Doctoral Dissertation

Who benefits when IKEA enters local markets in Sweden?

An empirical assessment using difference-in-difference analysis, synthetic control methods and Twitter sentiment analysis

Yujiao Li

Microdata Analysis
School of Technology and Business Studies

Dalarna University, Sweden
September 2018

Dalarna Doctoral Dissertations No.7
ISBN: 978-91-85941-85-8
Electronical version available at http://du.diva-portal.org
Printed by Printeliten, Falun
Dalarna University, Borlänge, Sweden

Abstract

Policy makers often spend considerable amounts of money in attracting IKEA to their region, despite having no empirical measurements on the expected contribution to the local economy. As such, an empirical study of the economic and social impact of new IKEA stores can aid political decision-making, and contribute to the literature regarding how big-box retail entry affects the regions they enter.

This dissertation aims to estimate: the impact of IKEA entry on incumbent retailers productivity, and investigate if the impact is heterogenus, depending on local market size, type of retail industry, distance to surrounding retailers, and firm size; IKEA-entry effects on the average labor productivity in durable goods retailing in the entry regions; and, finally, public opinion regarding IKEA entry.

Papers I-III separately examine four factors of potential heterogeneity to estimate IKEA-entry effects on incumbent retailers. Paper I finds that market size matters: smaller rural regions have greater IKEA effects. Paper II considers two factors: firm industry and distance, and confirms that IKEA-entry effects dissipate over distance. The positive impact of IKEA entry on incumbent retailers is limited to those selling complementary goods to IKEA. No positive effects were found for the urban entry in Gothenburg in the two first papers, which is somewhat surprising. Paper III found that a positive effect exists also in Gothenburg, but it is limited to relatively small incumbent retailers with a capital stock below 1 500 000 SEK. Policy-making tends to consider IKEA overall effects on entry municipalities, besides IKEA-spillover effects on firms. Paper V shows that rural regions are affected by IKEA entry, while larger urban markets are not.

For the social effects of IKEA, Paper VI uses Twitter text-mining to study public opinion regarding IKEA entry into local markets. The new IKEA stores under study attracted significant public attention at the time of entry, with mostly positive attitudes toward the new stores. The favorite topics for discussion at the time of the different IKEA entries were heterogeneous, depending on location.

Methodologically, Paper I uses traditional Difference-in-Difference (DID) to gain an initial understanding of IKEA-entry spillover effects in four regions; Paper II extends to Spatial DID to measure the spatial interaction between firms; Paper III uses Panel Smooth Transition Regression to identify heterogenous effects due to firms size. Paper IV investigates a new treatment effects estimation aproach, Synthetic Control Method (SCM), to explore when the SCM is powerful, and how to improve its performance; Paper V uses SCM to estimate IKEA effects at municipality level. In addition, to making SCM-developed research readily available for other researchers, the author of this thesis also published one web-application for implementing a synthetic control method power test, and another application to implement parametric and non-parametric estimation and inference.

These findings confirm that IKEA has a positive effect on the regions they enter. Nevertheless, in their decision to allow a big-box retail entry into their local community, governments should be aware that the impact of such entry will depend on the size of the existing retail market, the type of existing retail industry, and the size of existing retailers in the entry region.

Keywords: big-box entry effect; spatial spill-over effect; heterogeneous causal effect; difference-in-difference; synthetic control method; sentiment analysis.

Acknowledgements

During work on this dissertation, I have received much help and encouragement from my family, friends and colleagues at Dalarna University. Research funding from the Swedish Retail and Wholesale Development Council (Handelns Utvecklingsråd, Grant Number 2015:4) is also gratefully acknowledged.

I would especially like to mention Niklas Rudholm and Kenneth Carling who have worked with me the most. Niklas, you have continually helped me to improve, and coached me with your patience and kindness. Kenneth, you have trained me in becoming an independent researcher and taught me the spirit of research.

I could not achieve this work without the help of my supervisors: Mark Dougherty, thank you for guiding me when I felt lost. Hasan Fleyeh, you helped me avoid mistakes and taught me to be grateful, which is invaluable in life. Moudud Alam, thank you for your intelligent suggestions when I found myself stuck in my research. Oana Mihaescu, your sincere and warm thoughtfulness were essential for my PhD life, your excellent work and kindness have empowered me.

I would like to express my appreciation to my co-authors: Lars Rönnegård, Johan Håkansson, Kristin Svenson, Mengjie Han and Zuzana Macuchova. Your cooperation and contributions have been very important. Siril Yella, your support regarding research data and academic suggestions have been very useful. Thanks to Anders Forsman, Roger Nyberg and Renée Flacking for providing me with administrative assistance and frank advice.

I would also like to express a warm thanks to the following:
To my PhD colleagues in Microdata Analysis: Åsa Grek, Ilias Thomas, Vijay Pratap Paidi, Asif M Huq, Murshid Saqlain, Ross May, Somayeh Aghanavesi, Eugenio Conti, and Magnus Fahlström. Also to my colleagues and friends: Ola Nääs, Catia Cialani, Xiaoyun Zhao, XingXing Zhang, Yukai Yang, Yanina Espegren, and Pascal Rebreyend.

To my friends in HUI: Sven-Olov Daunfeldt, Helena Nilsson, Anders Bornhäll, and Hans Seerar Westerberg: The HUI workshop provided me with a memorable presentation and party.

To my students in the Business Intelligence Department. Your participation in my course gives me much pleasure in the sharing of knowledge. Your appreciation of my teaching makes me very happy.

To my charming piano teacher, Marcus Moszny, who motivated me to play the piano more, while my neighbors encouraged me to move. Thanks to my ballet teacher, Irina Basova, who made my life in Dalarna more colorful and joyful, even though I was the oldest dancer in the class. To my Latin dance teacher, Hang Zheng, for inspiring me to “dance by heart”.

To my parents, for your love and for giving me an education, which encourages me to be happy and to use my talents. Thanks to my boyfriend, Charlie Lindgren, and your family, for,your big heart and concern for everyone, and your love. You gave me the passion to learn Swedish, although my Swedish did not live up to my passion.

Finally, I would like to thank my lucky stars to have been able to participate in the PhD programe at Dalarna University. I studied *big-box entry*, and coincidentally learnt how to think *outside the box*.

List of Papers

I. Håkansson, J., Li, Y., Mihaescu, O., Rudholm, N., 2016, Big-box retail entry in urban and rural areas: Are there productivity spillovers to incumbent retailers? HUI Working Paper No. 118. *Submitted to International Review of Retail Distribution and Consumer Research.*

II. Han, M., Mihaescu, O., Li, Y., Rudholm, N., 2018, Comparison and one-stop shopping after big-box retail entry: A spatial difference-in-difference analysis, *Journal of Retailing and Consumer Services,* 40: 175-187.

III. Li, Y., Mihaescu, O., Rudholm, N., Håkansson J., 2018, Agglomeration economies in urban retailing: Are there productivity spillovers when big-box retailers enter urban markets? HUI Working Paper No. 137. *Submitted to Applied Economics Letters.*

IV. Carling, K. and Li, Y., 2017, The Power of the Synthetic Control Method, Working papers in transport, tourism, information technology and microdata analysis No. 2016:10. *Submitted to Sociological Methods and Research.*

V. Rudholm, N., Li, Y., Carling, K., 2018, How Does Big-Box Entry Affect Labor Productivity in Durable Goods Retailing? A Synthetic Control Approach. HUI Working Paper No. 130. *Submitted to Oxford Bulletin of Economics and Statistics.*

VI. Li,Y. and Fleyh. H., 2018, Twitter Sentiment Analysis of New IKEA Stores Using Machine Learning, International Conference on Computer and Applications, *IEEE conference paper.*

Contents

Part I

1. Introduction

This study assists policy-makers in deciding whether it is worth investing in new IKEA stores in their local community. Gains and losses due to recent IKEA entries in Sweden are assessed by using statistical models, machine-learning methods and other microdata analysis approaches. This doctoral dissertation is in the subject of microdata analysis, a subject which comprises a number of collaborating fields, such as Artificial Intelligence, Decision Support Systems, Simulation and Statistical Inference, Data Modelling, Experimental Design, and Economic Decision-Making. Research is conducted in close collaboration with regional and national companies and organizations in industries, such as transport, manufacturing and tourism and retail, and these industries also often act as financers and data providers.

IKEA is one of the world's largest big-box retailers, having 403 stores in 49 countries, and with retail sales of EUR 38.3 billion in 2017.[1] New IKEA stores are believed to positively affect the regions where they enter, and local policy-makers often spend considerable amounts of money in attracting IKEA, despite having no empirical measurements regarding its contribution to the local economy. Thus, estimating how IKEA entry affects municipalities and incumbent retailers where IKEA enter the market not only contributes to big-box research literature, but also provides decision support for policy makers. Finding positive spillover effects on incumbent retailers in the entry regions could justify local governments in subsidizing IKEA entry, but in the absence of such positive spillover effects, the use of taxpayers' money for such subsidies cannot be justified on economic efficiency grounds.

The literature on how big-box retail entry affects the economy of the communities where they entrer is inconclusive, and the social effects of big-box entry have not, to my knowledge, been investigated previously. Previous studies have mainly focused on the entry of Wal-Mart stores in the USA, but with inconclusive results regarding the impact on retail revenues or employment of incumbent retailers. For example, Basker (2005), and Hicks (2007) both report positive effects regarding retail employment, whereas negative effects are reported by Jones and Doucet (2000), and Neumark et al., (2008). Moreover, the results from the Wal-Mart studies cannot be generalized in respect of entry by IKEA stores. IKEA is more focused towards selling durable goods than Wal-Mart, and consumers travel farther distances to buy durables compared to non-durable goods (Brown, 1993).

One reason why current results are inconclusive is that big-box entry effects are generally complicated to study, and the traditional econometric methods used in most previous studies have their limitations. IKEA-entry economic effects might be heterogeneous depending on local market size, type of industry, incumbent firms' size and other unknown factors. Regarding IKEA-entry effects on regional economic growth, which might decline as time passes, but it is assumed to be constant in most previous studies. Recent development of research in this field mostly extends the study object to other big-box brand retailers, instead of improving accuracy and effectiveness of estimation.

[1] https://highlights.ikea.com/2017/facts-and-figures/.

The essential idea of treatment-effect estimations is to measure the difference in outcome between treated units and well-chosen control units not exposed to treatment. In our case, observations regarding the IKEA-entry regions are available, but choosing suitable counterfactual outcomes without IKEA entry (i.e. the control group) is more challenging. In this study, we identified counterfactual outcomes by applying approaches from the perspective of economic theory as well as data-driven optimization.

This study aims to make a contribution to the following three research areas, with the major research questions related to each area presented after each bullet point. Figure 1 visualizes how the six papers in this thesis are related to the study of either economic or social effects of IKEA entry.

- Economic effects:
    - How does IKEA entry affect incumbent retailers in the entry regions? Specifically, how does market size, type of retail industry, and incumbent firm size affect the impact of IKEA entry on the productivity of incumbent retailers?
    - How does entry by IKEA affect average labor productivity in durable goods retailing in the entry regions?
- Social effects:
    - Do people care when IKEA enter their local market?
    - Are residents of the entry region positive or negative to IKEA entry?
    - What do local people talk about when they talk about IKEA?
- Methods:
    - For IKEA-entry spillover effects, how do we measure these correctly, taking into account potential heterogeneity in the effects?
    - For dynamic IKEA-entry effects measured on the regional level, how much shall one trust inference in the standard synthetic control method, and how can we improve its performance?
    - For public opinion regarding IKEA entry, how do we build a valid and accutrate sentiment prediction model for Swedish text messages on Twitter?

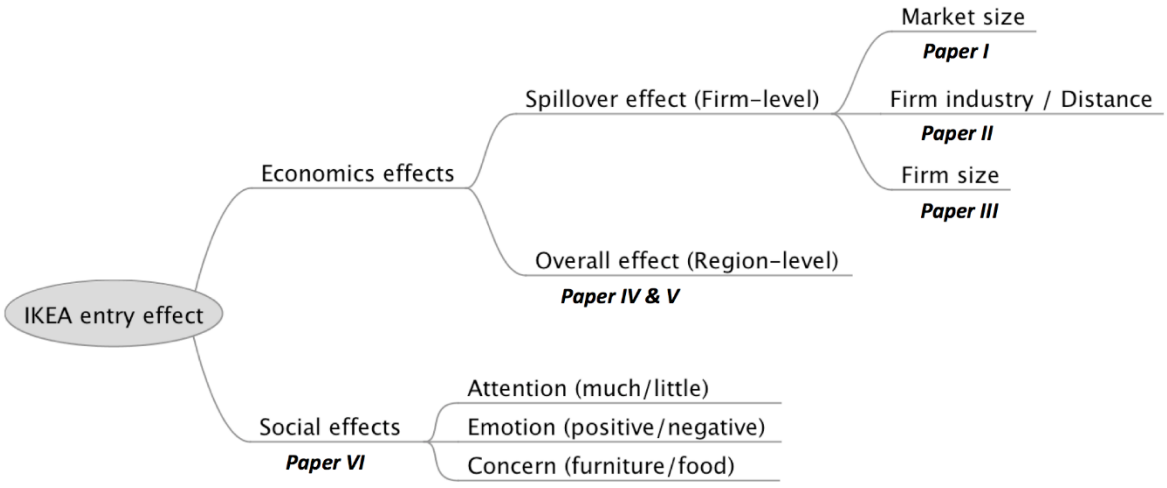

Figure 1. Structure of the papers.

Three different datasets are collected to empirically study the research questions: micro- and macro-level economic data, and opinions text data. To estimate the economic effects of IKEA entry on incumbent retailers in the four Swedish IKEA-entry regions from 2004-2007, a dataset including location coordinates and all information in the annual reports of the retail firms in the entry and control regions is used. When estimating the impact of IKEA entry at the regional level, data regarding durable goods retail sales and employment was collected from Trade in Sweden (Handeln i Sverige), a free database accessible at http://www.handelnisverige.se/ (in Swedish). This data was combined with data from Statistics Sweden, regarding the percentage of the population with a university education, the number of patents, and quality of the infrastructure in the municipalities. Finally, to study public opinion regarding IKEA entry in the local communities, Twitter data posted in four IKEA entry regions containing the term "IKEA" during the IKEA entry period was collected.

Papers I-III identified the effects of IKEA entry on incumbent retailers using DID based models. The results show that there are positive effects of IKEA entry on incumbent firm productivity, that the effects are larger in rural areas, where the size of IKEA in relation to the existing local retail market is large. Significant IKEA effects exist in retailers selling complements to products sold by IKEA, instead of substitutes. Finally, for the urban entry in Gothenburg, there is a positive impact of IKEA entry on incumbent retailers, but only for relatively small retailers.

To investigate IKEA effects at municipality level, the above DID model based methods are inferior, due to strict assumption and limited applied scope. Recent big-box effect studies scarcely improve accuracy and effectiveness of estimation. Thus, Paper IV investigates a new and popular approach: the Synthetic Control Method (SCM). To improve SCM's performance, Paper IV applies simulations to explore which data features affect the SCM's power and improve its performance, by developing a parametric specification for the treatment effects. The author of this thesis has also published two web-

applications[2] related to the SCM, where researchers can automatically implement power test and parametric and non-parametric estimation of the treatment effects.

Paper V used the SCM developed in Paper IV to estimate the effect of IKEA entry on municipality-level labor productivity, measured as consumer price-adjusted sales per employee in durable goods retailing. The results show that there is a positive effect of IKEA entry on productivity, but only when entry takes place in smaller markets, with no positive effects found for the urban entry in Gothenburg.

Paper VI investigated IKEA social effects by Twitter text-mining. The conclusion is that a new IKEA store receives much public attention on Twitter, indicated by a significant increase in the frequency of tweets. Most of these tweets expressed positive sentiment toward the new IKEA stores.

2. Theoretical background

Theories of agglomeration economies (Marshall, 1890; Hotelling, 1929; Weber, 1929) indicate that there are at least four different pathways for IKEA entry to have an impact on productivity in the entry regions, either directly by itself, or indirectly through affecting incumbent retailers in the entry regions.

The first pathway is the direct effects of IKEA entry on productivity. From a purely mathematical perspective, the big-box entry will increase average productivity in the region, if it is more productive than the surrounding incumbent retailers. Wal-Mart alone was responsible for nearly half of the productivity growth in the general merchandise sector in the US retail market between 1982 and 2002, and the real value-added per worker in Wal-Mart was 40% higher than that of other US general merchandise retailers (Basker, 2007). By analogy, IKEA entry is expected to increase average productivity in the entry regions, especially since IKEA is more efficient than most other Swedish durable goods retail firms.

The second pathway is an indirect effect due to the displacement of low-productivity retailers in the entry regions. Although this has not been investigated for IKEA entry specifically, Jia (2008) reported that Wal-Mart entries caused 50–70% of the net exit of small discount retailers in the US market, and that exiting establishments were 25% less productive than the surviving incumbents. In the Swedish retail food market, big-box entry forced low-productivity stores to exit, and surviving stores experienced productivity increases of approximately 3% (Maican and Orth, 2012). These previous studies indicate that entry by big-box retailers, such as IKEA, are likely to result in low-productivity retailers exiting the market, thus raising the average productivity level. Note that this study did not focus on firms' entry and exit due to IKEA entry, but the averaged change of productivity output.

The third way IKEA entry could affect productivity in the entry regions is through supply-side

[2] https://wishes.shinyapps.io/intervention_effect_study/
https://wishes.shinyapps.io/scm_simulation/

spillovers to incumbent retailers. According to agglomeration theories, firm co-location decreases input costs, facilitates labor matching, and creates knowledge spillovers (McCann, 2001; O'Sullivan, 2003). Other literature has focused on knowledge spillovers as a way of increasing productivity (Lucas, 1988; Grossman and Helpman, 1991; Glaeser, 1999). These researchers have pointed out that inter-firm learning might be affected not only by geographical, but also by other types of proximity (e.g., cognitive, organizational, social and institutional).

Finally, entry by IKEA will cause an increase in the degree of retailer co-location in the entry regions that is likely to create demand-side spillovers to incumbent retailers. Demand-side spillovers occur when the output of one retailer is positively affected by the locationand output of other retailers. There are two types of demand-side spillovers, depending on whether the products sold are complements or substitutes. The co-location of retailers selling close substitutes establishes the basis for comparison shopping, while the co-location of retailers selling complements establishes the basis for one-stop shopping (Han et al., 2018), and both of these could have a positive effect on incumbent firm productivity when IKEA enters the market.

Daunfeldt et al. (2017) investigated the average effect of the four IKEA entries under study in this thesis on revenues in durable goods retailing in the entry areas, and found that IKEA entry increases revenue by an average of 20%. However, one of the main points made in this study is that IKEA entry will have different effects on the local markets where they enter depending on four factors: the size of the market, the type of incumbent retailers active in the entry markets at the time of IKEA entry, distance to the new IKEA and the size of the incumbent retail firms. But, the average effect found by Daunfeldt et al. (2017) that ignores factors for heterogeneous effects could be insufficient. Thus, it is necessary to estimate IKEA-spillover effects on firms by distinguishing the subgroup of the firm, based on the above-mentioned four factors.

The first factor to consider is market size (rural/urban). IKEAentry effects on incumbent firms should be heterogeneous with respect to local market size, if one assumes diminishing returns to retail capital. In smaller, more rural markets, the existing total capital stock in retailing is low at the time of IKEA entry, making each new unit of retail capital effective in increasing output. As the IKEA stores are of (roughly) equal size, the impact in a market, such asHaparanda, where total durable goods sales before IKEA entry was 140 million SEK, will then probably be more pronounced than for the entry in Gothenburg, where durable goods retail sales was 12 945 million SEK the year before IKEA entry.[3] Thus, IKEA entry has different effects depending on whether the entry market is large or small in size. To test this factor's role in empirical data, Papers I, II, III and V studied effects by distinguishing local market size.

The second factor to study is the type of industry (substitutes/complements). Co-located retailers

[3] Data regarding retail sales in the municipalites is from Trade in Sweden (http://www.handelnisverige.se/).

selling substitutes to IKEA perhaps receive increased demand as more customers are attracted to the new IKEA entry area, but it also faces strong competition in the region after IKEA entry. For retailers selling complements, IKEA creates instead an increased demand for their products, without creating increased competition. Paper II in the thesis will study this factor.

The third factor of importance is the incumbent retailer's distance to the new IKEA store. Several previous studies indicate that each individual firm's productivity is affected by nearby firms. Cardamone (2014) found evidence in support of productivity spillovers across firms due to spatial proximity, and a number of papers have pointed out the nexus between spatial agglomeration and knowledge spillovers (Aldieri and Cincera, 2009; Arrow,1962; Audretsch and Feldman, 2004; Koo, 2005; Orlando, 2000; Romer, 1986) as a cause of productivity spillovers. As such, spatial interaction among firms should not be ignored in an estimation of IKEA entry effects, and this is also taken into consideration in Paper II in the thesis.

The final factor potentially affecting the impact of IKEA entry on incumbent retailers is firm size (small firm/large firm). When IKEA enters the market, incumbent firms often adjust their use of capital and labor, and if there are diminishing returns to capital and labor in retailing, this will have a larger impact on smaller firms. Firm growth increases at a diminishing rate with firm size (Evan, 1987), and if this occurs, IKEA entry will have a larger impact on smaller firms in the entry areas. This is investigated in Paper III in the thesis.

All of the above-mentioned effects of IKEA entry deal with the economic impact of IKEA entry on local communities, but one should also consider the social impact on the entry regions. This could perhaps be studied using Twitter sentiment regarding new IKEA stores, and this study is carried out in Paper VI in the thesis. Curtin (1992) claimed that consumer sentiment is the most closely-watched and intensely debated indicator of future economic trends, while Carroll et al. (1994) found that consumer sentiment indexes have predictive power on future changes in consumption spending. He stated two possible interpretations: sentiments are an independent driving factor in the economy, and that changes in sentiments not only forecast changes in spending but also cause them, or that sentiments reflect the overall outlook for the economy, i.e. when consumers are optimistic about the outlook for the economy, they give upbeat responses to interviewers. Hence, on average, optimistic expectations are substantiated, and spending eventually increases as foretold by the sentiments. Local residents' sentiments regarding IKEA indicate then their potential behavior as consumers in the future.

3. Data

Data quality decides the reliability of all empirical research, and in this thesis data quality will be assessed using the EuroStat Standard Quality Indicators (Eurostat 2005d). Data quality refers to five aspects: Relevance, Accuracy and Reliability, Timeliness and Punctuality, Comparability and Coherence, Accessibility and Clarity.

The study objects in this thesis are the four IKEA store entries in Sweden during the period 2004 to 2007. Three datasets are used to estimate IKEA entry spillover effects on incumbent retailers, IKEA entry overall regional effects, and IKEA entry social effects. The datasets used are: firms' financial data as recorded in their annual reports and collected by the Swedish Companies Registration Office; regional economic data collected by HUI Research and Statistics Sweden; and IKEA Twitter data collected by Twitter.

3.1 Firm-level annual report data

To investigate if there are any IKEA spillover effects on incumbent retailers in the entry areas, the collected data needs to contain information regarding the business development of incumbent retailers, both prior- and post- IKEA entry, as well as for a well-chosen control group consisting of retailers from non-IKEA entry regions of the country. All limited liability firms in Sweden are required to submit annual reports to the Swedish Companies Registration Office (CRO).[4] The data was collected by PAR, a Swedish consulting firm that compiles this information from CRO, and delivered to the research group.

The dataset thus consists of annual report data from limited liability companies in the retail trade industry that were active either in the IKEA-entry regions, or the corresponding control regions during the period 2001 to 2010. The entry regions are Haparanda, Kalmar, Karlstad and Gothenburg, while control regions are Borlänge, Uddevalla, Umeå and Stockholm. The main variables of interest include the retail firms' annual sales, use of labor and capital, SNI-codes determining industry, coordinates determining location, an organization number determining the firm's ID, and the year of the annual report. In Papers I, II and III, a change in productivity is measured as a change in consumer price index (CPI) adjusted sales at the time of IKEA entry, holding the use of inputs, capital and labor constant. The data over CPI was provided by Statistics Sweden, and the coordinates used to determine location were generated by ArcGIS in the "World Geodetic System 84 (WGS84)", which is a standard for use in cartography, geodesy, and navigation by GPS.

3.2 Municipality-level economic data

To study IKEA-entry effects on regional level productivity, measured as increases in CPI-adjusted sales per employee in durable goods retailing, municipality-level data on sales and employment in durable goods retailing is required. The dataset needs cover all 290 municipalities in Sweden because the suitable control regions to the IKEA entry ones are unknown, and will be selected from these municipalities by a data-driven approach. Besides similar regional productivity development before IKEA entry, other features of the municipalities should also be used in the matching procedure to obtain an optimal counterfactual. Thus, data regarding durable goods sales, durable goods employment, and population in the municipality was collected from Trade in Sweden (http://www.handelnisverige.se/), a

[4] Until 2005, the annual reports were instead delivered to the Patent and Registration Office (PRV).

database that originates from Statistics Sweden. The sales and employment variables are from the annual reports, if a firm in a specific municipality is a limited liability firm, and from other Statistics Sweden databases, if the firm is not a limited liability firm required to submit an annual report. Nevertheless, the sales data of the firms is used for purposes of taxation, and should thus be of good quality, although it should be noted that not all firms in this dataset are exposed to external audits. In addition, data regarding the percentage of the population with a university education, the number of patents awarded to firms in the region, and the quality of the infrastructure in the municipalities was all collected from Statistics Sweden.

3.3 Twitter text data

To understand public opinion regardinge IKEA entry, data is collected from Twitter, a micro-blogging platform, which has been widely used for sentiment analysis. Twitter data from four IKEA-entry regions was collected by retrieving tweets text containing the term "IKEA", and posted within a one-month interval of the IKEA-entry date. Four cities are investigated, namely Umeå-Sweden, Halifax-Canada, Fishers-Ireland and Sheffield-UK. The tweets in Umeå are written in Swedish, while the other tweetsare in English. IKEA established a new store in Umeå in 2016, while in the three other regions IKEA was established between September to October, 2017. One-week historical data can be collected in Twitter free of charge, while there is a charge to access earlier data. For the three IKEA stores that opened in 2017, data was collected using web-scraping, several times over the period August-November, 2017. However, data for the IKEA entry in Umeå in 2016 was too early to access, and data was thus acquired from the professional Twitter data platform named "Discovertext" (https://discovertext.com/).

3.4 Data quality assessment

According to the five EuroStat Standard Quality Indicators, we deem the datasets used to be of good quality.

- Relevance

Relevance denotes whether all necessary statistics are produced and related to the research goal, i.e. why the collected data suits the research questions. As explained, to estimate IKEA-spillover effects on firms and on regions, Swedish firm-level financial data, and Swedish municipality-level economics data is required because it contains information regarding the business/economics development of all firms/regions prior- and post- the IKEA-entry year. Tweets containing "IKEA" are currently the best data source for public opinion studies, because residents can freely express their opinions about IKEA entry during the entry dates. By setting location, time and keyword in web-scraping, the collected data focused particularly on residents' comments on the new IKEA stores.

- Accuracy and reliability

Accuracy denotes the closeness of computations or estimates to the (unknown) exact or true values.

Reliability means that the source data, intermediate results and statistical outputs are regularly assessed and validated, or revisions are regularly analysed in order to improve statistical processes. Data in the annual reports is exposed to external audits and also scrutanized by the Swedish Tax Authority when levying taxes on the limited liability firms. Trade in Sweden is built to a large extent on annual report data, and should therefore be of similar quality, and the data collecting processes of Statistics Sweden are periodically scrutinized and revised to uphold a high quality. Twitter data is directly collected without any artifical revision. As such, we deem that there is a high likelihood that the datasets used are both accurate and reliable.

- Timeliness and Punctuality

Timeliness reflects the length of time between its availability and the event or phenomenon it describes, and Punctuality refers to the time-lag between the release date of data and the target date when it should have been delivered. The annual report data is collected during the year following the year of business that it describes, but otherwise the three datasets used do not have any time-lag problems.

- Comparability and Coherence

Comparability aims at measuring the impact of differences in applied statistical concepts and measurement tools/procedures when statistics are compared between geographical areas, non-geographical domains, or over time. Coherence of statistics is their adequacy to be reliably combined in different ways and for various uses. The methods of collecting and compiling data in the annual report and Trade in Sweden datasets are similar all over Sweden, and any changes to the rules or regulations regarding annual reports are also imposed at the same point in time. The only major change in the collection of this data is that during the final three months under study in our data, external auditing was made voluntary for Swedish micro-firms with fewer than three employees. Twitter data is assessed as being comparable and coherent because it is collected from a public platform without differences in measurement methods during the data collection period.

- Accessibility and Clarity

Accessibility refers to the physical conditions under which users can obtain data. Clarity refers to the data's information environment, whether data is accompanied with appropriate documentation, metadata, and illustrations. The annual report data is publicly available. The Trade in Sweden database is also publicly available, and data can be downloaded from the website http://www.handelnisverige.se/. Finally, Twitter data can be web-scraped by users, and historical data is available for researchers and others for a fee from Discovertext (https://discovertext.com/).

4. Methods

For IKEA economic effect estimations, the fundamental evaluation problem is that a region that receives a new IKEA store cannot also be observed in the counterfactual state of not receiving that entry.

IKEA does not enter regions randomly, but rather after a process of finding the most profitable entry regions. To identify valid counterfactuals, we have in the first three papers of the thesis selected regions that IKEA itself deemed suitable for entry, either before or after the study period, and used retailers in these regions as controls. In the SCM papers, the counterfactual was created by synthesizing a best-matched unit before IKEA entry, in terms of the development of the outcome variable and several covariates. Regarding the IKEA social effect assessment paper, the challenge is to figure out a well-functioning Swedish sentiment analysis model. Figure 2 presents the methods applied in all papers.

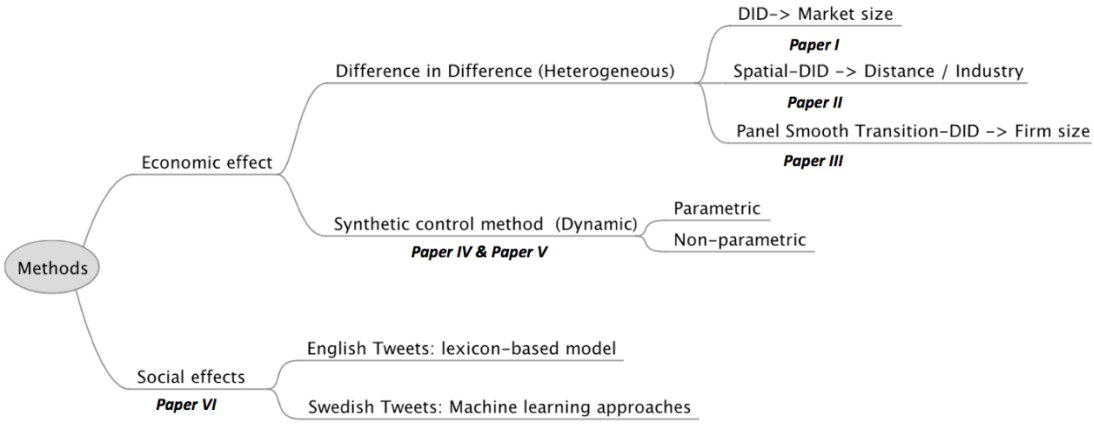

Figure 2. Methods used in the different papers

4.1 Difference-in-Difference based methods

A number of recent developments have taken place in recent years in the econometrics toolkit for addressing treatment effects estimation of policy changes, and one standard method is to use Difference-in-Difference (DID) estimation. Yet, treatment effect estimation often benefits from combining DID with other approaches (Athey and Imbens, 2016). Thus, this dissertation proposes a combination of statistical models, machine learning methods, and non-parametric approaches to assess the effects of IKEA entry on the local entry communities.

- **DID**

The essence of treatment effect estimation is to establish the counterfactual state of a treated region or firm, had it not received treatment (IKEA entry). DID estimation is a standard tool for evaluating the effects of public interventions and other treatments of interest on relevant outcome variables. In Paper I of the thesis, we use a production function difference-in-difference model to identify the effect of IKEA entry on incumbent retailers in the regions where IKEA enters the market. Using a Cobb-Douglas production function as an example, such a model can be written as follows:

$$y_{it} = \beta_0 + \beta_1 L_{i,t-1} + \beta_2 K_{i,t-1} + \alpha_1 TR_i + \alpha_2 TP_t + \alpha_3 (TP_t \cdot TR_i) + \varepsilon_{it}. \quad (1)$$

where y_{it} is the natural logarithm of output in retail firm i at time t, and $L_{i,t-1}$ and $K_{i,t-1}$ are the corresponding natural logarithms of labor and capital inputs, both lagged one period to alleviate a potential endogeneity problem. Output is measured as CPI-adjusted sales of firm i at time t. TR_i is an indicator variable equal to one, if firm i is located in the entry region, and zero otherwise, and TP_{it} is an indicator variable equal to one, for time periods after IKEA entry, and zero otherwise. The parameter α_3 (the coefficient of interaction of terms TP_t and TR_i,) estimates the effect of IKEA entry on the output of retailers located in the entry regions, compared to their own output in the period before entry, and the output of retailers in the control regions throughout the study period. In empirical studies, $100(e^{\alpha_3} - 1)$ is calculated to represent the change in productivity due to IKEA entry, in percentage terms. The intuition behind the setup of this regression model is simple: if retailers in the entry areas are able to increase their output more than those in the control area, while holding the use of inputs constant, they have become more productive due to IKEA entry.

- **Spatial DID**

In Paper II of the thesis, the impact of IKEA entry on an individual retail firm is also assumed to be affected by how adjacent retail firms are affected by IKEA, and closely located firms affect each other more than those located further away. Ignoring such spatial endogeneity will bias the IKEA-entry effects estimation. Spatial endogeneity has been addressed in some previous papers using geographic variables as instruments for capturing the probability of big-box entry, such as the distance from the store's corporate headquarters (Basker, 2007; Neumark et al., 2008).

To model the negative correlation between the retail firms' output and distance from each other, the distance decay function $\mathrm{corr(d)} = \exp\left(-\frac{\mathrm{d}}{\varphi}\right)$ is used, where d is the distance between firms, and where the distance threshold φ represents the distance, where the correlation will decrease to zero. This range parameter is specified by the empirical data-driven semi-variogram approach. Following Delgado and Florax (2015), the correlation function forms the spatial lag operator W_s (spatial weight matrix), and it is a block-diagonal row standard matrix. The interaction term $\mathrm{TP_t \cdot TR_i}$ in Eq. (1) is thus replaced by a spatial interaction $\mathrm{(I + \rho W_s)TP_t \cdot TR_i}$. The parameter ρ specifies how strong the correlation between co-located retailers is, and it is estimated by an iterative approach to the maximum likelihood function (Bailey and Gatrell, 1995). The coefficient α_3 in the spatial difference-in-difference model (Eq. 2) estimates the average treatment effect while considering the retail firms' spatial endogeneity.

$$y_{it} = \beta_0 + \beta_1 L_{i,t-1} + \beta_2 K_{i,t-1} + \alpha_1 TR_i + \alpha_2 TP_t + \alpha_3(\mathrm{I + \rho W_s}) \cdot (TP_t \cdot TR_i) + \varepsilon_{it}. \quad (2)$$

- **Panel smooth transition regression - DID**

As discussed in the theoretical framework, the elasticity of output to changes in labor and capital can be heterogeneous depending on firm size. If the output of small firms is more sensitive to a one-unit

increase of capital (or labor) than that for big firms, this means than applying a unified DID for all firms will bias the IKEA-entry effects estimation. To solve the problem of potentially heterogeneous output elasticities between big and smal firms, it is necessary to identify subgroups before applying the DID model.

In Paper III, we applied panel smooth transition regression (PSTR; Gonzàlez et al., 2005), which allows parameter heterogeneity in the regression coefficients. This model aims to separate data into several homogenous subgroups (regimes), by identifying which covariates should be used for separation, and estimating the separation threshold value. Model diagnostics are then used to show whether the number of regimes is correct by testing if there is any remaining non-linearity. In the case of IKEA entry in Gothenburg, there are two different regimes (small/big firm subgroups) with different regression coefficients in the Cobb-Douglas production function (Eq. 3). Capital is identified as a more suitable separation variable than labor, and the model diagnostics confirms that two subgroups should be used.

The advantage of PSTR is that its additive non-linear term (continuous functions of capital or labor) can fluctuate between two subgroups so as to represent heterogeneity. The applied PSTR model is specified as follows:

$$y_{it} = \mu_i + \beta_1 L_{it-1} + \beta_2 K_{i,t-1} + \left(\beta_3 L_{it-1} + \beta_4 K_{i,t-1}\right).g(q_{it},\gamma,c) + \varepsilon_{it}\,,$$

$$\text{where } g(q_{it},\gamma,c) = \left(1+\exp\left(-\gamma(q_{it}-c)\right)\right)^{-1} \qquad (3)$$

where y_{it} is the log-sales of firm i at time t, L_{it-1} and K_{it-1}represents labor and capital, and μ_i reflects firm-level fixed effects, and ε_{it} is an independent error term. $g(q_{it};\gamma,c)$ stands for a continuous transition function bounded between 0 and 1. q_{it} is the transition variable (either capital or labor), and c is the threshold value. This model can identify heterogeneity in two extreme values of g-function (i.e. $g = 0$ or 1): when $g = 0$, the coefficient of $(L_{it-1}, K_{it-1})'$ is $(\beta_1, \beta_2)'$, and when $g = 1$, the coefficient change to $(\beta_1 + \beta_3, \beta_2 + \beta_4)'$. The slope parameter γ determines the smoothness of transition function: when $\gamma = 0$, the model collapses into a homogeneous or linear regression model, and when $\gamma = \infty$, the model transition function switching back and forth between zero and one, because g-function becomes an indicator function $I(q_{it} > c)$.

4.2 Synthetic Control Method

One limitation of the DID-based methods is that they assume parallel trends in the outcome variable in the treated and control groups in the absence of treatment, an assumption that cannot be tested in empirical studies. The SCM does not require that assumption and thus has been widely used for treatment effect estimation in recent years. For example, Abadie et al. (2010) used SCM to study the effect of California's tobacco control program, and the control unit for California is synthesized using a weighted average of 38 US States. Roesel (2017) used SCM to identify the effect of mergers of large local governments in Germany (districts) on public expenditures. Birdsall (2017) used SCM to estimate

the effect of managerial Discretion Under Performance Management.

If one wants to estimate the effect of IKEA entry at municipality level, as we do in Paper IV in the thesis, the SCM is probably a good solution, since the four IKEA stores under study are established in dissimilar regions, and should be treated as individual case studies rather than similar events. The SCM estimates the intervention effect by measuring the difference in the outcome variable in the treated unit and a synthetic control unit. Other similar, unexposed units are used to synthesize a control unit to mimic the evolution of the outcome variable in the exposed unit, had it not been subject to exposure. The SCM constructs a counterfactual unit by weighing control units, such that the weighted average of outcomes and relevant covariates in the pre-intervention period is similar to the factual ones of the treated unit.

The counterfactual control unit is constructed as $\hat{\eta}_{0t} = \sum_{j=1}^{J} \hat{w}_j y_{jt}$, where y_{jt} denotes the observed outcome variable for unit j at time t, where $j = 0$ refers to the treated unit, and $j = 1, \dots, J$ refers to the J control units in the "donor pool". Furthermore, $t = 1, \dots, T_0, \dots, T$ is the time period during which the variables are observed, and T_0 is the time point the intervention occurs. η_{0t} denotes the outcome of the variable of interest, had the treated unit not been treated (i.e. the counterfactual). The estimator of $\hat{\eta}_{0t}$ is the outcome for the synthetic control unit, being the weighted average of control units, supposedly mimicking the evolution of the treated unit in the absence of treatment. The constraints of the weight vector is $\sum_{j=1}^{J} \hat{w}_j = 1$, and $0 \leq \hat{w}_j \leq 1$.

It is stipulated that the intervention bears effect at the time point of its introduction and afterwards, and, therefore, for $t < T_0$, it holds that $y_{0t} - \eta_{0t} = 0$. While, for $t \geq T_0$, the parameter of intervention effect equals $\alpha_t = y_{0t} - \eta_{0t}$. The estimator of the intervention effect at time t is in Eq. (4).

$$\hat{\alpha}_t = y_{0t} - \hat{\eta}_{0t} = y_{0t} - \sum_{j=1}^{J} \hat{w}_j y_{jt} \quad (4)$$

where $\hat{w}_j$ is obtained by optimizing an objective function, which minimizes the discrepancy between the observed treated unit and the SC unit before the intervention, with regard to the outcome variable and covariates. Abadie et al. (2010) also considered covariates in the model, where the outcome variable is explained by covariates using a linear factor model. The linear model for control units is expressed in Eq. (5), and the counterfactual of the treated unit in Eq. (6). δ_t is an unknown common factor with constant factor loadings across unit, z is a $r \times 1$ vector of r observed covariates, μ is a $F \times 1$ vector of unobserved covariates, and ϵ_{jt} and ϵ_{0t} are error terms with common variance σ^2.

$$y_{jt} = \delta_t + \theta_t z_j + \lambda_t \mu_j + \epsilon_{jt}, \quad (5)$$

$$\eta_{0t} = \delta_t + \theta_t z_0 + \lambda_t \mu_0 + \epsilon_{0t},$$

4.3 Sentiment prediction model

To understand people's opinion regarding IKEA entry into their local communities, Twitter texts

are a data source where users express their opinions concerning this event in public. The study object is local tweets containing "IKEA" posted during or around the IKEA-entry date. The goal is to compute the sentiment polarities (positive, negative or neutral) for each tweet relating to IKEA. There are two computing approaches: lexicon-based and machine-learning methods.

For English tweets, lexicon-based approaches are applied. To compute one sentence's sentiment, one can retrieve each word's associated sentiment polarity from the existing sentiment dictionary, then sum up the sentiment of all the words by incorporating negation and intensification (Thelwall, et al., 2012; Maite et al., 2011).

For Swedish tweets, the sentiment of each word cannot be retrieved, due to lack of a Swedish sentiment dictionary. But their sentiment can be learned from data which provides matchup between all the words in one sentence and their corresponding sentiment. To build the relationship between words and intrinsic sentiment, several machine-learning methods are fitted and compared. The optimal model is the Elastic Net Penalty method, which outperforms models Logistic Regression, Naïve Bayes, Support Vector Machines, Neural Network and Random Forest.

5. Summary of the papers

Paper I in the thesis is titled ***"Big-box retail entry in urban and rural areas: Are there productivity spill-overs to incumbent retailers?"***. In this paper, the impact of IKEA entry on the productivity of incumbent retailers is estimated using a translog production function difference-in-difference regression. Previous studies of big-box retail entry on local communities use similar models, and the results from Paper I in the thesis can then be seen as a baseline measure using traditional methods for estimating the impact of IKEA entry on productivity. The results show that IKEA entry increases incumbent firms' productivity in two smaller rural entry regions, by 44% in Haparanda, and by 13% in Kalmar, but no effects were found in Karlstad, or the urban entry in Gothenburg.

From Paper I we know that there are spillover effects to incumbent retailers, at least when entry occurs in smaller markets. Thus, Paper II examines two additional factors that could affect how IKEA entry affects incumbent retailers: the retail firm's industry and distance. The paper is titled "***Comparison and one-stop shopping after big-box retail entry: A spatial difference-in-difference analysis***". It adds to the findings from Paper I, by also considering that firms selling substitute or complement goods to those found in IKEA could be differently affected by IKEA entry, as could firms at different distances from the new IKEA. A Spatial DID model is applied because it includes the possibility of firms' sales being determined by a process with spatially interactive responses, and the results show that ignoring such issues can underestimate the IKEA-entry effect by as much as 3%. The results also show that for complementary goods retailers in Haparanda and Kalmar, productivity increased by 35% and 18%, respectively, while no effects were found in Karlstad or Gothenburg. In summary, retailers in smaller entry regions selling complementary goods and located in the vicinity of the new IKEA stores benefit

from IKEA entry.

Papers I and II in the thesis indicate that the effects of IKEA entry are delimited to smaller entry regions, while no effects have been found in any of the estimated models for the entry in Gothenburg, and similar findings have also been presented for Swedish food retailing (Maican and Orth, 2012). Regarding entry of WalMart in the US, Artz and Stone (2006), also report that the effect appears to be weaker in metropolitan areas, compared to smaller and less densely-populated regions.

That entry by IKEA does not seem to have any impact in urban areas is still somewhat surprising. After all, a new IKEA store attracts consumers spending approximately 800 million SEK in the region entered by IKEA, and this should also affect nearby retailers. One possible reason is that IKEA effects are heterogeneous in different subgroups, making the effect difficult to detect using traditional econometric methods, such as difference-in-difference estimation on the full sample of firms. Paper III ***"Agglomeration economies in urban retailing: Are there productivity spillovers when big-box retailers enter urban markets?"*** uses panel smooth transition regression to divide the dataset into large and small incumbent retailers, and finds that the IKEA entry in Gothenburg has a positive impact, but only for relatively small retailers. The results show that for incumbent retailers in Gothenburg, with a capital stock less than approximately 1 500 000 SEK, prior to IKEA entry, productivity increased by approximately 9% due to IKEA entry.

All three papers above find positive IKEA-entry effects on incumbent retailers. But the analysis in these papers is at firm level, and it could also be of interest to investigate the effect at municipality level, including the effect of IKEA itself. After all, the largest impact on average productivity in the retail sector in the entry regions is likely due to the big-box entrants themselves. For Wal-Mart, Basker (2007) reported that the real value-added per worker was 40% higher than that of other US general merchandise retailers, and that Wal-Mart was responsible for nearly half of the productivity growth in the US retail market between 1982 and 2002. By considering only the impact on incumbents as in Papers I, II and III, rather than the total impact of big-box entry on productivity in the entry regions, we are likely to have underestimated the impact of big-box retail entry on productivity in the entry regions, a potential caveat addressed in Paper V.

However, since the SCM is a non-parametric method, reliability and statistical inference are potentially problematic, which motivates Paper IV ***"The Power of the Synthetic Control Method"***. SCM is a new and popular method for estimating intervention effects when only one single unit has been exposed. Other similar, unexposed units are used to synthesize a control unit to mimic the evolution of the exposed unit, had it not been subject to exposure. As inference relies on only a single treated unit, the statistical inferential issue is a challenge. Unconventional model assumptions and non-parametric estimation in the SCM imply that its power must be tested when used in practical applications. Thus, we study the SCM's applicability by examining a number of features potentially yielding uncertainty in

the estimator, we discuss the rationale for statistical inference in relation to the SCM, and provide a Web-app to aid researchers in their decision of whether SCM is sufficiently powerful for a specific case study. We find that the SCM is powerful even for a small number of controls in the donor pool and a fairly short pre-intervention time period. This holds as long as a parametric specification of the intervention effect is used, the duration of the post-intervention period is reasonably long, and the fit of the synthetic control unit to the exposed unit in the pre-intervention period is good.

Applying the SCM that is developed in Paper IV, Paper V estimated the effect of IKEA entry on average productivity in durable goods retailing in the entry regions, in a paper titled **"*How Does Big-Box Entry Affect Labor Productivity in Durable Goods Retailing? A Synthetic Control Approach***". The effect of IKEA entry on labor productivity, measured as CPI-adjusted sales per employee in durable goods retailing, is estimated using data from 2001–2012, and the effects of IKEA entry are found to be consistent in direction, but larger in size (at lest in Haparanda), compared to the previous firm-level studies. The effects are larger in smaller entry regions (Haparanda 80%, Kalmar 8% and Karlstad 4%), than in larger regions, again with no statistically significant effects found for the entry in Gothenburg.

Finally, Paper VI studies the social impact of IKEA entry in a paper called ***"Twitter Sentiment Analysis of New IKEA Stores Using Machine Learning"***. Public opinion regarding IKEA-entry events could be essential for local governments in informing themselves of public attitude. Three specific research questions regarding IKEA entry are addressed: how much attention does a new IKEA entry attract?, how much positive and negative emotions are aroused by IKEA entry?; and, finally, what do residents in the entry areas talk about when IKEA enters the region? Data is sourced from the Twitter platform, where users express their opinions and sentiments in public. During IKEA opening periods, residents in the IKEA-entry areas post tweets related to the term "IKEA", and these tweets are collected for sentiment analysis in this study using data crawling.

To compute sentiment polarity of the tweets, we use a lexicon-based approach for English tweets, while for Swedish tweets, we use machine-learning methods. The lexicon-based approach retrieves every word's sentiment polarity from an English sentiment dictionary, and then adds them up (by incorporating negation and intensification words) as polarity of the current sentence. However, the lack of a Swedish sentiment dictionary requires new methods for predicting sentiments for Swedish tweets. We employ several machine-learning methods, trained on a large amount of Swedish tweets with labeled sentiments, to fit such a prediction model. The best method, the elastic net method, is then applied for analyzing the IKEA tweets. The results of the three research questions are: there is a significant increase in IKEA tweets at the time of entry, indicating that IKEA entry receives much attention; the sentiments expressed in the tweets are mostly positive; and the IKEA-related topics in the tweets are heterogeneous, with respect to the location of the new IKEA stores (i.e. from lifestyle and room rennovation, to problems concerning traffic and the environment).

6. Discussion

This dissertation has studied how IKEA entry affects local economic and social development, and the results are informative for policymakers. Governments should be aware that the big-box (IKEA)entry effects depend on the size of the existing retail market, the type of existing retail industry, the size of existing retailers in the entry region, and that the effects dissipate with distance from the entry location. This dissertation elaborately answered related research questions listed in Section 1.

The IKEA-entry effects for regional labor productivity in durable goods retailing, on average, are positive in three rural regions, when using the SCM to analyze these entries. The size of the effects is inversely proportional to local market size, and average productivity in Haparanda, Kalmar and Karlstad increased by 80%, 8% and 4%, respectively, whereas no effects are found for the urban entry in Gothenburg.

IKEA-spillover effects on incumbent retailers are estimated in three papers investigating different aspects potentially affecting the size of the effects of IKEA entry on productivity. Results using traditional estimation methods show that the average productivity of incumbent retailers in Haparanda and Kalmar increased by 44% and 13%, respectively, when IKEA entered these local markets, whereas no effects were found in Karlstad or Gothenburg. Secondly, the study progressed by then considering the incumbent retailers type of industry and distance to IKEA, using a spatial DID model. The results from this model show that the effects of IKEA entry dissipates at different distances in the four entry municipalities, and that the IKEA entry in Haparanda increased productivity in incumbent one-stop shopping (complementary goods) retailers by 35%, while the entry in Kalmar increased productivity by 18%. For the larger entry municipalities of Karlstad and Gothenburg, no significant effects were found for one-stop shopping retailers, and for comparison shopping (substitute goods) retailers, no statistically significant effects were found in any of the entry regions.

It is surprising that the urban IKEA entry in Gothenburg has no effects in these studies, after all, IKEA is a large entrant even in Gothenburg, and it should have an effect on incumbent retailers. In fact, an IKEA effect does exist also in Gothenburg, but since it is limited to smaller incumbent retailers it is difficult to detect using traditional DID models. In the small-firm subgroup consisting of incumbent retailers with a capital stock of less than approximately 1 500 000 SEK (138 000 EURO), the IKEA entry in Gothenburg increases productivity by 9%, but no positive effects were found for the large-firm group in Gothenburg.

In summary, IKEA-entry spillover effects are heterogeneous, depending on firms' size, industry, distance, and regional market size. Thus, political decision-makers should comprehensively consider the above aspects, as many Swedish municipalities spend large amounts of money in attracting IKEA entry into their own municipality.

In addition to the papers in the thesis, two web applications have also been developed to facilitate

SCM implementation in empirical studies. This method has little power in detecting an intervention effect in some cases, so the first web-app can determine its odds in advance of planning a study and collecting data. The second web-app can estimate the intervention effect with both parametric & non-parametric specifications, and provide statistical inference, not only through placebo testing but also using bootstrapped confidence intervals for inference. Users need only to upload their data and set some parameters.

By contributing both to empirical studies and method development in the field of big-box retail analysis, this dissertation lays a foundation for related research. Empirically, the analysis in the thesis discovered that IKEA-entry effects were heterogeneous, and methodologically, the analysis improves on traditional approaches using more flexible and reliable methods.

One limitation of this study is that all results pertain specifically to IKEA entry, while the impact of big-box retail entry, in general, could also be expected to be heterogeneous with respect to which big-box retailer enters the market. To find the general patterns of big-box entry effects on the communities where they enter, this requires additional empirical work focused on other big-brand retailers in the future. The merit of this study for other big-box researchers is that it proposes a comprehensive analysis framework and powerful empirical approaches and tools within this field.

Future work could also study the interaction between economic effects and social effects of big-box entry. The hypothesis to study would be that the more positive public opinion is, regarding the big-box entry, the larger the positive economic effects it will generate. The aim is to fit a model for the quantitative relationship between the economic and social effects, and the relationship between consumer sentiment and economic variables may also take into account the characteristics of the economic environment, such as the degree of competition in the local markets, the flexibility of the economy (especially the labor market), the nature of the welfare state, and the strength of political and economic institutions (Acemoglu et al., 2002). If one could create a well-functioning model, local governments could then use public sentiment measurements as an indicator for potential economic growth by applying the fitted model. The results are then not only important for academia, but also for local political decision-makers in deciding whether to allow entry by big-box retailers in their local communities.

References

Abadie, A., Diamond, A., and Hainmueller, J., (2010), Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program, *Journal of the American Statistical Association*, 105(490): 493-505.

Acemoglu, D., Johnson, S., Robinson J., and Thaicharoen, Y., (2002), Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth, NBER Working Paper 9124.

Aldieri, L., Cincera, M., (2009), Geographic and technological R&D spillovers within the triad: Micro evidence from US patents. *The Journal of Technology Transfer*. 34, 196-211.

Arrow, K., (1962), The economic implications of learning by doing. *Review of Economic Studies* 29, 155–172.

Artz, G. M., and Stone K., E. (2006), "Analyzing the Impact of Wal-Mart Supercenters on Local Food Store Sales", *American Journal of Agricultural Economics*, 88(5), 1296-1303. doi: 10.1111/j.1467-8276.2006.00948.x.

Audretsch, D. B., Feldmann, M. P., (2004), Knowledge Spillovers and the Geography of Innovation, *Handbook of Regional and Urban Economics*. 4, 2713-2739

Athey, S., and Imbens G. W., (2016), The State of Applied Econometrics: Causality and Policy Evaluation, *The Journal of Economic Perspectives*, 31(2): 3-32

Birdsall, C., (2017), The synthetic control method for comparative case studies: An application estimating the effect of managerial discretion under performance management, *International Public Management Journal*, 20 (1): 49-77.

Bailey, T.C., Gatrell, A.C., (1995), *Interactive spatial data analysis*.

Basker, E., (2005), Job creation or destruction? Labor market effects of Wal-Mart expansion. Rev. *Econ. Stat*. 87, 174–183.

Basker, E., (2007), The causes and consequences of Wal-Mart's growth. J. Econ. *Perspec*. 21, 177–198.

Cardamone P., (2014), R&D, spatial proximity and productivity at firm level: evidence from Italy, MPRA Paper No. 57149, University Library of Munich, Germany.

Curtin, R. T., (1992), The Index of Consumer Sentiment at Forty, *American Enterprise* 3:3, 18-24.

Carroll, D. C., (1994), Does Consumer Sentiment Forecast Household Spending? If So, Why? *American Economic Review*, 84 (5): 1397-1408.

Daunfeldt, S.-O., Mihaescu, O., Nilsson, H., and Rudholm, N. (2015), What happens when IKEA comes to town? *Regional Studies*, 51(2): 313-323, DOI: 10.1080/00343404.2015.1100287

Davidson S., M., and Rummel A., (2000), Retail changes associated with Wal-Mart's entry into Maine, *International Journal of Retail & Distribution Management*, 28:4, 162-169, doi: 10.1108 /09590550010319913.

Delgado, M. S., Florax, R. J. G. M., (2015), Difference-in-differences techniques for spatial data: local autocorrelation and spatial interaction. *Econ. Lett*. 137, 124–126.

Dhawan, R., (2001), Firm size and productivity differential: theory and evidence from a panel of US firms, *Journal of Economic Behavior & Organization*, 44: 269–293

Eurostat, (2007), *Handbook on Data Quality Assessment Methods and Tools*

Evan, D. S., (1987), The Relationship Between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries, *Journal of Industrial Economics*, 35 (4): 567-581.

Foster, L., Haltiwanger, J., Krizan, C.J., (2006), Market selection, reallocation, and restructuring in the U.S. retail trade sector in the 1990s, Rev. *Econ. Stat*. 88, 748–758.

Glaeser, E. L. (1999), Learning in cities, *Journal of Urban Economics*, 46 (2): 254-277.

González, A., Teräsvirta, T., van Dijk, D., and Yang, Y., (2005), Panel Smooth Transition Regression Models, SSE/EFI Working Paper Series in Economics and Finance 604, Stockholm School of Economics.

Grossman, G. M. and Helpman, E. (1991), *Innovation and Growth in the Global Economy*, Cambridge, MIT Press.

Han, M., Mihaescu, O., Li, Y., and Rudholm, N. (2018), Comparison and one-stop shopping after big-box retail entry: A spatial difference-in-difference analysis. *Journal of Retailing and Consumer Services*, 40: 175-187.

Hernandez, T., (2003), The impact of big-box internationalization on a national market: a case study of Home Depot Inc. in Canada. Int. Rev. Retail Distrib. Consum. Res. 13, 77–98.

Hicks, M.J., (2007), Job turnover and wages in the retail sector: the influence of Wal-Mart. J. Priv. Enterp. 22, 137–160.

Hicks, M.J., (2008), Estimating Wal-Mart's impact in Maryland: a test of identification strategies and endogeneity tests. East. Econ. J. 34:56–73.

Hotelling, H., (1929), Stability in Competition, *Economic Journal*, 39 (153): 41–57.

Håkansson, J., Li, Y., Mihaescu, O., Rudholm, N., (2016), Big-box retail entry in urban and rural areas: Are there productivity spillovers to incumbent retailers? HUI Working Papers No. 118.

Jia, P., (2008), What happens when Wal-Mart comes to town: an empirical analysis of the discount retailing industry. *Econometrica* 76:1263–1316.

Jones, K., Doucet, M., (2000), Big-box retailing and the urban retail structure: the case of the Toronto area. J. Retail. Consum. Serv. 7, 233–247.

Koo, J., (2005), Technology Spillovers, Agglomeration, and Regional Economic Development. *Journal of Planning Literature*. 20, 99-115.

Lucas, R., (1988), On the Mechanics of Economic Development, *Journal of Monetary Economics*, 22, 3-42.

Maican, F., Orth, M., (2012), Productivity dynamics and the role of big-box entrants in retailing. Research Institute of Industrial Economics Working Paper No. 898. Research Institute of Industrial Economics, Stockholm.

Marshall, A., (1890), *Principles of economics*. McMillan and Co., London.

McCann, P., (2001), *Urban and regional economics*. Oxford University Press, Oxford.

Merriman, D., Persky, J., Davis, J., and Baiman R. (2012), The Impact of an Urban WalMart Store on Area Businesses: *The Chicago Case, Economic Development Quarterly*, 26(4), 321-333. doi: 10.1177/0891242412457985.

Neumark, D., Zhang, J., Cicarella, S., (2008), The effect of Wal-Mart on local labor markets. J. Urban Econ. 63, 405–430.

Orlando, M., (2000), On the importance of geographic and technological proximity for R&D spillovers: An empirical investigation. Federal Reserve Bank of Kansas City, RWP 00-02.

Thelwall, M., Buckley, K., and Paltoglou, G., (2012), Sentiment strength detection for the social web, *Journal of the Association for Information Science and Technology*, 63(1): 163-173.

Romer, P., (1986), Increasing returns and long-run growth. *Journal of Political Economy* 94, 1002-1037.

Roesel, F., (2017), Do mergers of large local governments reduce expenditures? – Evidence from Germany using the synthetic control method, *European Journal of Political Economy*, 50, 22–36.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., and Stede, M., (2011), *Association for Computational Linguistics*, 37:2.

Weber, A., (1929), *Theory of the location of industries*, Chicago: University of Chicago Press.