)	(158.6)	(0.763)	(0.763)
Control variables	Yes	Yes	Yes	Yes
Year and grid cell FE	Yes	Yes	Yes	Yes
Observations	634,033	634,033	646,345	646,345
R-squared	0.579	0.579	0.764	0.764

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Control variables include Population (log) and Income (log) at the grid cell level, and Population (log), Income (log), Share urban population, Share male, Share lower-, medium-, higher-, and college educated, Share young and old, No. of establishments, No. of employees per establishment, the aggregate income, and 21 industry share variables at the level of labour market areas.



(a)



(b)

Figure 3: Distance decay of broadband effects on commuting distance (a) and the employment rate (b)

Sensitivity analysis: Internet technologies – Up until now we have made no distinction in the speed of broadband internet. In this subsection we show the robustness of our results to accounting for the difference in speed between DSL technology (14 mbit/s) and Cable and Fibre technologies (100 mbit/s). In Table 4 we include 2 dummies that denote the *fastest fixed connection* at the residence, being either DSL or Cable/Fibre, instead of a simple broadband dummy. We also include the Mobile (3G) variable, and we allow it to vary over space. We still include the overall place-level broadband penetration rate as a separate variable, and again we lag the internet variables by two years. Column (1) again shows that the only positive effects of broadband at the residence are found outside FUAs, and it seems this is driven by Cable/Fibre technologies. Column (3) shows that the employment effects of broadband are driven by place-level broadband rates rather than residential availability. This effect is only present in intermediate areas (within FUA, outside core) and in these areas Mobile (3G) internet has a small but significant effect on employment. Figure 4 shows that the respective decay patterns of the effects of the different technologies on commuting distance (a) and employment (b) are very similar, to each other and to the earlier estimates. Compared to the fixed technologies, Mobile (3G) internet has a consistently smaller effect, and shows the least variability over space. Our results are thus robust to accounting for differences in speed of broadband technologies.

4.6 Conclusion

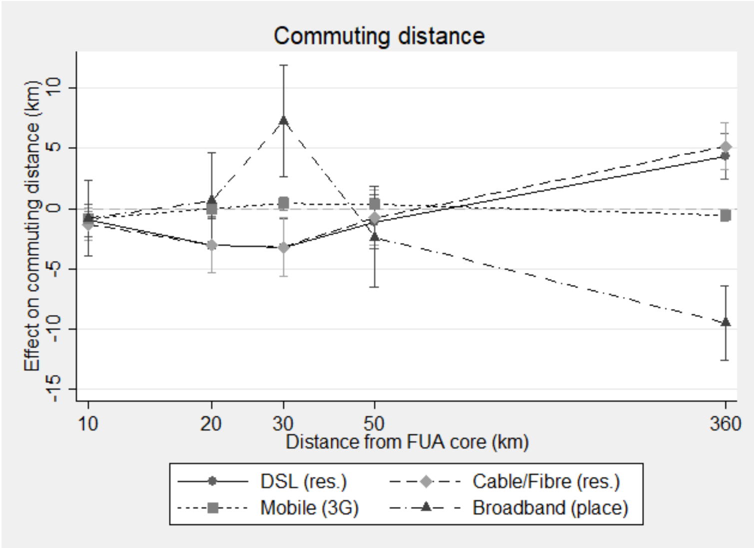
In this paper we have estimated the relation between broadband penetration and commuting behaviour and employment. The main goal was to test the hypothesis that broadband internet aids the process of borrowed size (Phelps 1995, Dijkstra et al 2013, Hesse 2016). To this end we estimated the effect separately for urban cores, urban areas, and rural areas, and we allowed the effect to differ with distance to the nearest core of a Functional Urban Area. We innovated by explicitly modelling a potential mechanism behind borrowed size (broadband), by using time-variation in broadband, and by distinguishing between broadband effects at the residence, and place-level.

The results suggest that residential broadband availability is associated with longer commutes in remote areas, which can be explained as a *working-from-home* effect (De Vos et al 2018).Within urban cores, residential broadband has a positive effect on employment, which suggests that broadband has enabled better labour market searching and matching in and closely around urban cores (Autor 2001). Considering the absence of significant results in proximity of urban cores, our results suggest no causal relation between residential broadband connections and ‘borrowed size’ in the labour market.

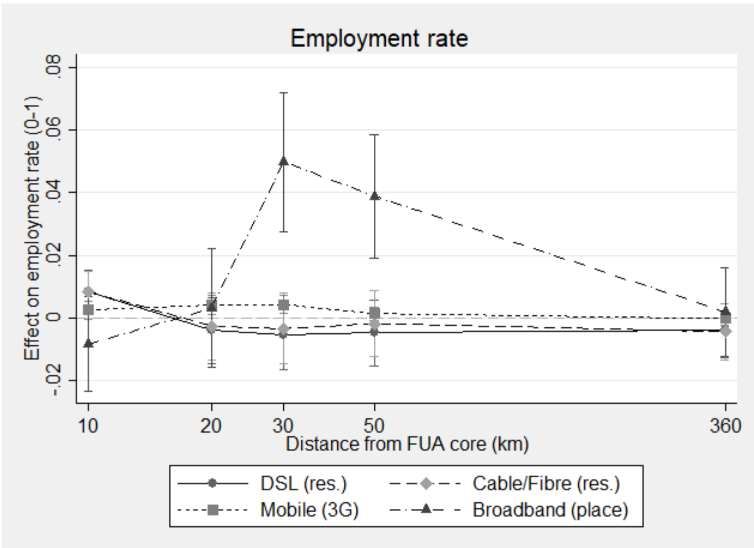
Table 4: Sensitivity analysis

VARIABLES	(1) Commuting distance	(2)	(3) Employment rate	(4)
FUA core				
DSL (t-2)	-0.880		0.00597	
	(1.062)		(0.00507)	
Cable/Fiber (t-2)	-1.098		0.00546	
	(1.057)		(0.00505)	
Local Broadband (t-2)	-4.092		0.150**	
	(15.44)		(0.0747)	
FUA outside core				
DSL (t-2)	-2.152***		0.00321	
	(0.610)		(0.00285)	
Cable/Fiber (t-2)	-2.350***		0.00479*	
	(0.617)		(0.00289)	
Local Broadband (t-2)	0.377		0.0128**	
	(1.189)		(0.00563)	
Outside FUA				
DSL (t-2)	1.095		0.000545	
	(0.695)		(0.00321)	
Cable/Fiber (t-2)	1.698**		0.00126	
	(0.711)		(0.00329)	
Local Broadband (t-2)	-4.373***		0.00376	
	(1.201)		(0.00561)	
Distance decay	See Fig. 4 (a)		See Fig 4 (b)	
3G (t)	-0.411	-0.378	0.00690***	0.00702***
	(0.435)	(0.435)	(0.00200)	(0.00200)
Constant	-253.8*	-257.9*	2.181***	2.147***
	(153.7)	(153.7)	(0.739)	(0.739)
Control variables	Yes	Yes	Yes	Yes
Year and grid cell FE	Yes	Yes	Yes	Yes
Observations	634,033	634,033	646,345	646,345
R-squared	0.579	0.579	0.764	0.764

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Control variables include Population (log) and Income (log) at the grid cell level, and Population (log), Income (log), Share urban population, Share male, Share lower-, medium-, higher-, and college educated, Share young and old, No. of establishments, No. of employees per establishment, the aggregate income, and 21 industry share variables at the level of labour market areas.



(a)



(b)

Figure 4: Distance decay of technology effects on commuting distance (a) and the employment rate (b)

The spatial variation in the effect of place-level broadband penetration on employment does show a clear borrowed size pattern. Growth in broadband penetration only affects employment rates in areas between 20 and 50 kilometres from urban cores, and commuting distances in areas between 20-30 kilometres from urban cores. This suggests that increased broadband penetration makes these places more attractive for firms and people to locate. From these locations they can tap into the benefits of the urban labour market nearby. Broadband internet thus seems an important driver behind the spatial expansion of the urban labour market. This is consistent with the assertion of Dijkstra et al (2013) that broadband access has aided the growth of small- and medium sized cities in proximity of larger cities.

There are limitations to the research approach of this paper that could inspire further research. First, Sweden is a specific country with a very high urbanization rate, combined with long distances between central and remote places. Future research should uncover whether our results also hold in other countries. Second, growth in broadband rates may be correlated with other dynamics that promote employment in cities, including new infrastructure and housing projects. We note that our control variables capture a lot of these potential confounders, but more research is needed into the causal nature of this effect. Third, we measure better matching in the labour market using the employment rate, but this may also translate into a better fit between workers and employers (lower overeducation, higher salaries). Future research may uncover whether there are also effects of broadband on these measures. Fourth, while we believe that working from home is a highly plausible explanation of the effect on broadband internet on commuting distance in remote regions, we lacked data on working from home rates or the number of commutes, that may give more conclusive evidence. Fifth, the small positive effects of Mobile (3G) internet on employment suggest that during the study period mobile internet was not a strong substitute for broadband internet, but with the arrival of 4G this may have changed. Finally, the concept of borrowed size needs further empirical substantiation, most notably explaining how local factors affect the capacity to borrow size and a better view of which agglomeration benefits can be 'borrowed'. And while this paper is one of the first to explicitly model an underlying, other studies could explore other mechanisms potentially driving processes of borrowed size.

We distil three policy implications from our research. First, if governments aim to increase the economic fortunes of places by introducing broadband, they should be aware that at least in the short run, the effects on employment are small, and vitally dependent on proximity to other employment centres. Second, our results suggest that governments should consider prioritizing fixed connections, as the effects of mobile internet are much smaller (but note that this was before the introduction of 4G). At the same time we do not see effects of broadband upgrades from DSL to Cable or Fibre. Finally, outside cities we find no effects of residential broadband availability on employment prospects for residents, but we do find significant effects of place-level broadband penetration on these prospects. This suggests that policies that aim to connect rural regions to fast internet to increase employment prospects should initially prioritize connecting places rather than connecting every single inhabitant.

References

Alonso, W. (1973). Urban Zero Population Growth. Daedalus, 102(4), 191–206. Retrieved from https://www.jstor.org/stable/20024174

Andersson Joona, P., & Wadensjö, E. (2013). The best and the brightest or the least successful? Self-employment entry among male wage-earners in Sweden. Small Business Economics 40(1), 155–172. https://doi.org/10.1007/s11187-011-9365-0

Atasoy, H. (2013). The effects of broadband internet expansion on labor market outcomes. ILR Review, 66(2), 315–345. http://doi.org/10.1177/001979391306600202

Autor, D. H. (2001). Wiring the Labor Market. Journal of Economic Perspectives, 15(1), 25–40. http://doi.org/10.1257/jep.15.1.25

Combes, P. P., & Lafourcade, M. (2005). Transport costs: Measures, determinants, and regional policy implications for France. Journal of Economic Geography, 5(3), 319–349. https://doi.org/10.1093/jnlecg/lbh062

Czernich, N. (2014). Does broadband internet reduce the unemployment rate? Evidence for Germany. Information Economics and Policy, 29, 32–45. http://doi.org/10.1016/j.infoecopol.2014.10.001

De Vos, D., Meijers, E., & van Ham, M. (2018). Working from home and the willingness to accept a longer commute. Annals of Regional Science 61(2), 375–398. http://doi.org/10.1007/s00168-018-0873-6

Dijkstra, L., Garcilazo, E., & McCann, P. (2013). The Economic Performance of European Cities and City Regions: Myths and Realities. European Planning Studies, 21(3), 334–354. http://doi.org/10.1080/09654313.2012.716245

Duranton, G., & Overman, H. G. (2005). Testing for localization using micro-geographic data: A new approach. The Review of Economic Studies, 72(4), 1077–1106. https://doi.org/10.1111/0034-6527.00362

Eriksson, R. H., Hansen, H. K., & Lindgren, U. (2014). The Importance of Business Climate and People Climate on Regional Performance. Regional Studies, 48(6), 1135–1155. http://doi.org/10.1080/00343404.2013.770140

Gaspar, J., & Glaeser, E. L. (1998). Information Technology and the Future of Cities. Journal of Urban Economics, 43(1), 136–156. http://doi.org/10.1006/juec.1996.2031

Government Offices of Sweden. (2017). A Completely Connected Sweden by 2025 − a Broadband Strategy. Retrieved from https://www.government.se/496173/contentassets/afe9f1cfeaac4e39abcdd3b82d9bee5d/sweden-completely-connected-by-2025-eng.pdf

Hesse, M. (2016). On borrowed size , flawed urbanisation and emerging enclave spaces : The exceptional urbanism of Luxembourg , Luxembourg. European

Urban and Regional Studies, 23(4), 612–627. http://doi.org/10.1177/0969776414528723
Kroft, K., & Pope, D. G. (2014). Does Online Search Crowd Out Traditional Search and Improve Matching Efficiency? Evidence from Craigslist. Journal of Labor Economics, 32(2), 259–303. http://doi.org/10.1086/673374
Kuhn, P., & Mansour, H. (2014). Is Internet Job Search Still Ineffective? Economic Journal, 124(581), 1213–1233. http://doi.org/10.1111/ecoj.12119
Kuhn, P., & Skuterud, M. (2004). Internet job search and unemployment durations. American Economic Review, 94(1), 218–232. http://doi.org/10.1257/000282804322970779
Mack, E. A. (2014). Businesses and the need for speed: The impact of broadband speed on business presence. Telematics and Informatics, 31(4), 617–627. https://doi.org/10.1016/j.tele.2013.12.001
Mack, E. a, & Grubesic, T. H. (2009). Broadband Provision and Firm Location in Ohio: An Exploratory Spatial Analysis. Tijdschrift Voor Economische En Sociale Geografie, 100(3), 298–315. https://doi.org/10.1111/j.1467-9663.2008.00487.x
Meijers, E. J., & Burger, M. J. (2017). Stretching the concept of 'borrowed size.' Urban Studies, 54(1), 269–291. http://doi.org/10.1177/0042098015597642
Meijers, E. J., Burger, M. J., & Hoogerbrugge, M. M. (2015). Borrowing size in networks of cities: City size, network connectivity and metropolitan functions in Europe. Papers in Regional Science, 95(1), 181–198. http://doi.org/10.1111/pirs.12181
Meijers, E. J., Hoogerbrugge, M., & Cardoso, R. (2018). Beyond polycentricity: Does stronger integration between cities in polycentric urban regions improve performance? Tijdschrift Voor Economische En Sociale Geografie, 109(1), 1–21. http://doi.org/10.1111/tesg.12292
Phelps, N. A. (1992). External Economies, Agglomeration and Flexible Accumulation. Transactions of the Institute of British Geographers, 17(1), 35–46. http://doi.org/10.2307/622635
Phelps, N. A. (1998). On the edge of something big: Edge-City Economic Development in Croydon, South London. The Town Planning Review, 69(4), 441–465. https://doi.org/10.3828/tpr.69.4.dv1t387m20078jjp
Phelps, N. A., Fallon, R. J., & Williams, C. L. (2001). Small firms, borrowed size and the urban-rural shift. Regional Studies, 35(7), 613–624. http://doi.org/10.1080/00343400120075885
Sinai, T., & Waldfogel, J. (2004). Geography and the Internet: Is the Internet a Substitute or a Complement for Cities? Journal of Urban Economics, 56(1), 1–24. http://doi.org/10.1016/j.jue.2004.04.001
Sohn, J., Kim, T. J., & Hewings, G. J. D. (2003). Information technology and urban spatial structure: A comparative analysis of the Chicago and Seoul regions. The Annals of Regional Science, 37, 447–462. https://doi.org/10.1007/3-540-28351-X_18
Tranos, E., & Mack, E. A. (2016). Broadband Provision and Knowledge-Intensive Firms: A Causal Relationship? Regional Studies, 50(7), 1113–1126. https://doi.org/10.1080/00343404.2014.965136

Vilhelmson, B., & Thulin, E. (2016). Who and where are the flexible workers? Exploring the current diffusion of telework in Sweden. New Technology, Work and Employment, 31(1), 77–96. http://doi.org/10.1111/ntwe.12060

Appendix A: Functional Urban Areas

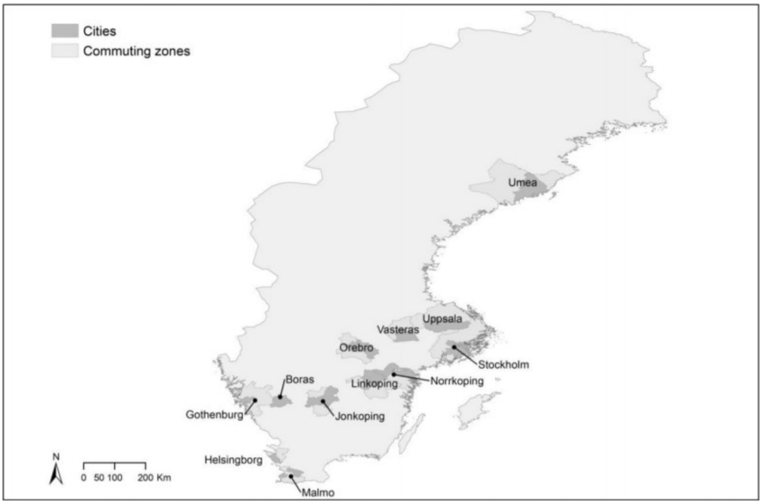


Figure 5: Functional Urban Areas in Sweden. Source: www.oecd.org/cfe/regional-policy/Sweden.pdf

Part 2: Local product markets

5. Information technology and local product variety: Substitution, complementarity and spillovers

Published in Tijdschrift voor Economische en Sociale Geografie 110.4 (2019).
Written by Duco de Vos and Evert Meijers.

Abstract
This paper addresses the interaction between information technology (IT) and agglomeration. The literature distinguishes two types of interactions, namely a substitution effect and a complementarity effect. We conceptualise a third effect, namely a 'spillover' mechanism, by which IT allows places in close proximity of large cities to 'borrow size' and sustain greater product variety. We test these mechanisms using detailed data on restaurant cuisine variety in the Netherlands, and the IT dimension is measured through the use and penetration of online restaurant reviews. We find that IT complements cuisine variety in cities, and induces spillovers to smaller places near larger ones, allowing smaller places to sustain 'rare' cuisines that were traditionally only present in larger cities. As such, IT leads to the spread of agglomeration benefits such as local product variety over larger territories.

5.1 Introduction

Many industries exhibit a positive relationship between market size and product variety. A wide variety of local products is therefore seen as a typical urban benefit (Glaeser *et al.* 2001; Rosenthal & Strange 2004). The way in which information technology (IT) interacts with this local market size externality is part of a larger, polarised debate on the interaction between agglomeration economies and IT, that is in need of new empirical substantiation now that (mobile) access to digital information and communication channels has reached unprecedented levels in recent years.

The literature that studies the interplay between IT and offline local variety offers mixed results. On the one hand, it is shown that internet technologies are more intensively used in rural areas, where offline alternatives of many services are not at hand (Sinai & Waldfogel 2004). On the other hand, it is often argued that the internet is more useful and beneficial in urban environments, where it can provide information on the bewildering amount of opportunities nearby, spurring demand that leads to the entry of new varieties of local products (Anenberg & Kung 2015).

Geographically, the essence of IT is that it eliminates spatial information frictions (Jensen 2007): access to most forms of information is almost ubiquitous. Decreasing spatial information frictions may not only complement the benefits of agglomeration in cities themselves, but may also allow these benefits to spill over to adjacent areas. By disseminating information about specific destinations, online information may decrease the uncertainty associated with travelling to these destinations, and thereby affect the generalised costs of travel (Mokhtarian & Tal 2013). If such an effect were sufficiently high, we would expect that places adjacent to large cities are increasingly able to capitalize on the population of their neighbours, for instance by drawing on their critical mass to sustain more specialised urban functions for which they themselves do not provide sufficient mass. Alonso (1973) coined the term 'borrowed size' for such spillover processes that weaken the strong relation between size and function of places as for instance predicted by Central Place Theory (Christaller 1966). Within the debate on the substitution- or complementary effect of IT on cities, the option of 'borrowed size' has not been considered.

We investigate the interaction between IT and local product variety using the restaurant industry in the Netherlands as a case study. This market is known for its strong relationship between market size and horizontal product differentiation (Berry and Waldfogel 2010), and the presence of a large variety of restaurant cuisines is understood as an urban consumption benefit (Couture 2013; Glaeser et al. 2001; Schiff 2015).[1] The restaurant market is also heavily affected by internet technologies. With the rise of IT, a large part of marketing activities of restaurants has moved online. Moreover, this market enjoys extensive online word-of-mouth, through major review websites such as Yelp, TripAdvisor, and national level counterparts, such as Iens.nl in the Netherlands (the geographical focus of our research). Research has shown that online reviews substantially affect purchase behaviour in markets for 'experience goods' in general (Huang *et al.* 2009), and in restaurant markets in particular (Anderson & Magruder 2011). Data from Google Trends shows that since 2004 there has been a dramatic increase in the use of restaurant review databases in the Netherlands, most notably the culinary website Iens.nl, and the general tourism review website TripAdvisor (see Figure 1).[2]

[1] Growing demand does not necessarily lead to higher product variety. Shaked & Sutton (1987) propose that this depends crucially on whether quality is predominantly produced with fixed rather than variable costs. In the former case quality tends to increase with market size while the industry remains concentrated, in the latter case the range of qualities tends to increase with market size. Berry & Waldfogel (2010) take this prediction to the data, and find that newspaper markets tend to remain concentrated, while restaurant markets offer more variety with increasing market size. Schiff (2015) shows that the relationship between market size and variety holds both with a general 'love of variety', and in the case that every consumer has a preferred restaurant type, given that there is a taste hierarchy.

[2] Since September 2016, Iens.nl has been taken over by TripAdvisor. This explains the sudden bifurcation of the two trends in Figure 1 around the end of 2016.

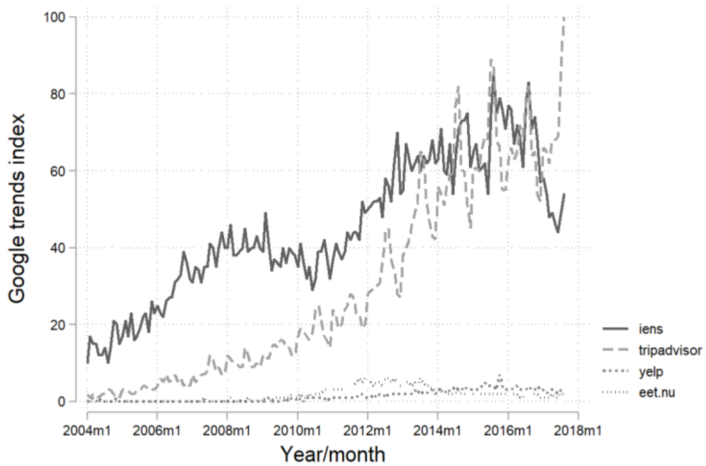


Figure 1. Google trends results for online review databases in the Netherlands.

The research question guiding this paper is as follows: to what extent does IT interact with agglomeration economies of consumption? Thereby, we distinguish and explore the presence of three main mechanisms:

1. A substitution effect; this occurs when IT is of greater benefit in smaller places, e.g. because online information has a higher value when offline substitutes such as physical comparison shopping and word-of-mouth are less available.
2. A complementary effect; this occurs when IT is of greater benefit in larger places, e.g. because online information on local goods is more valuable in dense, congested areas.
3. A spillover effect; this occurs when IT allows smaller places in the vicinity of larger places to 'borrow size', and sustain more agglomeration benefits locally.

By answering our research question we contribute to at least three strands of literature. First, we add to the literature that deals with the question whether IT substitutes or complements the consumption benefits of cities by examining the effect of a *direct* measure of online information on cuisine variety (Anenberg & Kung 2015; Gaspar & Glaeser 1998; Sinai & Waldfogel 2004). Second, we add to the literature that investigates the alternative sources of agglomeration externalities through borrowed size (Meijers 2007; Meijers & Burger 2017; Phelps *et al.* 2001) by highlighting potential ways in which the extent of markets may increase and enable positive agglomeration spillovers between places. Finally, we test whether the earlier established relationship between cuisine variety and local population in American metropolitan areas (Couture 2013; Schiff 2015; Waldfogel 2008) also holds in the Netherlands, and whether it is also present in a non-metropolitan context. Where earlier research was mainly focused on metropolitan areas, our analysis includes all places in the Netherlands with at least 1 restaurant.

The remainder of this paper is structured as follows. The next section further elaborates on the mechanisms by which online reviews interact with agglomeration and product variety. In the third section we discuss the data. In fourth section we present our empirical approach. In the fifth section we present the results, and we subject our results to several sensitivity checks. The final section concludes.

5.2 Agglomeration, online reviews, and cuisine variety

During the 1990s and early 2000s several authors declared that the ‘death of distance’ was near (Cairncross 1997), the world was increasingly becoming ‘flat’ (Friedman 2005), and that geography and proximity seemed to matter less and less, as a result of advances in IT and continuing globalisation. The economic and geographic literature concerned with this question continued to emphasize the ways in which distance still seems to matter (Goldfarb & Tucker 2017). It is, for instance, generally argued that tele- or online communication complements face-to-face encounters, and that the increasing need for encounters offsets the decentralizing effects of IT (Gaspar & Glaeser 1998; Leamer & Storper 2001). Next to that, it is argued that while the usage of IT is usually higher when real-world offline options are less available, for instance in rural areas, the welfare benefits of IT are greater in urban areas where information on local offline options is more valuable because it spurs demand for local goods, and allows the entry of new varieties (Anenberg & Kung 2015; Sinai & Waldfogel 2004).

The rise of online reviewing of restaurants is a prime example of how information technology has made information on local consumption options accessible. We believe there are three mechanisms by which these online reviews interact with agglomeration economies of consumption in the restaurant market.

First, online reviews may provide a substitute for word-of-mouth advertising and comparison shopping, traditionally fostered by physical proximity. Word-of-mouth advertising is local by definition. It requires face-to-face contact, and it can only involve restaurants that have been visited in person, which is more likely for restaurants close by. The effect of word-of-mouth may be declining, in favour of online reviews. Similarly, online reviews substitute physical comparison shopping: one of the traditional reasons for restaurants to cluster together in cities. By allowing potential customers a detailed comparison of restaurants before arriving at the scene, online presence in general, and online reviews in particular may substitute the need for restaurants to locate near one another. Thus the attractiveness of places where restaurants cluster – typically cities – would decrease in favour of places with a lower restaurant density. Following this mechanism the advent of online restaurant reviewing would substitute agglomeration economies of consumption.

Second, online reviews may be of greater value and effect in cities. It is argued that one of the main effects of IT is the reduction of spatial information frictions (Jensen 2007). Information about local goods has traditionally depended critically on proximity. By loosening the relationship between proximity and information, online information in general, and online reviews in particular may lead to increased demand for local goods, and thereby spur product variety. If, as Anenberg and Kung (2015) argue, spatial information frictions are more abundant in dense, congested cities with a bewildering amount of offline opportunities, then the variety economies of cities may be complemented by IT. Anenberg and Kung (2015) find such evidence for the food truck industry in metropolitan areas in the US.

Third, online reviews may influence the willingness to travel for restaurants, not only within, but also between places. In the transportation literature it is well established that ICT does not only substitute making trips, but also complements travel in many ways. Particularly, IT has a role in "disseminating information about specific destinations" (Mokhtarian & Tal 2013). By providing extensive information about the quality of restaurants, culinary review websites circulate such information. This may well increase the spatial extent of restaurant markets, because people are willing to travel further for restaurant visits if they face less uncertainty about their offer, while navigation devices make them easier to find. This may give rise to a complementary effect, because cities can draw better upon population in the surroundings. However, in places not too distant from large cities, an increased spatial extent of restaurant markets will entail a greater increase in consumer potential, compared to places in scarcely populated areas. This affects the location choice of new restaurants, for whom smaller places in the vicinity of larger ones have become more attractive. Evidence from Couture (2013) shows that almost all benefits of density in the restaurant market arise from the fact that people living in dense urban areas visit more restaurants, rather than from travel time savings. In the same manner, the abundance of information on the internet may induce people to travel further for restaurant visits, rather than to visit nearby restaurants more often.

This spillover mechanism may give body to the claims of several authors that the traditional benefits of urban size are increasingly substituted by 'borrowed size' (e.g. Meijers 2007; Meijers & Burger 2017; Phelps *et al.* 2001). This literature suggests that higher-order functions are increasingly found in places that are smaller than (central place) theory would suggest. They argue that the central place hierarchy of functions between cities – in which the most populated places tend to have the most higher order functions – is declining, in favour of stronger, more complex and complementary relationships between cities causing a decoupling of size and function.

To investigate the existence and interactions of the proposed mechanisms (substitute, complement, spatial spillover effect) we focus on two relationships. First, to investigate whether there is predominantly substitution or complementarity between online reviews and agglomeration economies of variety, we estimate the effect of online review intensity on cuisine variety, and we assess whether this effect is greater in larger or in smaller places. Second, to investigate whether online reviews enable places in the vicinity of larger cities to borrow size, and thus sustain a greater variety of cuisines than one would expect given their size, we examine to what extent the effect of online reviews on cuisine variety depends on distance to the nearest population centre.

5.3 Data

Background – The history of restaurant cuisine variety in the Netherlands started in cities. The first Chinese restaurant in the Netherlands was opened in 1920 in Rotterdam, and the first Chinese restaurant that was aimed at Dutch customers was opened in 1928 in Amsterdam.[3] An excess of Chinese cooks during the economic crisis of the 1930s, and the increasing demand for Asian food by colonists returning from Indonesia stimulated the subsequent rise of Chinese restaurants: by 1945 there were 30 Chinese restaurants in the Netherlands, and by 1965 there were 225, of which many advertised themselves as Chinese-Indonesian (Van Voskuilen 2013). Other ethnic cuisines followed during the 1960s and 1970s, as a result of labour migration from southern Europe (Italian, Greek, Spanish, Turkish) and decolonisation (Surinam).

Next to the specificities related to trade, colonialism, and work migration, there are a few other peculiarities concerning restaurant cuisine variety in the Netherlands. First, this market may be particularly prone to diversification because of a weak national food identity (Terhorst & Erkuş-Öztürk 2018; DeSoucey 2010). Indeed, the limitations of the Dutch cuisine have led to many crossovers with other cuisines, most notably the French. Second, the geography of restaurant variety in the Netherlands is not only shaped by the market. In the 1980s, the Dutch Ministry of Economic Affairs devised special rules that allow maximally 1 foreign owned Chinese-Indonesian restaurant for each 10.000 inhabitants in a municipality (Koopman 2002). Similar rules continue to apply nowadays: new migrants that apply for a permit in the Netherlands need to fulfil several characteristics that show that they provide a genuine

[3] 'Cheung Kwok Low' was the first Chinese restaurant in the Netherlands, with a clientele consisting of Chinese sailors (http://www.isgeschiedenis.nl/nieuws/de-geschiedenis-van-de-afhaalchinees-in-nederland/). 'Kong Hing' in Amsterdam was the first restaurant that actively aimed to serve Dutch customers (http://www.iisg.nl/collections/chinezen-zeedijk/chinezen-nl.php).

merit to the Dutch economy.[4] EU citizens, on the other hand, have the freedom to start a business in the Netherlands since the Rome Treaty,[5] conditional on some provisions. De Lange (2016) notes that, due to trade agreements, entrepreneurs from Japan and the United States are almost as privileged as EU citizens in this respect. Finally, there are several other notable cuisine varieties in the Netherlands, not necessarily related to decolonisation, large-scale immigration or specific trade agreements, but rather to globalization. These include Argentinian steakhouses, Mexican *cantinas*, Indian curry houses, and Thai and Vietnamese restaurants. Recent research into the Dutch restaurant industry calls attention to the role of urbanisation and gentrification in the continuing growth and diversification of this market (Terhorst & Erkuş-Öztürk 2018), and in the changing geography of haute cuisine restaurants (Boterman 2018).

Measuring cuisine variety – To study the current dynamics in the geography of cuisine variety in the Netherlands, we collected data on the universe of restaurants in the Netherlands, in 2000 and 2016, subdivided into 20 cuisine categories that broadly refer to different nationalities. The data are provided by HorecaDNA, and is constructed using the business registry of the Dutch Chamber of Commerce (Kamer van Koophandel). The information is available at the 4 digit postcode level (9 km^2 on average), and also includes the number of inhabitants at this scale. We recoded the data to the level of places (villages, towns, and cities in the Netherlands). There are 2,456 such settlements in the Netherlands. Population data is available for all places with more than 100 inhabitants, and Amsterdam had the largest population (747,725) in 2016. Summary statistics of these data are in Appendix 1. For the main analysis we use the same data on the number of cuisines in settlements, but only for the year 2016. Here we consider it as a continuous measure of cuisine variety. Figure 2 shows the distribution of this variable.

While this categorization of 20 cuisine types is not exhaustive, and only measures variety across (groups of) nationalities, it is the only categorisation that covers all restaurants in the Netherlands, and it is consistent across places. We found that this categorization correlates highly (0.9405) with a measure of cuisine variety based on 92 cuisine types, calculated from the Iens.nl data rather than from registry data. We do find that our measure of cuisine variety underestimates variety in places with the highest variety. Therefore, in our sensitivity analysis we check whether our results change when we exclude large cities with a high variety.

[4] "Wezenlijk Nederlands belang", see Artikel 13 sub b Vreemdelingenwet 2000.
[5] Treaty Establishing the European Economic Community, Mar. 25, 1957

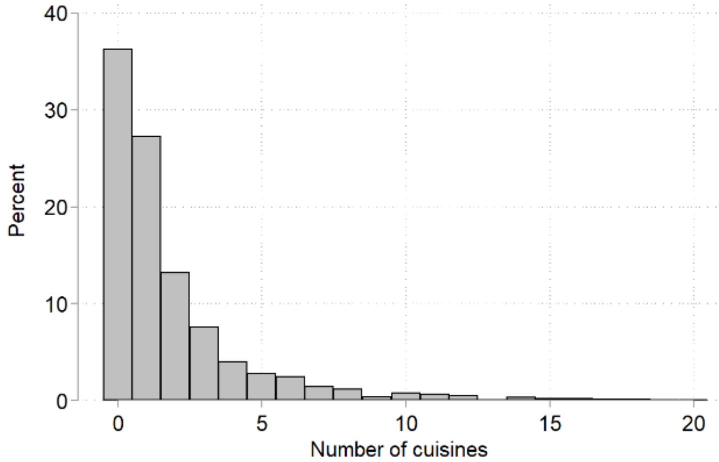


Figure 2. Histogram of the number of cuisines and the share of reviewed restaurants.

Measuring information technology – We measure information technology as the relative amount of online information on local restaurants, for each place in our dataset. We do this by calculating the share of restaurants in a place with at least 1 review on Iens.nl. Given the prominence of Iens.nl at the time of collecting the data (see Figure 1), we argue this is a reasonable proxy for online presence across all online channels. We scraped the entire website, consisting of georeferenced information of 21,795 Dutch restaurants, on September 8th, 2016. Together with the detailed information on the total amount of restaurants in a settlement we are thus able to calculate the share of reviewed restaurants.[6] Figure 3 shows the distribution of this variable.

The indicator for the share of reviewed restaurants may depend on the turnover rate of restaurants. This may bias our results if the turnover rate is related to variety. One could indeed argue that vivid places with a high cuisine variety probably have a higher turnover rate. Since brand new restaurants may not be present on Iens.nl immediately, we run the risk of underestimating information technology presence in places with the most variety. This will bias our results downward, so in this sense our estimates may be seen as a lower bound of the real effect. We do believe that the extent of this bias will be limited, because in 2016 Iens.nl deleted reviews that were older than 4 years.

[6] Because we measure the universe of restaurants at the beginning of 2016, and the number restaurants registered at Iens.nl more than half a year later, there may be a slight measurement error. Indeed, 23 out of 2,249 places show a share of reviewed restaurants higher than 1, which we normalise to 1. This does not affect the results.

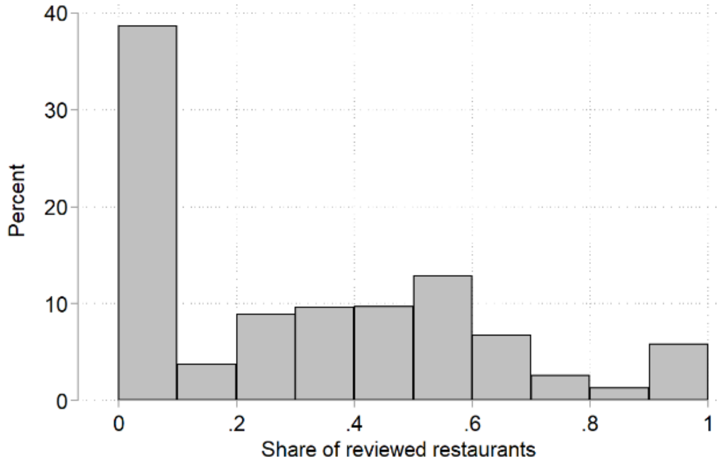


Figure 3. Histogram of the share of reviewed restaurants.

Data description – We enrich the data with demographic composition characteristics from Netherlands Statistics (CBS). We include household size, household income, the share of high educated-, young- and old people, and the share of western- and non-western foreigners. For household income and the share of higher educated, data from 2016 is missing, and we use the income level of the nearest preceding year available (2012). Information on the share of high educated people is only available at the municipality level, and only for the year 2015. Netherlands Statistics' definition of western foreigners refers to persons of whom at least one parent is born outside of the Netherlands, in either Europe (excl. Turkey), North America, Oceania, Indonesia or Japan. Other foreigners include people with at least one foreign born parent from any other country. Furthermore we use data on the amount of hotel beds (also from HorecaDNA.nl).

The summary statistics are presented in Table 1. Dutch places have a modest variety of cuisines on average, with 1.88 cuisines on 20.16 restaurants. In 2000 the average number of cuisines was 1.56 so there has been a slight increase. The average share of reviewed restaurants is 30 percent (43 percent for all settlements with at least 1 restaurant), with considerable variation between places, and about 0.02 reviews per capita. In Appendix 2 we report summary statistics for the part of the sample for which the (historical) number of cuisines and the average household income are greater than zero, because we only use this set of data in the analysis. This leaves us 1,280 places, about 57 percent of all places. Arguably, excluding all (genuinely tiny) places without restaurants should not be problematic, because the share of reviewed restaurants is only informative if there is at least 1 restaurant.

Table 1. Summary statistics.

VARIABLES	(1) mean	(2) sd	(3) min	(4) p50	(5) max
Number of cuisines	1.881	2.814	0	1	20
Number of cuisines 2000	1.557	2.268	0	1	20
Number of restaurants	20.16	112.7	0	5	4,088
Number of restaurants/capita	0.00351	0.00499	0	0.00240	0.0741
Hotel beds/capita	0.0232	0.106	0	0	2.477
Share of reviewed restaurants	0.301	0.299	0	0.286	1
Number of reviews/capita	0.0196	0.0561	0	0.00200	0.992
Population	7,540	28,831	100	1,690	747,725
Population density	0.397	0.646	0.00478	0.165	6.415
Area (hectare)	1,517	1,641	18.92	1,070	26,058
Average household size	2.393	0.216	1.670	2.380	3.420
Average household income	34,352	10,616	0	36,300	81,800
Share high educated	24.29	7.046	0	23.70	96.80
Share young	0.163	0.0306	0.0345	0.162	0.349
Share old	0.197	0.0459	0.0333	0.194	0.581
Share western foreign	0.0653	0.0476	0	0.0549	0.624
Share nonwestern foreign	0.0306	0.0411	0	0.0204	0.675
N	2,249				

5.4 Methods

Empirical approach – We begin our empirical analysis by exploring how cuisine variety changed across places, between 2000 and 2016. We use maps of cuisine variety, and we plot the smoothed relationship between population and cuisine variety for both years, to examine the geographical pattern of changes in cuisine variety, and to what extent it varies with the size of local population. As during this period there were many major IT advances (broadband internet, Wi-Fi, 3G, smartphones, etc.), the patterns we find can give us some initial expectations about how IT affected cuisine variety.

In the exploratory part we also investigate to what extent the distribution of cuisines across places conforms to a *central-place hierarchy*, according to which the rarest cuisines are found in places that have all other, more common cuisines. For this purpose we use a hierarchy diagram and a hierarchy statistic, initially proposed by Mori et al. (2008). The hierarchy statistic captures the share of cuisine-place pairs for which all places with more cuisines, also have the respective cuisine. Following Schiff (2015) we test the null hypothesis that cuisines are randomly assigned across places. We do this by running 1,000 simulations in which we randomly draw cuisine types for every settlement, based on the observed number of cuisines, estimating the distribution of the hierarchy statistic in a counterfactual situation in which a place with only one cuisine should not be more likely to have an Italian cuisine, than a Japanese cuisine.

The next step is to investigate the relationship between the share of reviewed restaurants and local cuisine variety. The idea is that if relatively more restaurants in a place are reviewed at this website, online information dissemination is more effective, and demand increases, spurring the entry of new varieties. Following Schiff (2015) and Berry and Waldfogel (2010), the dependent variable is the logarithm of the number of cuisines. The independent variable that denotes population is also in logs, so the resulting coefficient can be interpreted as an elasticity. The specification is as follows:

$$\log(C_i) = \alpha + \beta \log(Pop_i) + \gamma \log(Area) + \delta R_i + \kappa X_i + \epsilon_i \qquad (2)$$

where C is the number of cuisines. We add an interaction between the share of reviewed restaurants and population to assess whether the effects of reviews on variety are greater in larger or in smaller places. Our conclusion depends on the value of $\partial^2 E[\log(C) \mid \log(Pop), R]/\partial R \partial \log(Pop)$. If this is positive, IT and local population complement each other in generating local variety, if this is negative the effect of IT on local variety is greater in smaller places.

Finally we investigate whether IT can be linked to increased spillovers between places. This is the third mechanism through which IT may impact agglomeration. Here we allow the effect of IT on cuisine variety to vary across distance from nearby population centres ($Dist_i$), using different population thresholds that define such a centre (20k, 50k, and 100k). If IT is associated with spillovers between places – that is with increased demand from nearby places – we expect the effects of IT to be greater in the vicinity of large population centres, i.e. $\partial^2 E[\log(C)|R, Dist]/\partial R \partial Dist < 0$.

Identification issues – We aim to measure the causal effect of the share of reviewed restaurants on cuisine variety. However, this relationship poses several identification problems, including a potential bias due to omitted variables (Angrist & Pischke 2008). In our regression models therefore, we control for a variety of socio-demographic variables. In addition, we include the number of hotel beds per capita to control for the effects of elevated demand due to local tourism.

There remain other issues. Severe local competition may for instance both trigger restaurants to differentiate their cuisine from other restaurants, and to register at online review databases. Evidence from the management literature indeed shows that the extent of online information about products can be related to the extent of competition (Li *et al.* 2011). We resolve this issue by taking two steps. First, by focussing on the share of *reviewed* restaurants we have an independent variable that cannot be easily controlled by restaurant owners. While restaurant owners may in principle register their restaurant at Iens.nl, reviews have to come from customers, and registration at

Iens.nl is not required to leave a review.[7] Second, we control for the extent of local competition by including the number of restaurants per capita as a proxy.

There also exists the possibility of selection. Online reviewers may for instance be disproportionately attracted to cities with a culinary culture, with a lot of restaurants and different cuisines. Following this reasoning, typical ‘foodie cities’ would have both more cuisines and more reviewed restaurants, without there being a causal relation from IT to variety. An ideal strategy to control for such time-invariant characteristics of places would be to use temporal variation to estimate the effect. Unfortunately we only have cuisine data at the level of places until 2016, and Iens.nl data starting 2016. Therefore we employ an alternative strategy in which we include the cuisine count from 2000 as a proxy for ‘culinary culture’. The key assumption here is that variation in the share of reviewed restaurants is random, conditional on cuisine variety in 2000 and the other control variables. Since the Iens.nl website started in 1999, cuisine variety in 2000 is likely unaffected by Iens.nl reviews.

Finally there is the issue of reverse causality. This would be the case if greater variety in cuisines would trigger consumers to leave reviews, or alternatively, if restaurants with rare cuisines are somehow more attractive to review. The ideal strategy to deal with this type of bias is to find an instrumental variable that only affects cuisine variety through its effect on the share of reviewed restaurants. We experimented with a number of candidates, including the number of reviews per capita, the penetration of glass-fibre internet, and distance to Amsterdam (the city where Iens.nl started). We could not find an instrument that was both strong and clearly exogenous to cuisine variety. We also experimented with a model based on time variation in reviews and variety, using municipality level data between 2016 and 2017, but there was not enough variation for any significant results. While we are confident that controlling for ‘foodie cities’ captures a large part of the reverse causality problem, it cannot be ruled out that changes in cuisine variety between 2000 and 2016 affected the share of reviewed restaurants rather than the other way around. Our results remain sensitive to this issue, which would bias our estimates upwards.

We ignore potential reverse causality between cuisine variety and local population because earlier results on the relationship between population and cuisine variety suggest that controlling for reverse causality does not alter results (Schiff 2015), in line with results from e.g. Chen and Rosenthal (2008) who find that local amenities, among which cuisine variety, are in fact a minor pull-factor of cities compared to business factors.

[7] A potential bias may still exist if Iens registration lowers the bar for people to leave a review, and restaurant owners with rare cuisines may push customers more to leave reviews. As we control for competition, this bias is only relevant in case of registrations by restaurants not triggered by competition.

5.5 Results

In this section we proceed with the empirical analysis. First we explore changes in the geographical distribution of cuisines between 2000 and 2016. Then we estimate the effect of the share of reviewed restaurants on cuisine variety, and we investigate to what extent this effect varies across smaller and larger places, and with distance from population centres. Finally, we subject our results to several sensitivity checks.

Cuisine variety across places – Figure 4 shows the geographic distribution of cuisine variety over Dutch settlements. This map clearly shows that cuisine variety rose countrywide between 2000 and 2016, and that cuisine variety seems to have radiated out from places that had a varied supply of cuisines in 2000. We examine whether the (observational) relationship between municipal population and cuisine variety has significantly changed during this period. Figure 5 depicts the fitted relationship between (the logarithm of) population and cuisines, using a local polynomial function. For places with populations between 1,000 and 100,000, cuisine variety has increased significantly.

Figure 6 shows hierarchy diagrams for the year 2000 and 2016. Here cuisine-place pairs are ranked according to (1) the rarity of the cuisine (y-axis, 0 is rarest), and (2) the amount of cuisines in a place (x-axis). Both years show a clear hierarchical pattern in which rarer cuisines are more likely to be present in places that have more cuisines. Following Mori et al. (2008), we use the term 'hierarchy event' for cuisine-place pairs for which all higher ranked places also have the respective cuisine, and 'non-hierarchy event' for all other pairs. The difference in hierarchy for both years can be assessed using a hierarchy statistic that denotes the share of hierarchy events. This is equivalent to the share of dots in the graph for which all 'slots' on the right are occupied. In 2000 23.2 percent of cuisine-municipality pairs were hierarchy events, versus 18.2 percent in 2016. While both years show a significant hierarchy compared to the hierarchy that results from a random cuisine distribution (7.1%), the difference in hierarchy between 2000 and 2016 is significant (p=0.000) and this means that the functional hierarchy cuisine variety across cities has decreased during this period.

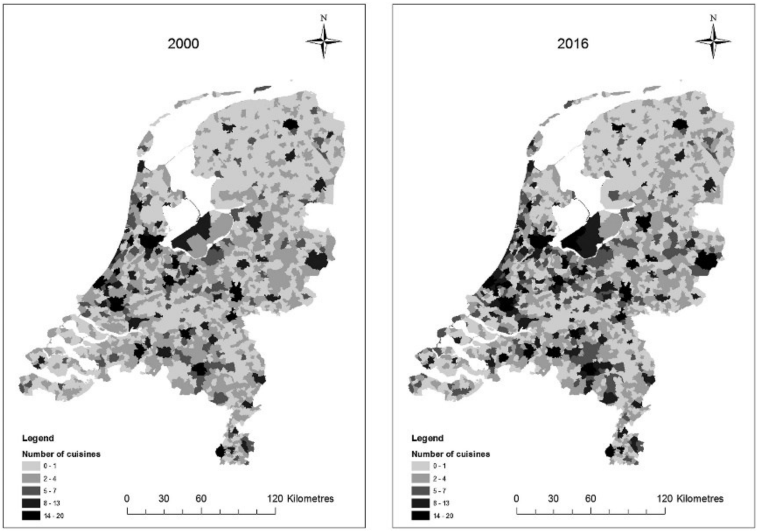


Figure 4. Maps of cuisine variety.

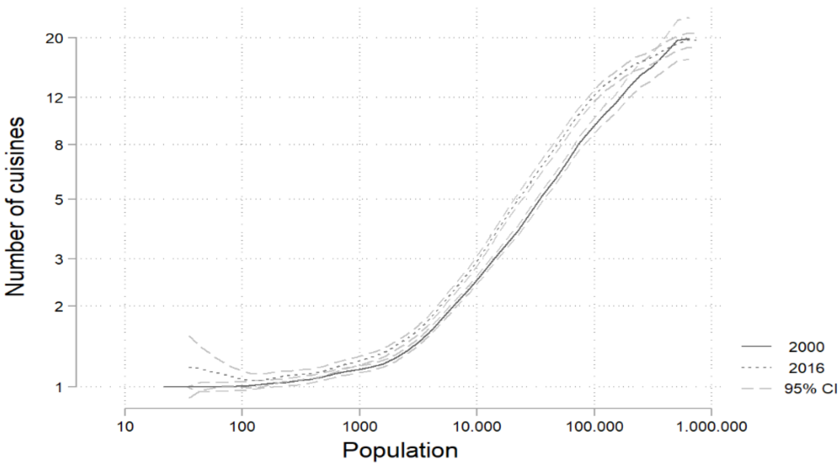


Figure 5. Nonparametric fit of the relationship between population and cuisine variety in Dutch places in 2000 and 2016.

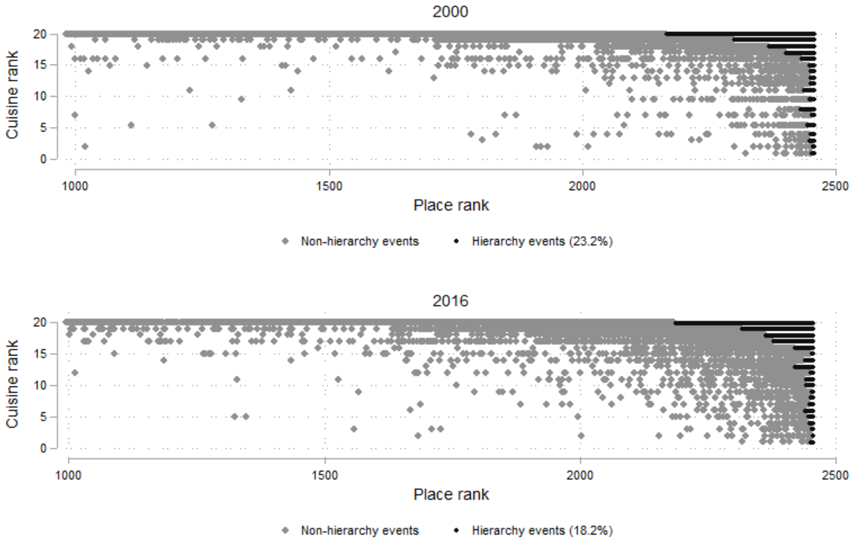


Figure 6. Hierarchy diagrams, 2000 and 2016. The observations are instances in which a cuisine is present in a place. Along the y-axis, cuisines are ranked according to their year specific rarity (where 1 denotes the rarest cuisine). Along the x-axis, places are ranked according to the number of cuisines they have. The place rank starts at 996, because 995 places have no cuisines at all.

We draw three conclusions from this exploratory analysis. First, the finding that cuisine variety seems to have grown most extensively near places where variety was already high is in line with the hypothesis that the advent of IT in the restaurant industry led to an increase in the geographical extent of restaurant markets, and increased spillovers between places. Second, the functional form, and size of the relationship between local population and variety did not change dramatically during the study period, which indicates that local population is still a major factor explaining cuisine variety. Finally, the finding that the hierarchical distribution of cuisines across places has decreased may indicate that IT has made it possible for (residents of) nearby places to make better use of complementarities in the supply of cuisines.

Substitution or complementarity – Table 2 presents the results of regression models based on eq. (2). It should be noted that the dependent variable is in logs, so a dR increase in the share of reviewed restaurants leads to a $(e^{dR*\delta} - 1) * 100\%$ increase in cuisine variety. Column (1) reports a coefficient of 0.254 for the share of reviewed restaurants, which means that a 0.25 increase (about a standard deviation) leads to a 6.5 percent increase in cuisine variety. This is comparable to the effect of increasing the amount of inhabitants in a settlement with 29 percent. We find positive effects of restaurants per capita, and the number of cuisines in 2000, so both competition and 'culinary culture' seem closely related with cuisine variety. The effect of hotel beds per capita is only significant at the 10% confidence level.

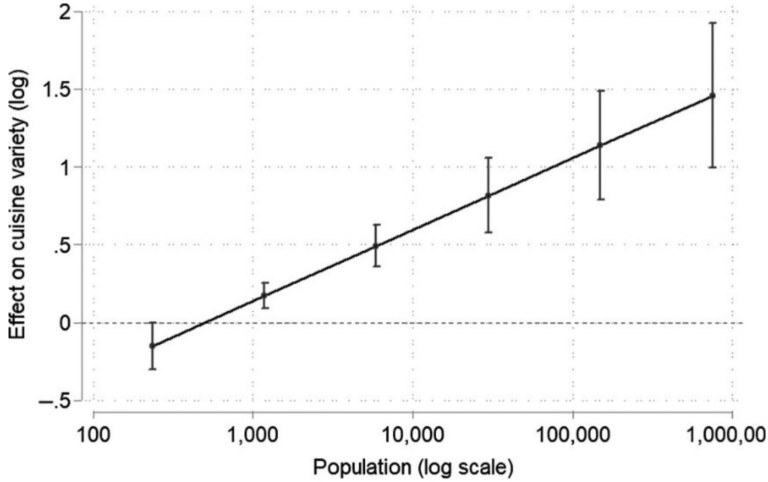
Results from column (2), in which we use population density instead of population and area separately, results in similar findings, although it seems that variables that distinguish urban from non-urban environments (such as the share on non-western foreigners and cuisine count in 2000) matter more in this regression. According to this model a 0.25 increase in the share of reviewed restaurants results in a 5.4 percent increase in cuisine variety. The effect of population density suggests that all else equal, places that are twice as dense have about 7% more cuisine variety.

In columns (3) and (4), we allow for a linear interaction between the share of reviewed restaurants and population and population density respectively. In these models population and population density are centred around their mean, for ease of interpretation. The resulting estimates are shown graphically in Figure 7, and they are unequivocal: the effect of the share of reviewed restaurants is higher in both larger and denser places. Column (3) provides a range of estimates from -0.15 in the smallest place, to 1.45 in the largest. In column (4) the estimates range from -0.32 in the least dense places to 0.81 in the densest. The emerging pattern suggests a complementary relationship between the cities and IT when it comes to variety in the restaurant market.

Table 2. OLS Regression models cuisine count and reviewed restaurants.

	(1)	(2)	(3)	(4)
Share of reviewed restaurants	0.254***	0.211***	0.404***	0.295***
	(0.0442)	(0.0469)	(0.0570)	(0.0576)
*Population (log)			0.200***	
			(0.0368)	
*Population density (log)				0.164***
				(0.0442)
Population (log)	0.225***		0.146***	
	(0.0163)		(0.0206)	
Area (log)	0.0385***		0.0365**	
	(0.0149)		(0.0144)	
Population density (log)		0.0722***		0.00508
		(0.0145)		(0.0227)
Restaurants/capita	27.16***	13.70***	28.48***	14.37***
	(3.182)	(3.044)	(3.084)	(3.074)
Number of cuisines 2000 (log)	0.576***	0.843***	0.551***	0.830***
	(0.0291)	(0.0225)	(0.0296)	(0.0231)
Share young	0.585	1.154*	0.441	0.970
	(0.575)	(0.602)	(0.567)	(0.605)
Share old	0.278	0.172	0.367	0.195
	(0.307)	(0.334)	(0.302)	(0.334)
Average household size	-0.187	-0.115	-0.161	-0.0974
	(0.114)	(0.122)	(0.111)	(0.120)
Average household income	0.0455	-0.00251	-0.0170	-0.0575
	(0.127)	(0.135)	(0.124)	(0.136)
Share western foreign	-0.0889	-0.205	-0.0864	-0.214
	(0.204)	(0.223)	(0.203)	(0.221)
Share non-western foreign	0.397	1.342***	0.235	1.273***
	(0.296)	(0.401)	(0.291)	(0.389)
Share high educated	0.00146	0.000569	0.000822	6.42e-05
	(0.00152)	(0.00163)	(0.00154)	(0.00162)
Hotel beds/capita	-0.253*	-0.144	-0.254**	-0.122
	(0.129)	(0.152)	(0.122)	(0.153)
Constant	-0.672	0.126	-0.0713	0.674
	(1.202)	(1.231)	(1.167)	(1.248)
Observations	1,280	1,280	1,280	1,280
R-squared	0.836	0.803	0.840	0.806

Dependent variable: Cuisine count (log) Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Population (log) and Population density (log) are mean-centred, for ease of interpretation of the main effects.



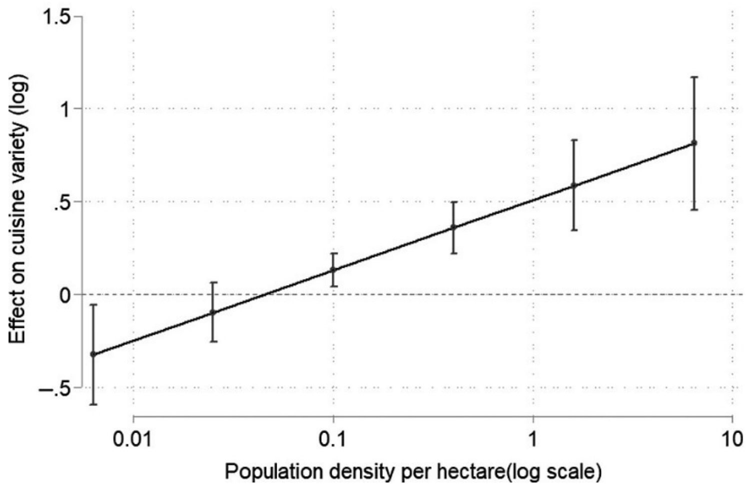


Figure 7. Marginal effects of the share of reviewed restaurants across population (l) and population density (r). The dots represent the point estimates, the vertical lines represent the 95% confidence intervals, and the dashed line represents 0.

Spillovers – Recall that Figures 3,4 and 5 provided exploratory evidence for the existence of agglomeration spillovers, as cuisine variety seems to grow most near larger places and that the relationship between size and cuisine variety was strong, but declining. Here, we estimate models in which we let the effect of the share of reviewed restaurants vary according to the distance to the nearest population centre. We include interactions with the share of reviewed restaurants and quintile categories of distance to the nearest centre, and we vary the minimum size of these population centres between 20,000, 50,000 and 100,000 inhabitants. The results of these models are shown graphically in Figure 8, and in Appendix 3. Panel (a) shows the decay pattern of the effect as it varies over distance from places with at least 20,000 inhabitants. A steep decay pattern is visible that plateaus after 8 kilometres. The effect of reviewed restaurants thus vitally depends on proximity to population centres. Panels (b) and (c) show the results of interactions with distance to centres of 50,000, 100,000 inhabitants or more, respectively. As we increase the threshold population that defines a population centre, the decay pattern becomes weaker. However with all definitions, the effect is consistently higher in the closest 2 quintiles than in the two farthest quintiles. These results favour the idea that part of the effect of online review intensity on cuisine variety is due to spatial spillovers.

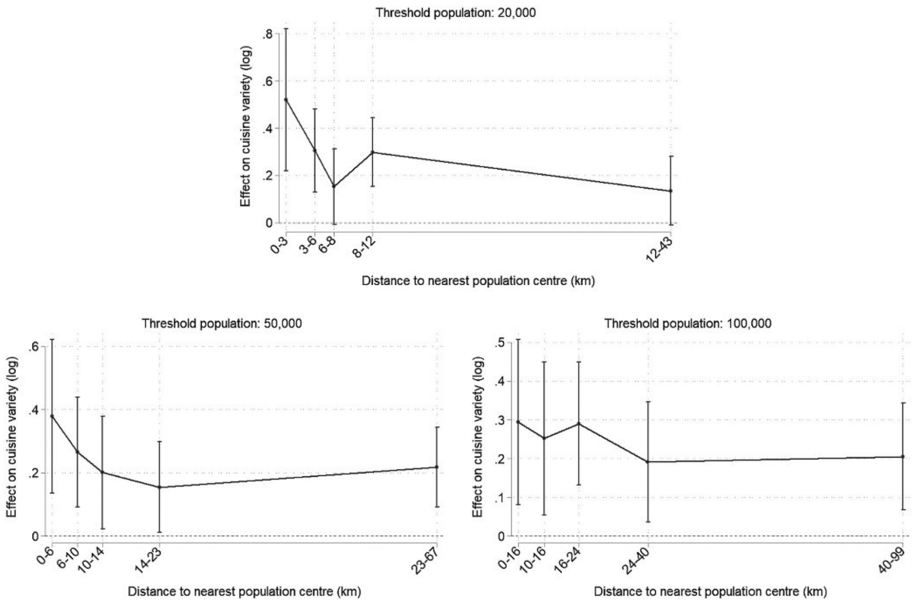


Figure 8. The effect of the share of reviewed restaurants and distance to population centres. The dots represent the point estimates, the vertical lines represent the 95% confidence intervals, and the dashed line represents 0.

Sensitivity analysis – In this subsection we submit our results to several sensitivity checks. Table 3 presents the sensitivity analysis for which we use Table 2, column (3) as a baseline, so we include an interaction between the share of reviewed restaurants and population.

One potential drawback of the analysis conducted in the previous subsection is the limited categorisation of cuisines in the dependent variable (20 categories). This may lead to measurement error, especially in the largest cities that have close to 20 cuisines. In column (1) we check whether our results are not dependent on observations from these largest cities. Therefore we limit the sample to places that do not exceed 100,000 inhabitants. The results are similar to the baseline estimate, with a comparable interaction between IT usage and population.

In column (2) we use an alternative measure of review website penetration: the number of reviews per capita. In the average city, a standard deviation (about 0.06) increase in the number of reviews per capita, results in a 4.7 percent increase in cuisine variety, which is comparable to the results in the previous subsection. The estimated interaction term is positive, in line with the other models.

In column (3) we check whether our results are sensitive to choosing another geographical scale of analysis. In the previous regression models cuisine variety was analysed at the level of places. Arguably, this is a lower bound of the geographical extent of restaurant markets. Therefore we analyse cuisine variety at the level of municipalities (covering 86 square kilometres on average). The summary statistics of these data are in Appendix 4. We find a positive effect of the share of reviewed restaurants on cuisine variety that is somewhat greater than the baseline estimate. An increase of 0.25 in the share of reviewed restaurants leads to a 12.5 percent increase in cuisine variety (vs. 10.6 percent in Table 2 column (3). At the municipality level the interaction term is insignificant. This result suggests that the interaction effect between size and information depends crucially on the spatial scale of analysis.

Several conclusions can be drawn from our sensitivity analysis. First, the notion that there is a positive and significant effect of online review intensity on cuisine variety is particularly robust to our sensitivity checks. Second, whether the effect of online information is greater in larger places depends crucially on the spatial scale at which the analysis is done. Finally, at the level of places, the positive interaction effect remains a robust finding.

Table 3. Sensitivity regressions population.

VARIABLES	(1) Pop<=100k	(2) Reviews/ capita	(3) Municipality
Share of reviewed restaurants	0.409***		0.473***
	(0.0552)		(0.130)
*Population (log)	0.228***		-0.273
	(0.0362)		(0.166)
Reviews per capita		0.767**	
		(0.374)	
*Population (log)		0.340**	
		(0.158)	
Population (log)	0.133***	0.210***	0.407***
	(0.0204)	(0.0162)	(0.0900)
Control variables	Yes	Yes	Yes
Observations	1,258	1,280	389
R-squared	0.824	0.832	0.792
F-statistic			
Endogeneity test (p)			

Dependent variable: Cuisine count (log) Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Population (log) and Population density (log) are mean-centred, for ease of interpretation of the main effects.

5.6 Conclusion

In the literature it is often shown that while IT acts as a substitute for many urban benefits, IT also complements the variety economies of agglomeration, because a greater extent of online local information stimulates demand for local goods, thereby promoting the entry of new varieties (Anenberg & Kung 2015; Sinai & Waldfogel 2004). We find that IT usage in the restaurant market (reviews) has a significant and positive effect on local cuisine variety, that increases with place size. This suggests a complementary relationship between agglomeration and IT, when it comes to local variety. However, the complementary relation between online information and variety decays with distance from large population centres, which indicates that IT allows places in the vicinity of these centres to ‘borrow size’ of their neighbours, and sustain a higher variety of cuisines. IT decreases spatial information frictions, and makes these places more accessible for residents of neighbouring areas. The complementarity between IT and local variety thus not only comes from more efficient information exchange, and increased local demand, but also from an increased geographical extent of markets.

We have three suggestions for further research into the link between IT and agglomeration. First of all, we identify a geographical effect of IT, referred to as ‘borrowed size’: agglomeration spillovers to smaller nearby places causing the increasing disconnection between size and function. However, conclusive evidence of this mechanism should ideally come from travel behaviour data that shows that (1) inhabitants of large towns are indeed travelling longer for consumption purposes, (2) this is particularly so for users of online platforms, and (3) this coincides with the penetration of online reviews. Second, our measure of reviewed restaurants only includes the entries on one website. For a long time Iens.nl was the market leading website, but there are rapid changes in the world of online culinary reviews, and information can now be found through various sources including Google Places, TripAdvisor, Facebook, and through personal blogs. While we feel that Iens.nl penetration may be a good proxy for the usage rate of all these sources, future measures of online information may have to take into account this increasingly scattered landscape. Third, variety increases may not benefit all consumers equally. Several scholars have proposed that high skilled individuals have a higher love of variety, and may thus benefit most (Adamson et al. 2004; Lee 2010). Future research may be aimed at finding out whether the benefits of IT have a skill-bias.

The results in this paper have multiple implications for strands of literature that do not focus on the effects of IT on agglomeration externalities of consumption. This paper is the first to estimate the relationship between cuisine variety and population size in a European context with many small settlements. The estimated elasticities of cuisine variety with respect to population are between 0.133 and 0.407. This is lower than the elasticity estimates reported by Schiff (2015) for the US (between 0.35 and 0.49), most likely due to the inclusion of smaller places. Relatedly, there seems to be a central place pattern to the distribution of cuisine varieties among Dutch municipalities. However, by adding a temporal dimension to this relationship, we show a declining relationship between size and variety, and a decreasing hierarchy among settlements. This may indicate that central place theory is becoming less useful in describing spatial patterns of consumption, and our results strengthen the idea that the focus should move more towards complementarity instead of hierarchy (Meijers 2007).

Finally, our results call for a renewed debate on the effects of IT on agglomeration. Substitution, complementarity, and spillover mechanisms are most likely present in the interaction between IT and other aspects of agglomeration, including face-to-face contact, knowledge spillovers, sharing of public goods and other amenities.

Appendix 1. Summary statistics cuisine data, 2000 and 2016.

	2000				2016			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	mean	sd	min	max	mean	sd	min	max
Number of cuisines	1.434	2.210	0	20	1.737	2.739	0	20
Cuisines								
Dutch-French	2.410	7.910	0	213	2.116	7.623	0	247
Italian	0.401	3.469	0	152	0.562	5.196	0	231
Greek	0.145	0.663	0	12	0.155	0.691	0	12
Turkish	0.0651	0.784	0	27	0.120	1.459	0	49
Southern-European	0.0289	0.452	0	20	0.140	1.244	0	49
Spanish-Portuguese	0.0314	0.488	0	20	0.0656	0.737	0	31
Eastern-European	0.0322	0.273	0	7	0.0151	0.159	0	5
Other European	0.0550	0.432	0	13	0.143	0.993	0	36
International	0.195	2.340	0	108	0.591	5.834	0	255
Chinese-Indonesian	0.896	4.676	0	160	0.730	3.532	0	110
Japanese	0.0269	0.413	0	17	0.142	1.466	0	61
Other Asian	0.0305	0.679	0	32	0.0660	1.106	0	49
Thai	0.0415	0.709	0	32	0.0599	0.916	0	39
Indian	0.0334	0.657	0	30	0.123	1.501	0	60
Argentinian	0.0350	0.514	0	23	0.0566	1.157	0	56
Mexican	0.0371	0.370	0	12	0.0350	0.402	0	17
American (US)	0.0375	0.254	0	7	0.0399	0.319	0	10
Surinamese	0.0200	0.405	0	14	0.0248	0.471	0	18
Other American	0.0159	0.236	0	10	0.0261	0.377	0	16
Other Foreign	0.0334	0.413	0	17	0.0525	0.718	0	29
N	2,456				2,456			
Population	6,501	24,747	0	645,650	6,924	27,688	0	747,725
N	2,439				2,452			

Appendix 2. Summary statistics regression sample

	(1)	(2)	(3)	(4)	(5)
VARIABLES	mean	sd	min	p50	max
Number of cuisines	3.176	3.147	1	2	20
Number of cuisines 2000	2.594	2.534	1	2	20
Number of restaurants	33.71	147.9	1	11	4,088
Number of restaurants/capita	0.00388	0.00493	0.000292	0.00263	0.0651
Hotel beds/capita	0.0285	0.0934	0	0.00485	1.195
Share of reviewed restaurants	0.417	0.238	0	0.427	1
Number of reviews/capita	0.0273	0.0629	0	0.00891	0.992
Population	12,437	37,477	235	4,050	747,725
Population density	0.580	0.787	0.00629	0.275	6.415
Area (hectare)	1,956	1,976	20.86	1,453	26,058
Average household size	2.337	0.199	1.670	2.330	3.334
Average household income	36,727	4,592	25,200	36,300	81,800
Share high educated	25.29	7.686	0	24.90	96.80
Share young	0.161	0.0275	0.0545	0.161	0.349
Share old	0.207	0.0436	0.0533	0.204	0.451
Share western foreign	0.0733	0.0479	0	0.0640	0.583
Share nonwestern foreign	0.0407	0.0460	0	0.0263	0.636
N	1,280				

Appendix 3. Regression models spatial spillovers

	(1)	(2)	(3)
Share of reviewed restaurants			
*Distance quintile I	0.519***	0.357**	0.301**
	(0.153)	(0.139)	(0.121)
*Distance quintile II	0.304***	0.316***	0.211**
	(0.0899)	(0.0874)	(0.106)
*Distance quintile III	0.152*	0.247***	0.271***
	(0.0821)	(0.0922)	(0.0816)
*Distance quintile IV	0.298***	0.147**	0.248***
	(0.0740)	(0.0749)	(0.0804)
*Distance quintile V	0.134*	0.233***	0.264***
	(0.0740)	(0.0658)	(0.0730)
Distance quintile II	0.0721	0.0126	0.0193
	(0.0821)	(0.0776)	(0.0747)
Distance quintile III	0.173**	0.0481	0.0282
	(0.0818)	(0.0765)	(0.0723)
Distance quintile IV	0.105	0.0625	0.0479
	(0.0806)	(0.0750)	(0.0716)
Distance quintile V	0.176**	0.104	0.0442
	(0.0813)	(0.0749)	(0.0724)
Population (log)	0.229***	0.231***	0.230***
	(0.0171)	(0.0169)	(0.0173)
Control variables	Yes	Yes	Yes
Threshold population	20,000	50,000	100,000
Observations	1,280	1,280	1,280
R-squared	0.837	0.837	0.836

Dependent variable: Cuisine count (log) Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix 4. Summary statistics municipality data.

VARIABLES	(1) mean	(2) sd	(3) min	(4) p50	(5) max
Number of cuisines	6.813	3.886	1	1	20
Number of cuisines 2000	5.218	3.416	0	0	20
Number of restaurants	102.2	244.4	2	2	3,857
Number of restaurants/capita	0.00241	0.00236	0.000612	0.000612	0.0250
Hotel beds/capita	0.0175	0.0622	0	0	0.942
Share of reviewed restaurants	0.494	0.148	0.0833	0.0833	1
Number of reviews/capita	0.0206	0.0353	7.81e-05	7.81e-05	0.493
Population	43,536	67,610	919	919	833,624
Area (hectare)	8,635	7,507	702.6	702.6	45,979
Average household size	2.293	0.183	1.600	1.600	3.300
Share high educated	25.97	8.191	0	0	96.80
Share young	0.165	0.0206	0.109	0.109	0.290
Share old	0.200	0.0313	0.0884	0.0884	0.310
Share western foreign	0.0806	0.0422	0.0144	0.0144	0.441
Share nonwestern foreign	0.0634	0.0549	0.0114	0.0114	0.376
	390				

References

Adamson, D. W., Clark, D. E. and Partridge, M. D. (2004) Do urban agglomeration effects and household amenities have a skill bias?, Journal of Regional Science, 44(2), 201–223. https://doi.org/10.1111/j.0022-4146.2004.00334.x.

Alonso, W. (1973) Urban Zero Population Growth, Daedalus, 102(4), 191–206. Available at: https://www.jstor.org/stable/20024174.

Anderson, M. L. and Magruder, J. (2011) Learning from the Cloud: Regression Discontinuity Estimates of the Effects of an Online Review Database, The Economic Journal, 122, 957–989. https://doi.org/10.1111/j.1468-0297.2012.02512.x.

Anenberg, E. and Kung, E. (2015) Information technology and product variety in the city : The case of food trucks, Journal of Urban Economics, 90, 60–78. https://doi.org/10.1016/j.jue.2015.09.006.

Angrist, J. D. and Pischke, J. S. (2008) Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press.

Berry, S. and Waldfogel, J. (2010) Product Quality and Market Size, The Journal of Industrial Economics. Blackwell Publishing Ltd, 58(1), 1–31. https://doi.org/10.1111/j.1467-6451.2010.00404.x.

Boterman, W. R. (2018) A Flavour of Class-Based Spatial Change: Geographies of Haute Cuisine in the Netherlands, Tijdschrift voor Economische en Sociale Geografie, 109(1), 161–170. https://doi.org/10.1111/tesg.12268.

Cairncross, F. (1997) The death of distance: how the communication revolution is changing our lives. Boston, Massachusettes: Harvard Business School Press.

Chen, Y. and Rosenthal, S. S. (2008) Local amenities and life-cycle migration: Do people move for jobs or fun?, Journal of Urban Economics 64(3), 519–537. https://doi.org/10.1016/j.jue.2008.05.005.

Christaller, W. (1966) Central places in southern Germany. Prentice-Hall.

Couture, V. (2013) Valuing the Consumption Benefits of Urban Density.

DeSoucey, M. (2010) Gastronationalism: Food Traditions and Authenticity Politics in the European Union, American Sociological Review, 75(3), 432–455. https://doi.org/10.1177/0003122410372226.

Friedman, T. L. (2005) The World Is Flat: A brief history of the twenty-first century. New York: Farrar, Straus and Giroux.

Gaspar, J. and Glaeser, E. L. (1998) Information Technology and the Future of Cities, Journal of Urban Economics, 43, 136–156. https://doi.org/10.1006/juec.1996.2031.

Glaeser, E. L., Kolko, J. and Saiz, A. (2001) Consumer city, Journal of Economic Geography, 1(1), 27–50. https://doi.org/10.1093/jeg/1.1.27.

Goldfarb, A. and Tucker, C. (2017) Digital Economics, NBER Working Paper Series, 23684. https://doi.org/10.3386/w23684.

Huang, P., Lurie, N. H. and Mitra, S. (2009) Searching for Experience on the Web : An Empirical Examination of Consumer Behavior for Search and experience goods, 73, 55–69.

Jensen, R. (2007) The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector, The Quarterly Journal of Economics, 122(3), 879–924. https://doi.org/10.1093/cje/bem005.

Koopman, E. R. (2002) Wezenlijk Nederlands belang en economische migratie. University of Amsterdam (UvA).
De Lange, T. (2016) Wezenlijk Nederlands Belang: De toelating tot Nederland van ondernemers van buiten de EU. Oisterwijk: Wolf Legal Publishers.
Leamer, E. E. and Storper, M. (2001) The economic geography of the internet age, Journal of International Business Studies, 32(4), 641–665. https://doi.org/10.1057/palgrave.jibs.84909988.
Lee, S. (2010) Ability sorting and consumer city, Journal of Urban Economics. Elsevier Inc., 68(1), 20–33. https://doi.org/10.1016/j.jue.2010.03.002.
Li, X., Hitt, L. M. and Zhang, Z. J. (2011) Product Reviews and Competition in Markets for Repeat Purchase Products, Journal of Management Information Systems, 27(4), 9–42. https://doi.org/10.2753/MIS0742-1222270401.
Meijers, E. J. (2007) From central place to network model: theory and evidence of a paradigm change, Tijdschrift voor Economische en Sociale Geografie, 98(2), 245–259. https://doi.org/10.1111/j.1467-9663.2007.00394.x.
Meijers, E. J. and Burger, M. J. (2017) Stretching the concept of “borrowed size”, Urban Studies, 54(1), 269–291. https://doi.org/10.1177/0042098015597642.
Mokhtarian, P. L. and Tal, G. (2013) Impacts of ICT on Travel Behavior: a tapestry of Relationships., in Rodrigue, J.-P., Notteboom, T., and Shaw, J. (eds) The SAGE handbook of transport studies. https://doi.org/10.4135/9781446247655.n14.
Mori, T., Nishikimi, K. and Smith, T. E. (2008) The number-average size rule: A new empirical relationship between industrial location and city size, Journal of Regional Science, 48(1), 165–211. https://doi.org/10.1111/j.1467-9787.2008.00550.x.
Phelps, N. A., Fallon, R. J. and Williams, C. L. (2001) Small firms, borrowed size and the urban-rural shift, Regional Studies, 35(7), 613–624. https://doi.org/10.1080/00343400120075885.
Rosenthal, S. S. and Strange, W. C. (2004) Evidence on the nature and sources of agglomeration economies, Handbook of Regional and Urban Economics - Vol 4, 4(04), 2120–2167. https://doi.org/10.1016/S0169-7218(04)07049-2.
Schiff, N. (2015) Cities and product variety: Evidence from restaurants, Journal of Economic Geography, 15(6), 1085–1123. https://doi.org/10.1093/jeg/lbu040.
Shaked, A. and Sutton, J. (1987) Product Differentiation and Industrial Structure, The Journal of Industrial Economics, 36(2), 131–146.
Sinai, T. and Waldfogel, J. (2004) Geography and the Internet: Is the Internet a Substitute or a Complement for Cities?, Journal of Urban Economics, 56(1), 1–24. https://doi.org/10.1016/j.jue.2004.04.001.
Terhorst, P. and Erkuş-Öztürk, H. (2018) Carriers of Change and Continuity: New Restaurants in Amsterdam, Tijdschrift voor Economische en Sociale Geografie, 109(1), 144–160. https://doi.org/10.1111/tesg.12266.
Van Voskuilen, M. (2013) Van table d’hôte tot haute cuisine. Universiteit van Amsterdam.
Waldfogel, J. (2008) The median voter and the median consumer: Local private goods and population composition, Journal of Urban Economics, 63(2), 567–582. https://doi.org/10.1016/j.jue.2007.04.002.

6. Information technology and the geographical extent of local product markets

Written by Duco de Vos

Abstract
This paper investigates to what extent ubiquitous online information on local products and services has altered the geographical extent of these local markets. To this end I use data from a panel survey specifically designed to analyse the impacts of ICT on travel behaviour. For each period I measure whether or not respondents are daily internet users. I estimate the effect of changes in this measure on travel distance for service and personal care purposes. By using the longitudinal aspect of the data I control for all time-invariant characteristics of respondents. I find that respondents that start to use the internet daily, increase travel distance for local products by 44 percent on average. This effect is mainly driven by residents of non-urban areas, and it is robust to a variety of different measures of IT usage. These results suggest that the rise of information technology is a credible explanation for at least part of the observed rise in travel distance for service trips between 2000 and 2015.

6.1 Introduction

Information about destinations is an essential part of generalized transportation costs, and thus influences travel behaviour (Klaassen 1988). With the rise of information technology, and the increasing ubiquity of information on local opportunities, one would expect that generalized transportation costs decrease, especially for local products and services that do not have a close online substitute (Mokhtarian and Tal, 2013). But while there have been many inquiries into the effects of information technology on travel behaviour, this research has mainly centred on trip purposes that do have an online counterpart, such as work, grocery shopping, and social interactions (Andreev et al., 2010). At the same time, local products without a clear offline counterpart make up an increasing part of the economic landscape of cities. [1] Therefore, transportation costs for these products will increasingly dictate the spatial organization of the economy, at least at the level of daily urban systems.

This paper investigates the effect of IT usage on travel behaviour for local products without a clear online counterpart. I analyse data from a mobility panel, and I use trips

[1] Such local products include non-tradable goods and services such as hairdressers, bootmakers, restaurants, cafes, dentists, etc.

related to services and/or personal care as a proxy.[2] Consistent with the hypothesis of decreased generalized transportation costs, Dutch travel survey data depicted in Figure 1 shows that between 2000 and 2015, the distribution of trips for service purposes has shifted towards more long distance trips, and average travel distance has increased significantly with about 17 percent (from 7.63 to 8.94 kilometres).[3] The aim of this paper is to investigate whether the rise of information technology can explain this trend.

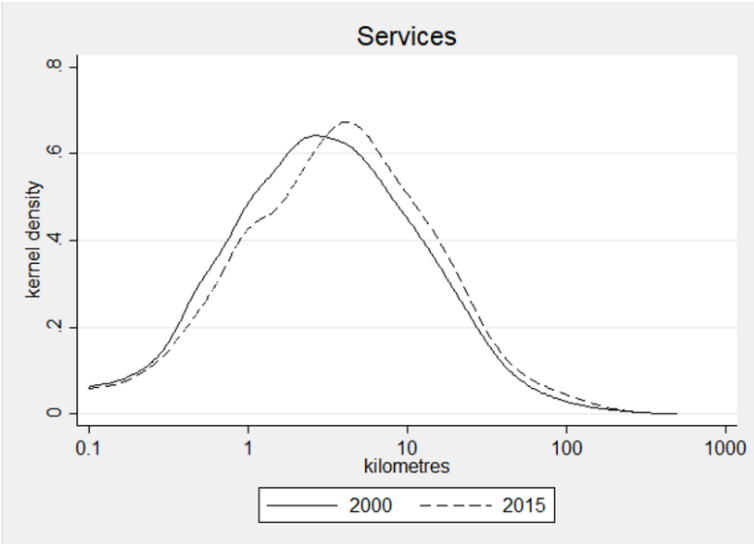


Figure 1: Travel distance for services in 2000 and 2015. Source: MON 2000 and OViN 2015

I investigate to what extent increased availability and usage of online information can explain the upward trend in travel distance for Service trips. To this end I use 4 waves from a novel panel data source from the Netherlands, the Mobility Panel Netherlands (MPN), that is specifically designed to assess the effects of information technology on travel behaviour. I measure whether or not respondents engage with online local information, and I use the panel dimension of the data to control for time invariant characteristics of respondents, and to measure the effect of changes in IT usage on changes in travel distance. I pay explicit attention to spatial differences in the relation between IT and travel behaviour, to assess potential effects on the spatial organization of the economy. Finally, I assess what part of the increase in travel distance for services can be explained by IT usage.

---

[2] The data set I use groups trips for service and personal care purposes together.
[3] This increase in travel distance is not a general phenomenon. A similar increase is present for commuting trips, but not for shopping or leisure trips.

This paper adds to the literature on the interaction between information technology and travel behaviour, that has traditionally focused on the effects of teleworking on commuting (e.g. Helminen and Ristimäki 2007, De Vos et al. 2018), the effects on social media usage on social trips (e.g. Delbosc and Mokhtarian 2018), and the effects of online shopping on offline shopping trips (e.g. Farag et al. 2006). In this literature trips for services that do not have a clear online counterpart have been ignored, and the current paper aims to fill this gap. As approaches that exploit time variation in IT usage are relatively rare, I also add to this literature by employing a fixed effects model. Furthermore, I add to a broader literature that deals with the interaction between information technology and urban agglomeration economies (Sinai and Waldfogel 2004, Anenberg and Kung 2015, De Vos et al. 2019). I do this by examining whether engaging with online information leads to a greater spatial reach of local markets.

The rest of this paper is organized as follows. In Section 2 I discuss the data used in the study, including some descriptive statistics. In Section 3 I discuss the empirical approach. Section 4 presents the results, and Section 5 concludes.

6.2 Data

I use data from the MobilteitsPanel Nederland (MPN), conducted by the Netherlands Institute for Transport Policy Analysis (KiM). This panel has been running since 2013, and I use the four available waves up until 2016.[4] The panel consists of around 2000 households with 4000 respondents, and it is based on annual three-day trip diaries between September and November, complemented with survey- and register data on individual characteristics (Hoogendoorn-Lanser et al., 2015). The data is available at the level of trips.

The MPN was especially designed to analyse the influence of social trends, including information technology, on travel behaviour. It includes very detailed questions about the intensity of internet usage and the usage of information technology for social networks, blogging, teleworking, online shopping, and medical services, among other purposes.[5] Similar to other trip diary surveys such as the Dutch mobility survey OViN, it includes detailed information on trip purposes. Importantly, and in contrast to some surveys such as the German *Mobilitätspanel*, the MPN distinguishes between shopping trips and trips for services and personal care. Trips are classified

[4] It is somewhat unfortunate that this survey only began after the rise of internet 2.0 and the big wave of internet adoption in the Netherlands (https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=NL). It should be noted that the effects estimated in this article are based on the behaviour of late adopters.

[5] Other categories are videochat, instant messaging, entertainment, dating, personal and work email, tele-meetings, learning, selling products, banking, and searching for information about products.

by respondents themselves, who can choose from (1) To and from work, (2) Business-related visit in work context, (3) Transport as a profession,(4) Dropping off, picking up people, (5) Delivering, picking up goods, (6) Following education study, courses, (7) Shopping, doing grocery shopping, (8) Visitation, (9) Touring, hiking, (10) Sports, hobby, (11) Other leisure time activities, (12) Services, personal care, and (13) Other purpose. The data used in the analysis consists of all 6,705 service and personal care trips in the MPN data, from 2013-2016. One important caveat of the data is that there is no information on the age of respondents. Given that the use of individual fixed-effects, and the short study period, my models will still control for a 'vintage' effect. I do not expect a strong bias related to the remaining effect of ageing.

I measure information technology usage based on the number of days of internet usage per week. I define daily internet users as respondents who use the internet 6 days or more. The data shows that these respondents are more likely to engage with local online content than respondents that make less use of the internet. They are for instance 3.8 times more likely to read or maintain blogs, 2.1 more likely to use social networks, and 1.7 times more likely to use medical services online. In a sensitivity analysis I experiment with alternative measures of IT usage, based on a lower threshold of weekly internet days (4), and engaging with local content through blogs.

Importantly, Figure 2 shows that for daily internet users, the distribution of travel distance for service trips includes more long-distance trips. Non-daily internet users travel 6.24 kilometres on average for service trips, whereas for daily internet users the average distance is 35 percent higher, with 8.44 kilometres. In the next section I investigate whether this relationship is also present when controlling for individual characteristics, and using time-variation in internet usage to estimate the effect on travel distance.

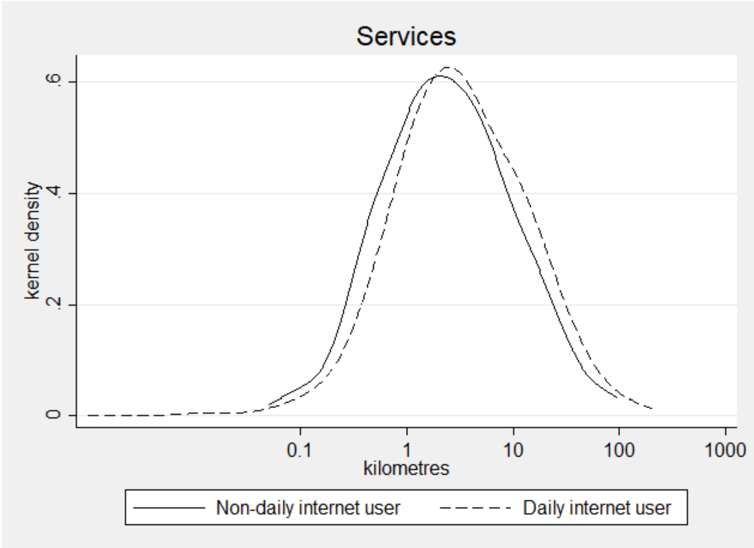


Figure 2: Travel distance for service trips for non-daily and daily internet users.

As mentioned in the introduction I pay attention to spatial variation in the effect of online information on travel distance for services. The MPN survey contains only one spatial indicator. This concerns the level of urbanization (*Stedelijkheid)*, and relates to the average address density in the municipality of residence. This measure corresponds to the urbanization measure of Statistics Netherlands (CBS), and discerns 5 types of urbanization. Because of my relatively small sample size, I recode this variable into a urban area dummy that equals 1 if the average address density is more than 1500 per square kilometre. This corresponds to the CBS measure of Strongly urban and Very strongly urban.

Table 1 shows the summary statistics of the data used in the analysis, consisting of all 6,705 service and personal care trips in the MPN data from 2013 until 2016. Average travel distance is 8.2 kilometres, with considerable variation. Internet usage is rather high, and on average respondents use the internet more than 6 days a week. 91 percent of respondents fit my definition of daily inter user, and 96 percent of respondents use the internet 4 days a week or more. I also use a stricter proxy for engaging with local online content, based on blog usage, and 22 percent of respondents are defined as such. The survey is well balanced in spatial terms, and about 49 percent of respondents resides in urban areas. Further notable summary statistics are that more than half of trips is done by car, and trips are mostly done during the day.

Table 1: Summary statistics service trip data

VARIABLES	(1) mean	(2) sd	(3) min	(4) max
Travel distance (km)	8.246	16.43	0.00100	206
Number of internet days	6.601	1.180	0	7
Daily internet user	0.911	0.285	0	1
Internet usage >= 4 days	0.959	0.197	0	1
Blog reader	0.224	0.417	0	1
Urban area	0.487	0.500	0	1
Female	0.592	0.492	0	1
Multiperson household	0.750	0.433	0	1
Children at home	0.386	0.487	0	1
Education: Basic	0.288	0.453	0	1
Education: Medium	0.390	0.488	0	1
Education: Higher	0.322	0.467	0	1
Education: Unknown	0.000597	0.0244	0	1
Employment: < 12h/w	0.132	0.339	0	1
Employment: >= 12h/w	0.447	0.497	0	1
Employment: None	0.421	0.494	0	1
Income: 0 - 12,500	0.0480	0.214	0	1
Income: 12,500 - 26,200	0.146	0.353	0	1
Income: 26,200 - 38,800	0.212	0.409	0	1
Income: 38,800 - 65,000	0.295	0.456	0	1
Income: 65,000 - 77,500	0.0780	0.268	0	1
Income: > 77,500	0.0864	0.281	0	1
Income: Unknown	0.134	0.341	0	1
Mode: Personal motorized	0.524	0.499	0	1
Mode: Public transport	0.0901	0.286	0	1
Mode: Light motorized	0.0192	0.137	0	1
Mode: Bicycle	0.255	0.436	0	1
Mode: Walking	0.112	0.316	0	1
Time of travel: Morning peak	0.111	0.314	0	1
Time of travel: Day	0.684	0.465	0	1
Time of travel: EveningPeak	0.152	0.359	0	1
Time of travel: Evening	0.0458	0.209	0	1
Time of travel: Night	0.00656	0.0807	0	1
Weekday	0.927	0.260	0	1
N	6,705			

6.3 Results

Empirical approach - I estimate the effects of changes in information technology usage on travel distance using fixed effects models. Observations corresponds to trips, and fixed effects are included for individuals and time. The model is as follows:

$$D_{rit} = \alpha + \beta IT_{it} + \gamma X_{rit} + \eta_i + \theta_t + \epsilon_{rit} \quad (1)$$

where D_{rit} is the natural logarithm of distance travelled during trip r by individual i in year t. IT_{it} denotes IT usage. X_{rit} are both trip level and individual level control variables. η_i and θ_t are individual and year fixed effects respectively. And ϵ_{rit} is the error term.

By including individual fixed effects, essentially including a dummy variable for every individual, identification of the effects is based on variation 'within' individuals over time. This model thus controls for all observed and unobserved characteristics of individuals that do not vary over time. The key identifying assumption is that variation in IT usage is not related to time varying unobserved variables that are correlated with travel distance. To minimize the extent of bias due to time varying variables, we include control variables for changes in household structure (dummies for single person household and children at home), education level (Basic, intermediate, and higher education), employment situation (unemployed, employed < 12h/w, employed >12h/w), and income (6 categories). We also control for trip-level variables, including mode of travel (walking, cycling, public transport, light motorized or personal motorized), a weekday dummy, and time of day dummies.

Regression results - Table 2 presents regression models in which I regress the logarithm of travel distance on a dummy variable for Daily internet usage, that equals 1 if a respondent uses the internet 6 days a week or more on average. Since the dependent variable is in logs, the coefficients of dummy variables should be interpreted as $\left(e^{\beta}-1\right) * 100\%$. Column (1) shows the results of a regression model with year dummies, and without further control variables. According to this model Daily internet users have 32 percent[6] longer travel distances for service trips on average. This is similar to the bivariate result of 35 percent from the previous section.

In column (2) I add individual fixed effects that account for all time invariant characteristics of individuals. According to this model, respondents that start to use the internet daily increase their travel distance for service trips by 65 percent on average. This model suggests that when estimating the relation between IT and travel distance for service trips, failing to account for time invariant characteristics biases results downwards. In column (3) we add controls for time varying characteristics of respondents, and trip level controls. Here the coefficient of 0.364 suggests that becoming a Daily internet user leads to a 44 percent longer travel distance for service tips on average. Whereas including individual fixed effects solved a downward bias, it seems that controlling for time varying variables solves an upward bias. Furthermore, the inclusion of control variables significantly improves the explanatory power of the model, evidenced by a higher overall and within r-squared.

In column (4) I test whether the effect of internet usage on travel distance for service trips differs between urban and non-urban residents. Urban residents are defines as

[6] $(e^{0.279}-1)$*100%

those who live in municipalities with an address density higher than 1500 per square kilometre, and I separate indicators for non-urban daily internet users and urban daily internet users. The results show that the positive relation between internet usage and travel distance is primarily driven by non-urban residents. In this group, those who start using the internet daily increase travel distance by 47 percent on average. Although the point estimate for urban residents is similar, the standard error is higher and the estimate is only significant at the 10 percent confidence level.

Taken together I find a strong and significant positive relationship between Daily internet usage and travel distance for service trips. This result is consistent with the hypothesis that the advent of information technology is a key driver of the increased travel distance for service trips. I find that respondents who start to use internet daily increase their travel distance by around 44 percent, and this is mainly driven by residents of non-urban areas. A potential explanation for this is that cities have an edge in both online and offline information provision. This is consistent with the finding of Sinai and Waldfogel (2004) that larger cities have disproportionately more local online content.

Table 2: Regression models

	(1)	(2)	(3)	(4)
Daily internet user	0.279***	0.503***	0.364***	
	(0.0597)	(0.132)	(0.109)	
Daily internet user*Non-urban area				0.386***
				(0.145)
Daily internet user*Urban area				0.325**
				(0.148)
Constant	0.881***			
	(0.0568)			
Year FE	Yes	Yes	Yes	Yes
Individual FE	No	Yes	Yes	Yes
Control variables	No	No	Yes	Yes
Observations	6,705	6,705	6,705	6,705
R-squared	0.003	0.655	0.772	0.772
Within r-squared		0.0087	0.3437	0.3445

Notes: Dependent variable: Travel distance (log). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Alternative measures of IT usage - In Table 3 I show the sensitivity of my results to changing the measure of IT usage. In column (1) I use a lower threshold of internet usage, based on 4 days a week or more. This model suggests 70 percent higher travel distances for respondents that pass the threshold. Column (2) again shows that this effect is mainly driven by non-urban residents. The difference between the effects in non-urban and urban environments is larger than in the model from Table (2) column (4).

In column (3) I include a linear measure of internet days per week, and I find that every increase with 1 day corresponds to a 10 percent increase in travel distance. According to column (4), and consistent with the previous results, this effect is driven by non-urban residents. The effect I find for non-urban residents is about 60 percent greater than the effect for urban residents, and the effect is not statistically significant for the latter group.

In column (5) I employ a different proxy for engaging with local online content, based on whether respondents read or maintain blogs. More than news or general interest websites, blogs are usually personal, and often have a specific local delineation (e.g. travel blogs, food blogs, shopping blogs, etc.) (Kurashima et al. 2006). This model shows that respondents who start to use blogs increase their travel distance by 22 percent. This is a more modest estimate compared to the estimates based on internet usage. Consistent with the earlier findings, column (6) shows that the effect is driven by non-urban residents.

Table 3: Sensitivity regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Internet days >=4	0.534***					
	(0.200)					
Internet days >=4*Non-urban		0.738***				
		(0.261)				
Internet days >=4*Urban		0.415*				
		(0.233)				
# internet days			0.0931***			
			(0.0344)			
# internet days*Non-urban				0.119***		
				(0.0446)		
# internet days*Urban				0.0741*		
				(0.0414)		
Blog reader					0.202**	
					(0.0917)	
Blog reader*Non-urban						0.254**
						(0.106)
Blog reader*Urban						0.153
						(0.109)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,705	6,705	6,705	6,705	6,705	6,705
R-squared	0.772	0.772	0.772	0.772	0.772	0.772
Within r-squared	0.3431	0.3442	0.3431	0.3441	0.3431	0.3440

Notes: Dependent variable: Travel distance (log). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The sensitivity analysis shows that there is a robust positive relationship between IT usage and travel distance for service purposes. Depending on the definition of IT usage to engage with local online information, the estimates range from 22 to 70 percent longer distance for IT users on average. The result from Table 2 column (3) points at a 44 percent increase, which holds the middle ground between these extremes. Although the effect is driven by non-urban respondents, the point estimates for urban residents are positive across all estimations, and occasionally significant at the 10 percent confidence level. One possible reason why results are less clear for urban residents may be that variation in daily internet usage is lower in this group.

6.4 Conclusion

This paper started with the observation that travel distances for service trips in the Netherlands have increased considerably between 2000 and 2015 with about 17 percent. I have shown with panel data from 2013 until 2016 that changes in IT usage are strongly related to changes in travel distance for service trips. My preferred estimate suggests that people who start to use the internet daily increase their travel distance for these trips by 44 percent on average. This effect is mainly driven by non-urban residents, but a somewhat weaker positive effect for urban residents is not ruled out. These results are robust to different measures of IT usage including using the internet 4 days a week or more, a linear measure of internet days per week, and engaging with blogs. These results thus suggest that the rise of information technology is a credible explanation for the observed rise in travel distance for service purposes over the last decades.

It should be noted that information technology is likely not the only explanation for the rise in travel distance for services. For instance, health-services such as hospitals have increasingly concentrated over time (Meijers 2007). The centralization of such public provisions have likely played a role in changes in travel behaviour for services. Importantly, the decreasing accessibility of services may have prompted people to use information technology and the internet as a substitute for such services. This points at a potential reverse causality whereby increasing travel distances for services have fostered internet usage, rather than the other way around. Here it should be noted that my estimates show an increase in travel distance that *coincides* with starting to engage with local online content. To the extent that there remains reverse causality, my results still show that information technology enables people to make use of local goods and services over a longer distance.

By showing a complementary relationship between IT and travel for service purposes, this paper adds to a broader literature on the interaction between IT and travel behaviour (Andreev et al. 2010, Ben-Elia and Zhen 2018). It is important to note that in this literature complementarity can either arise through increased travel distance and an increased number of trips, whereas I only estimated effects on the former. Future research should uncover whether a similar complementary relation exists for the number of trips.

This paper also adds to a literature concerned with the relation between IT and urbanization economies. The results from this paper partly confirm the idea that information technology has extended the range of local product markets, that may have resulted in a wider availability of scale economies associated with larger markets, including the availability of a variety of local products (Anenberg and Kung 2015, De Vos and Meijers 2019). The extent to which this goes for local products other than services and personal care (for instance musea, restaurants, local newspapers, street markets, etc.) is a fruitful topic for further research.

Policymakers and planners should be aware that information technology can increase the geographical extent of local markets. For some products this may mean that local provision can be substituted for provision from a distance. Planners should then focus on making these distant services accessible for all.

References

Andreev, P., Salomon, I., Pliskin, N., 2010. Review: State of teleactivities. Transp. Res. Part C Emerg. Technol. 18, 3–20. https://doi.org/10.1016/j.trc.2009.04.017

Anenberg, E., Kung, E., 2015. Information technology and product variety in the city : The case of food trucks. J. Urban Econ. 90, 60–78. https://doi.org/10.1016/j.jue.2015.09.006

De Vos, D., Meijers, E., 2019. Information Technology and Local Product Variety: Substitution, Complementarity and Spillovers. Tijdschr. voor Econ. en Soc. Geogr. 1–21. https://doi.org/10.1111/tesg.12365

De Vos, D., Meijers, E., van Ham, M., 2018. Working from home and the willingness to accept a longer commute. Ann. Reg. Sci. 61, 375–398. https://doi.org/10.1007/s00168-018-0873-6

Delbosc, A., Mokhtarian, P., 2018. Face to Facebook: The relationship between social media and social travel. Transp. Policy 68, 20–27. https://doi.org/10.1016/j.tranpol.2018.04.005

Farag, S., Krizek, K.J., Dijst, M., Farag, S., Krizek, K.J., Dijst, M., 2007. E - Shopping and its Relationship with In - store Shopping : Empirical Evidence from the Netherlands and the USA E-Shopping and its Relationship with In-store Shopping : Empirical Evidence from the Netherlands and the USA 1647. https://doi.org/10.1080/01441640500158496

Helminen, V., Ristimäki, M., 2007. Relationships between commuting distance, frequency and telework in Finland. J. Transp. Geogr. 15, 331–342.

Hoogendoorn-lanser, S., Schaap, N.T.W., Oldekalter, M., 2015. The Netherlands Mobility Panel: An innovative design approach for web-based longitudinal travel data collection. Transp. Res. Procedia 11, 311–329. https://doi.org/10.1016/j.trpro.2015.12.027

Klaassen, L.H., 1988. Exercises in spatial thinking: case studies of regional development issues. Gower Publishing Company, Aldershot.

Kurashima, T., Tezuka, T., Tanaka, K., 2006. Mining and Visualizing Local Experiences from Blog Entries, in: International Conference on Database and Expert Systems Applications. Springer, Berlin, Heidelberg, 213–222.
Mokhtarian, P.L., Tal, G., 2013. Impacts of ICT on Travel Behavior: a tapestry of Relationships., in: Rodrigue, J.-P., Notteboom, T., Shaw, J. (Eds.), The SAGE Handbook of Transport Studies. https://doi.org/10.4135/9781446247655.n14
Sinai, T., Waldfogel, J., 2004. Geography and the Internet: Is the Internet a Substitute or a Complement for Cities? J. Urban Econ. 56, 1–24. https://doi.org/10.1016/j.jue.2004.04.001

7. Discussion

7.1 Introduction

It is increasingly recognized that urbanization economies – the benefits of living in cities – can be generated by proximity to large cities (OECD 2015). Several scholars have put forward that places near other large cities are increasingly able to 'borrow size' of their neighbours to generate these economies (Dijkstra et al. 2013, Burger and Meijers 2016), due to improvements in physical and digital infrastructure. Indeed, there is plenty of evidence that increasing the *effective density* of regions by improving physical transportation infrastructure leads to higher levels of urbanization economies (Graham 2019). For improvements in digital infrastructure however, such evidence is missing. In this thesis I have attempted to fill this gap, and contribute to the discussion of whether information technology enables places in proximity of large cities to 'borrow' urbanization economies?

To understand the relation between IT and borrowed size it is important to have a plausible theoretical mechanism. In the introduction of this thesis I have put forward such a theoretical link, that is based on the relation between ubiquitous online information and travel behaviour. In short, I expect that in some cases IT may complement longer distance travel for jobs and local products, which means that in these markets urban scale economies are generated and enjoyed across a greater geographical scale. Based on this theoretical link, I devised two research questions.

1. To what extent does information technology increase the geographical extent of local labour and product markets?

2. To what extent has the advent of information technology led to better local (labour or product) market outcomes in places in proximity of large cities?

In the preceding Chapters I have answered these research questions by providing micro-level evidence on the link between information technology and travel behaviour, and macro-level evidence on the relationship between information technology and urbanization economies. In this concluding Chapter I synthesise the results. In Section 7.2 I summarize the results of the Chapters, and I discuss how and to what extent the research questions are answered. In Section 7.3 I point out a number of the limitations of my work that may inspire future research. In section 7.4 I discuss the policy implications of the findings. Section 7.5 provides a research agenda that points out some open and interesting questions concerning relationship between information technology and urbanization economies. Section 7.6 concludes.

7.2 Summary of results

This thesis was divided into two parts, whereby the Chapters in Part 1 answered the research questions for local labour markets, and the Chapters in Part 2 answered the research questions for local product markets. Table 1 shows this structure schematically. In this section I first summarize the results for each part, after which I discuss whether and to what extent the results contribute to the aim of this thesis.

Table 1: Thesis structure

	Part 1 *Local labour markets*	**Part 2** *Local product markets*
RQ1: IT and travel behaviour	Chapters 2 and 3	Chapter 6
RQ2: IT and urbanization economies	Chapter 4	Chapter 5
Synthesis	Chapter 7	

Local labour markets - In Part 1, Chapters 2 and 3 aimed to answer whether information technology has extended the geographical range of local labour markets (research question 1) by analysing the relationship between home-based teleworking and commuting distance. There already is a considerable literature that addresses the question of whether telework leads to more, or to less travel, but evidence on the relationship between telework and commuting distance has been mixed (De Graaff 2004). Furthermore, the majority of studies have suffered from issues with the definition of teleworking, external validity, theoretical substantiation of results, and empirical rigour (Andreev et al. 2010, Gubins et al. 2019). To come closer to an answer for research question 1 for local labour markets, it was necessary to address these issues. My aim was to overcome the most pressing issues by using longitudinal data from representative panel surveys.

Chapter 2 named *Working from home and the willingness to accept a longer commute* addressed the empirical issues that have plagued research on the relationship between teleworking and the length of commutes. These issues relate to (1) omitted variables, (2) reverse causality, and (3) selection based on unobserved preferences. In the literature there have been some attempts to tackle these issues by using Instrumental Variable approaches (e.g. Jiang 2008, Zhu 2012), but as Gubins et al (2019) note, it is hard to find a proper Instrumental Variable for IT usage, uncorrelated with commuting distance. In this Chapter I employed a different approach, using a panel data set that spans more than a decade, and estimating the effect of telework on commute lengths with fixed effects models. This approach primarily addresses the issue of selection based on preferences, but also makes headway in addressing omitted variables and reverse causality by including detailed job characteristics in the estimation, and by analysing the effect of changes in telework on changes in commuting.

The first result of Chapter 2 is that it is important to control for unobserved preferences. Using an existing approach to measure commuting preferences by the Marginal Cost of Commuting (Van Ommeren and Fosgerau 2009) I find that teleworkers have a stronger dislike for commuting, compared to non-teleworkers. But whereas this should theoretically bias observational results downwards, I find that adding fixed effects leads to a lower estimate. This suggests that another type of selection, based on residential preferences, may induce a stronger bias. The second and main result of this Chapter is that information technology has indeed increased the geographical scale of labour markets, as teleworking has a significant positive effect on commute lengths. Working from home allows people to accept 5% longer commuting times on average, and every 8 weekly hours working from home are associated with 3.5% longer commuting times.

The paper in Chapter 3 named *Working from home and commuting: Heterogeneity over time, space, and occupations* delved deeper into the relationship between telework and commuting. Specifically this Chapter reproduced the research of Chapter 3 with a different, more recent data set, and brought nuance to the results by showing the heterogeneity of the effect of teleworking. The key finding of this Chapter is that I can reproduce the positive relation between working from home and commuting found in Chapter 2 with another data set. I found a stronger effect than in Chapter 2: respondents who start teleworking increase their commutes by 12 percent. It should however be noted that the definition of teleworkers slightly differs between the two chapters. Across the study period of 2008-2018 we find that the effect remained remarkably similar.

Chapter 3 also made clear that there is not a one-size-fits-all relationship between telework and commuting. First, there is a consistent positive effect of telework on commuting times, but this effect is especially high in Moderately urban areas with address densities between 1000 and 1500 addresses per square kilometre. A visual inspection of the geography of these areas shows they are largely situated around and in between the larger cities in the main urban regions of the Netherlands. Between occupations there are starker differences in the relationship. Perhaps unsurprisingly, the effect of telework on commuting is driven by occupations where telework has grown during the study period. This involves occupations that make use of intellectual rather than manual skills.

The results of Chapters 2 and 3 shed light on whether information technology has increased the geographical scale of local labour markets. The results unequivocally show that commencing teleworking is associated with a greater geographical distance between home and workplace, as measured by commuting time. The information technologies that allow teleworking thus make it easier to sustain longer commuting distances. A big question in the literature is whether these long commutes would have taken place regardless of the possibility of teleworking. By making use of longitudinal data, and estimating the effect of changes in teleworking on changes in commuting distance, the results from Chapters 2 and 3 show that it is not only long-distance commuters that choose to telework, but rather commencing teleworking is associated with accepting longer commutes. The increase in the geographical scale of labour markets for teleworkers is substantial. Assuming that changes in commuting time are proportional to changes in commuting distance, a 12 percent increase in commuting time leads to a 25 percent increase in the area of the local labour market. What has become clear is that not all professions are conducive to teleworking, and information technology has thus only affected the geographical scale of markets for jobs conducive to teleworking.

Chapter 4 dealt with the question of whether improvements in digital infrastructure have also had measurable effects on labour market outcomes. This Chapter aimed to answer research question 2, whether IT has led to better labour market outcomes in places in proximity of large cities. It consists of a paper called *Does broadband internet allow cities to 'borrow size'? Evidence from the Swedish labour market*. This Chapter started from a discussion of recent literature that suggests that recent growth dynamics of European cities – whereby the largest cities did not necessarily have the highest economic growth – may be explained by increased possibilities for 'borrowing size'. In other words, that proximity to large cities may have become a better substitute for agglomeration in recent years. In this literature it is often noted that this may have to do with the increased availability of broadband internet (Dijkstra et al. 2013, Hesse 2016). I made use of a unique data set from Sweden that traces the geographical availability of different internet technologies at the level of 250m grid cells, between 2007 and 2015. These data were then connected to the Swedish population register to estimate the effects of broadband availability on labour market outcomes.

Chapter 4 again used fixed-effects models to estimate the effects of interest. Specifically, it examined whether changes in internet availability are related to changes in commuting, and to changes in the propensity to be employed. This approach thus tests whether the relationships found at the micro level, between teleworking and commuting, are also present at the macro level, between the infrastructure that enables telework and the length of commutes. By estimating the effect on local employment rates, I then investigated whether this also leads to better outcomes. Importantly, I distinguished between the effect of broadband at the residential location, and broadband penetration at the place level. By conditioning on the latter, the estimates pertaining to residential broadband should better approach the causal effect.

Whereas Chapters 2 and 3 showed strong evidence that information technology changes spatial behaviour, Chapter 4 at best provides suggestive evidence that broadband internet has changed urbanization economies (i.e. matching) in local labour markets. Outside of cities I find no relationship between residential broadband and the prospects of employment. There is however a weak but significant relation between place-level broadband penetration and employment, in places between 20 and 50 kilometres from cities. While this result suggests that broadband penetration makes places in proximity of cities more attractive for firms and (employed) people to locate, the finding that it is not being connected, but rather local connectivity that counts poses a puzzle.

Local product markets - Part 2 of this thesis dealt with the relation between IT and borrowed urbanization economies in local product markets, as opposed to local labour markets. Specifically, here I examined whether information technology alters the geographical extent of local product markets, and leads to a wider distribution of the consumption benefits of agglomeration. I built on a small literature that assesses the impact of IT and internet usage on consumption side agglomeration economies (Sinai and Waldfogel 2004, Anenberg and Kung 2015). One of the key findings of this literature is that the availability of online information may increase the efficiency of demand for local goods. The line of reasoning is then that in cities, where it is more difficult to navigate between numerous opportunities, this has stronger effects on outcomes in local markets such as product quality and product variety. According to this literature, IT thus complements the consumption benefits of agglomeration.

The paper in Chapter 5 titled *Information technology and local product variety: Substitution, complementarity and spillovers* used cuisine variety in local restaurant markets as indicator of consumption side agglomeration economies, following a longer tradition (Glaeser et al. 2001, Berry and Waldfogel 2010, Schiff 2015). I measured the usage and penetration of internet in local restaurant markets by the share of restaurants reviewed on Iens.nl, which at the time was the most popular culinary review website in the Netherlands. I controlled for unobserved culinary culture using historic levels of cuisine variety, and I found a strong and significant relationship between the share of reviewed restaurants and cuisine variety in Dutch places. A standard deviation increase (about 0.25) in the share of reviewed restaurants is associated with an increase of cuisine variety with more than 5 percent. The results confirm the finding of earlier studies that the effects of IT on product variety increase with city size.

When mapping changes in cuisine variety across places in the Netherlands between 2000 and 2016, it became clear that it has become a lot more common to have considerable cuisine variety, and variety increases have predominantly taken place between and around places where variety was already high in 2000. The results of Chapter 5 may provide an explanation for this pattern, as it is shown that the relationship between IT and variety decays with distance from large population centres. This suggests that information technology allows places near larger centres to better capitalize on the size of their neighbours, and sustain higher levels of cuisine variety. Information technology may increase the willingness to travel for restaurants, and make these places more accessible for residents from neighbouring places.

Chapter 6 titled *Information technology and the geographical extent of local product markets* aimed to answer sub-question 2 for local product markets. In this Chapter I first identified a lacuna in current research about the effect of information technology on travel behaviour: The literature has primarily dealt with trip purposes with a clear online counterpart, and it has ignored trips for purposes without a clear online pendant. At the same time, these activities make up an increasingly important share of the economic landscape of cities. I used trips related to services and personal care as a proxy for such activities, and I first showed that there has been a significant increase in travel distance for these purposes between 2000 and 2015, with about 17 percent. I then proceeded to test whether the rise of information technology can explain this trend.

Making use of a Dutch panel survey specifically designed to assess the impact of information technology on travel behaviour, I used fixed-effects models to estimate the effect of changes in engaging with local online content on changes in travel for services and personal care purposes. With slight variability depending on the measure of engaging with online content, I found a 40 percent increase in travel distance for service trips. The effect seems mainly driven by non-urban residents. I concluded that the rise of information technology is a credible explanation for the observed increase in travel distance for service trips over the last decades.

Conclusions – Chapters 2-7 have offered important empirical evidence about the extent to which information technology allows places to borrow urbanization economies from nearby cities. The results confirmed the hypothesized mechanism behind this link, that usage of information technology leads to longer travel distances for work and local consumption purposes. These results are quite dramatic: teleworkers increase their commutes by 12 percent on average, and engaging with local online content is associated with 40 percent longer trips for service purposes. Information technology has thus increased the geographical scale of distinct local labour- and local product markets (research question 1), which allows generating and enjoying urbanization economies across greater geographical areas.

Concerning the actual effects of information technology usage and availability on measures of urbanization economies, results are less dramatic but provide some evidence that borrowing urbanization economies is facilitated by digital technologies. Chapter 4 showed that in places nearby large cities, the propensity to be employed is correlated with changes in place-level broadband penetration. Chapter 5 showed that controlling for historic levels of cuisine variety, places with more available online information on the restaurant market have higher levels of cuisine variety. This relationship decays with distance from large cities, which suggests a positive interaction between proximity and information technology, that leads to more local product variety. Taken together, the results provide suggestive evidence that information technology has led to better outcomes in local labour- and product markets (research question 2).

While I found evidence that information technology enables borrowing urbanization economies, it is important to note that the results show information technology also positively affects urbanization economies in cities. In Chapter 4 it is shown that broadband availability at the residence only affects employment in cities. And in Chapter 6 the interaction between information technology and product variety is strongest in cities. It remains difficult to assess to what extent this affects the relative changes in welfare across space. Firstly, because small increases in prospects of employment or variety may be valued more in small places with initially lower levels. Secondly, because diseconomies of urbanization, such as pollution and congestion, may be more prevalent in cities compared to small hinterland places. How information technology interacts with these negative sides of urbanization may be a fruitful topic for further research.

The aim of this thesis was to contribute to the discussion of whether information technology enables places in proximity of large cities to 'borrow' urbanization economies*?* The results pertaining to research question 1 suggest that there may be significant effects, because information technology has important impacts on travel behaviour in local markets. I only found suggestive evidence concerning the link between information technology and market outcomes. This evidence is indicative of a small but positive and significant effect of information technology on borrowed urbanization economies. Considering the strong relation between information technology and travel behaviour, it should not be ruled out that in the long run, information technology may have stronger effects on the geographical scale of urbanization economies.

7.3 Limitations

As already elaborated upon at the end of each Chapter, the research done in this thesis is subject to several limitations, that may inspire future research. In this subsection I reflect on what I consider the main conceptual and empirical limitations of this thesis. I propose strategies to overcome these limitations, and come closer to definitive answers on the relationship between information technology and borrowed urbanization economies.

The first limitation follows from a somewhat narrow view of urban scale economies. Across this thesis I have considered matching on the labour market, and product variety in local product markets as main outcomes that signify urbanization economies, and I have not considered other often used indicators such as productivity and innovation. This choice can be partially justified by noting that I have taken the perspective of residents. From this perspective, the prospects of employment and local consumption opportunities seem more relevant (for location behaviour, policy) than local levels of productivity and innovation. However, the sources that suggest a link between information technology and borrowed urban scale economies do not necessarily take on this perspective. And some of the evidence that inspired such hypotheses is explicitly based on productivity data (e.g. Camagni et al. 2014). The correlations between local connectivity and employment rates in places near large cities warrant further investigations into the relation between information technology and borrowed urbanization economies from the perspective of firms rather than residents.

The second important limitation is that I have only focused on one mechanism that links information technology to urbanization economies. Namely that online information lowers (generalized) transportation costs and increases the geographical scale of local markets. I have focused on local labour- and restaurant markets where I deemed these effects highly likely. It should be noted that there may be local markets that are affected differently, or not at all. Moreover, information technology may alter urbanization economies in a number of different ways than through transportation costs alone. In Section 7.2 I propose more research into one such mechanism, whereby information technology alters search technologies in local markets.

The third limitation relates to the empirical models used. This thesis has made important advances by using longitudinal data to estimate the effect of IT on travel behaviour and on local economic outcomes. This strategy poses several advantages compared to strategies based on cross-sectional data as it allows to use fixed-effects models that control for time-invariant characteristics of units (individuals, places, grid cells). An important limitation is that time varying characteristics of units, that are somehow correlated with the outcome, will still bias estimates of the effects.

I expect this potential source of bias to be more problematic when units correspond to geographical areas (Chapters 4 and 5), compared to when units correspond to individuals (Chapters 2, 3, and 6). Especially if there is information on time varying characteristics of individuals, a stronger case can be made that the effects may be interpreted as causal. Still, the identification approach in Chapters 2, 3, and 6 ignores simultaneity bias, and improvements can theoretically be made by using exogenous variation in teleworking or other measures of IT usage. Chapter 4, about broadband internet and local prospects of employment, may therefore be improved by repeating the analyses at the level of individuals. Concerning Chapter 5, time varying data on geographical penetration of information technology is not available, and it is not straightforward to measure product variety at the level of individuals. Coming closer to causal estimates would be better served by finding a time varying measure, or a credible instrumental variable, for local usage of information technology.

Finally, there are important limitations to the data used in this thesis. Using panel surveys to measure changes in commuting behaviour (in Chapters 2 and 3) may be problematic because chances of attrition increase when respondents relocate (Van Ommeren 1998). In Chapter 4 this problem was circumvented because of the usage of register data, but such data contains much less specific information. For instance, there was no information on the number of hours worked, which required additional assumptions to measure whether someone is employed. Measurement error is another problem that affects the results of this thesis. This is especially relevant when it comes to explanatory variables, and non-random measurement error. In Chapter 5, the share of reviewed restaurants may for instance be influenced by restaurant turnover, which may be higher in vibrant places with more variety. More generally, models based on longitudinal data are more sensitive to errors in measurement, because estimates are based on changes rather than levels. In every Chapter I have therefore carefully assessed potential biases due to measurement error, and I have run sensitivity checks where applicable.

7.4 Policy implications

The results of this thesis have important general implications for spatial policies. First and foremost, the results contribute to the question how advances in information technology will affect future demand for urban density. The idea for the research in this thesis is largely based on a literature that to some extent questions the continued importance of urban density and agglomeration (e.g. Dijkstra et al. 2013, Burger and Meijers 2016). Indeed, I have found evidence that information technology has made proximity a better substitute to urban density: the effect of information technology on urbanization economies is highly dependent on proximity to urban areas. At the same time however my results emphasise the continued importance of urban density: Connecting people to the internet only affects employment rates positively in urban cores, and the effect on information technology on local cuisine variety increases with local population. Overall the results thus confirm the findings of existing studies, that information technology complements urbanization economies (Goldfarb and Tucker 2017). But policymakers should be aware that information technology likely complements 'proximity economies' as well.

This has some practical implications for urbanization policies in the Netherlands. According to the Dutch *National Housing Agenda (*Nationale Woonagenda 2018), 700,000 new homes are needed by 2025. This should be addressed by building 75,000 houses annually, within inner cities and existing built-up areas where possible, in line with the *Ladder for Sustainable Urbanization*. My results warrant an initial priority for urban locations. The advent of information technology seems to have only strengthened the benefits of urbanization. However, since fulfilling this need for housing solely in cities is deemed infeasible, the results provide hope for housing projects at other locations, especially those in proximity of urban areas. Providing digital infrastructure together with transportation infrastructure makes proximity a better substitute for urban density. Considering the steady increase in the usage and usefulness of information technology, I expect this trend to continue.

The second policy implication pertains to whether or not governments should subsidize broadband internet in (rural) places that are underserved by the market. In Chapter 4 I find no evidence that rolling out broadband in rural places affects the prospects of employment. The only effect I find is that broadband internet allows people to have longer commutes, potentially through allowing working from home. When considering to subsidize broadband initiatives, governments should not overstate the expected local economic effects.

The final policy implication relates to the finding that information technology, and the ability to perform some activities from a distance, in many ways allows people to cover more ground rather than less. This should be considered when implementing policies that aim to substitute travel for online activities. The analyses in this thesis do not shed definitive light on the effects of teleworking on transport emissions. Still, the results give ample reason to not overstate the effects of teleworking as a strategy to lower harmful emissions related to commuting.

7.5 Research agenda

The results of this thesis have given new insights in the relation between information technology and urbanization economies. In particular, it has highlighted that information technology allows people to cover more ground for work and consumption purposes, and that this potentially has consequences for the geographical scale at which the benefits of urban density can be enjoyed. In this subsection I highlight two lines of research that deserve particular attention for further research.

The first line of research is based on a lack of theoretical work that relates information technology to economic outcomes across space. In this thesis I have sidestepped this issue and assumed that information technology leads to a reduction in transportation costs and/or search costs. Hypotheses about the effects were then based on existing theoretical models and their comparative statics (e.g. Glaeser and Kahn 2004, Rhee 2008, Autor 2001, Anenberg and Kung 2015). The assumption that information technology more or less directly lowers transportation costs through teleworking is defendable, but when it comes to searching, information technology may be argued to *change* search technologies rather than to lower search costs. Specifically, the internet has changed searching in several markets by providing detailed ex-ante information about opportunities.

A new type of economic search models, known as directed- or competitive search, provides an opportunity to model the advent of ex-ante information about market opportunities, and asses the effects on market outcomes such as prices, product quality and variety (in product markets) and employment and matching (in labour markets). In traditional models of labour market search, buyers and sellers or firms and workers are assumed to meet each other at random, and deals are assumed to arise from bargaining. With directed search, these agents have information to *direct* the search process towards particular (types of) agent, and the terms of trade are posted in advance (Wright et al 2017). I think it is arguable that in many markets, the internet has changed search from being characterized by random search, to be more akin to directed search. Similar parallels between the internet and directed search are drawn in the literature on semi-directed search, where a share of buyers is informed, and the others search randomly (Bethune et al 2016, Lester 2011). Explicitly modelling the internet as the advent of directed search in for instance a 2-region economy (core and periphery, or CBD and suburbs) could provide valuable theoretical insights that can be tested using data.

The second avenue of interest for further research relates to how current dynamics in cities may expedite the process by which IT substitutes for urban density. Urbanisation is meeting its boundaries in a lot of western countries, characterised by a limited supply of (newly built) homes. Where in the 1980s suburban house prices were generally higher than central city house prices, nowadays central city house prices are going through the roof in European and North American cities. In essence, urban density has become more expensive over the last decades. This may well have emanated from an initial complementary relation between IT and proximity, but there are other potential causes including rising incomes, making commuting more costly and increasing demand for luxury goods (Glaeser et al., 2001). From a basic economic perspective, increases in the price of one good will induce substitution effects. It has already been argued by various authors that IT may in principle substitute some of the core functions of cities, including face-to-face contact and demand aggregation to support more varieties (Gaspar and Glaeser, 1998; Sinai and Waldfogel, 2004). Therefore, even if IT solutions did not continually improve, one would expect that high urban house prices induce people to move to or stay put in less dense areas, and to increasingly use IT in ways that substitute the urban density they can no longer afford.

Recent empirical work has already uncovered so called ripple effects of house prices, whereby house price increases that start out in central cities gradually spread out over nearby municipalities (Teye et al. 2017). This suggests that people substitute reasonable proximity to urban centres for density itself. IT can play a role in making this proximity a better substitute to urban density. For instance, teleworking allows people to commute less to their urban job. At the same time, ubiquitous online information, and shared experiences about local options decrease uncertainty about destinations that are further away. Both teleworking and online reviewing have increased steadily since 2008, in countries such as the Netherlands, Sweden, France, and the United States. A first step towards analysing the role of high (urban) house prices is to investigate whether IT usage or availability is associated with relocations to areas with cheaper housing. Data that can be used for these analyses may come from a variety of sources, including travel- and workforce surveys, especially if it can be combined with register data.

7.6 Concluding remarks

This thesis has examined the link between information technology and the extent to which living in proximity of large cities can substitute urban density itself. Several scholars have proposed that such a link exists. In this thesis I have put forward a mechanism that links information technology and 'borrowed' urbanization economies, based on the notion that digital technologies increases the geographical extent of local labour- and product markets. At the micro level I have found that usage of information technologies allows people to accept longer commutes, and to travel further for local consumption purposes. At the macro level I have found that the effects of information technology availability or usage on local outcomes such as employment and product variety, are only present in cities, and in close proximity of cities. Taken together I found evidence that information technology complements urban density, but likely also makes proximity a better substitute for urban density.

References

Andreev, P., Salomon, I., & Pliskin, N. (2010). Review: State of teleactivities. Transportation Research Part C: Emerging Technologies, 18(1), 3–20. https://doi.org/10.1016/j.trc.2009.04.017

Anenberg, E., & Kung, E. (2015). Information technology and product variety in the city : The case of food trucks. Journal of Urban Economics, 90, 60–78. https://doi.org/10.1016/j.jue.2015.09.006

Autor, D. H. (2001). Wiring the Labor Market Published by : American Economic Association. Journal of Economic Perspectives, 15(1), 25–40.

Berry, S., & Waldfogel, J. (2010). Product Quality and Market Size. The Journal of Industrial Economics, 58(1), 1–31. https://doi.org/10.1111/j.1467-6451.2010.00404.x

Bethune, Z., Choi, M., & Wright, R. (2016). Frictional Goods Markets: Theory and Applications.

Burger, M. J., & Meijers, E. J. (2016). Agglomerations and the rise of urban network externalities. Papers in Regional Science, 95(1), 5–15. https://doi.org/10.1111/pirs.12223

De Graaff, T. (2004). On the Substitution and Complementarity between Telework and Travel: A Review and Application (Vrije Universiteit Amsterdam). Retrieved from http://degree.ubvu.vu.nl/repec/vua/wpaper/pdf/20040016.pdf

de Graaff, T., & Rietveld, P. (2007). Substitution between working at home and out-of-home: The role of ICT and commuting costs. Transportation Research Part A: Policy and Practice, 41(2), 142–160. https://doi.org/10.1016/j.tra.2006.02.005

Dijkstra, L., Garcilazo, E., & McCann, P. (2013). The Economic Performance of European Cities and City Regions: Myths and Realities. European Planning Studies, 21(3), 334–354. https://doi.org/10.1080/09654313.2012.716245

Gaspar, J., & Glaeser, E. L. (1998). Information Technology and the Future of Cities. Journal of Urban Economics, 43(1), 136–156. https://doi.org/10.1006/juec.1996.2031

Glaeser, E. L., & Kahn, M. E. (2004). Sprawl and urban growth. In Handbook of Regional and Urban Economics (Vol. 4, 2481–2527). https://doi.org/10.1016/S1574-0080(04)80013-0

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. Journal of Economic Geography, 1(1), 27–50. https://doi.org/10.1093/jeg/1.1.27

Goldfarb, A., & Tucker, C. (2017). Digital Economics. In NBER Working Paper Series (No. 23684). https://doi.org/10.3386/w23684

Graham, D. J., Gibbons, S., London, I. C., & Kensington, S. (2019). Economics of Transportation Quantifying Wider Economic Impacts of agglomeration for transport appraisal : Existing evidence and future directions. Economics of Transportation, 19(May), 100121. https://doi.org/10.1016/j.ecotra.2019.100121

Gubins, S., van Ommeren, J., & de Graaff, T. (2019). Does new information technology change commuting behavior? Annals of Regional Science, 62(1), 187–210. https://doi.org/10.1007/s00168-018-0893-2

Hesse, M. (2016). On borrowed size , flawed urbanisation and emerging enclave spaces : The exceptional urbanism of Luxembourg , Luxembourg. European Urban and Regional Studies, 23(4), 612–627. https://doi.org/10.1177/0969776414528723

Jiang, Y. (2008). The Impact of Telecommuting on the Journey to Work : A Two-Sample Instrumental Variables Approach (University of Maryland). Retrieved from http://drum.lib.umd.edu/bitstream/handle/1903/8506/umi-umd-5557.pdf

Lester, B. (2011). Information and prices with capacity constraints. American Economic Review, 101(4), 1591–1600. https://doi.org/10.1257/aer.101.4.1591

OECD. (2015). The Metropolitan Century. https://doi.org/10.1787/9789264228733-en

Rhee, Hyok Joo (2008). Home-based telecommuting and commuting behavior. Journal of Urban Economics, 63(1), 198–216. https://doi.org/10.1016/j.jue.2007.01.007

Rhee, Hyok Joo. (2009). Telecommuting and urban sprawl. Transportation Research Part D: Transport and Environment, 14(7), 453–460. https://doi.org/10.1016/j.trd.2009.05.004

Sinai, T., & Waldfogel, J. (2004). Geography and the Internet: Is the Internet a Substitute or a Complement for Cities? Journal of Urban Economics, 56(1), 1–24. https://doi.org/10.1016/j.jue.2004.04.001

Teye, A. L., Knoppel, M., Haan, J. De, & Elsinga, M. G. (2017). Amsterdam house price ripple effects in The Netherlands. Journal of European Real Estate Research, 10(3), 331–345. https://doi.org/10.1108/JERER-11-2016-0041

Van Ommeren, J., & Fosgerau, M. (2009). Workers' marginal costs of commuting. Journal of Urban Economics, 65(1), 38–47. https://doi.org/10.1016/j.jue.2008.08.001

Wright, R., Kircher, P., Benoit, J., & Guerreri, V. (2017). Directed Search: A Guided Tour. In NBER Working Paper Series (No. 23884).
Zhu, P. (2012). Are telecommuting and personal travel complements or substitutes? Annals of Regional Science, 48(2), 619–639. https://doi.org/10.1007/s00168-011-0460-6

About the author

Duco de Vos was born in Dordrecht on September 30, 1990. He grew up in Streefkerk and after spending a year working in Berlin, he went to attend the Erasmus University Rotterdam where he studied economics. After obtaining his bachelor degree, he took a break to travel and work in Switzerland and France, and subsequently he entered the Spatial, Transport, and Environmental Economics Master programme at VU Amsterdam. He continued to work there, as a researcher at the Spatial Economics department, awaiting a suitable PhD position.

In April 2016 Duco started his PhD at Delft University of Technology under the supervision of Evert Meijers and Maarten van Ham. During the PhD Programme, he presented his work at various international conferences and he gave tutorials and lectures across different faculties and universities. In 2018 Duco was a visiting PhD student at the European University Institute.

Publications

Does Broadband Internet Allow Cities to 'Borrow Size'? Evidence from the Swedish Labour Market (Regional Studies)

Working from Home and Commuting: Heterogeneity over Time, Space, and Occupations (with Maarten van Ham and Evert Meijers; IZA Discussion Paper No. 12578)

Information Technology and Local Product Variety: Substitution, Complementarity and Spillovers (with Evert Meijers; Tijdschrift voor Economische en Sociale Geografie, 110(4), 486-506)

Why bright city lights dazzle and illuminate: A cognitive science approach to urban promises (with Rodrigo Cardoso, Evert Meijers, Maarten van Ham, and Martijn Burger; Urban Studies, 56(2), 452-470)

Working from home and the willingness to accept a longer commute (with Evert Meijers and Maarten van Ham; The Annals of Regional Science 61(2), 375-398)

Parking Occupancy and External Walking Costs in Residential Parking Areas (with Jos van Ommeren; Journal of Transport Economics and Policy (JTEP) 52(3), 221-238)