Journal of Retailing and Consumer Services 45 (2018) 33–41

Contents lists available at ScienceDirect

Journal of Retailing and Consumer Services

journal homepage: www.elsevier.com/locate/jretconser

Identifying the geography of online shopping adoption in Belgium

Joris Beckers*, Ivan Cárdenas, Ann Verhetsel

Department of Transport and Regional Economics, University of Antwerp, Prinsstraat 13, 2000 Antwerpen, Belgium

ARTICLE INFO

Keywords:
Geography
E-commerce
City logistics

ABSTRACT

The widespread adoption of the internet as retail channel is impacting a range of stakeholders. Retailers are expected to sell online, logistics operators are required to reconfigure their supply chain and public authorities try to keep local retail competitive while simultaneously attempt to manage the increase in freight transport. Within this context, a growing body of research is studying the socio-economic profile of the online shopper and the spatial variation in the demand for B2C goods. Yet, as can be expected for a relatively new evolution, little consensus exist. Therefore, in this paper, with data from the national retail federation on online shopping behaviour, we add to this growing field by first analysing the relation between socio-economic characteristics and the willingness to shop online. By mapping these characteristics, we then construct the geography of online shopping adoption in Belgium. Finally, we assess the impacts of this specific geography for the stakeholders that are adapting to this new reality. We conclude firstly that the well-educated man in his thirties with a well-paid job has the highest probability to shop online, independent of the level of urbanisation of the area he resides. Secondly, we predict over- and underestimations of the potential online buyers of up to 50% when assuming a homogeneous e-commerce penetration, especially in poorer urban areas. This implies a serious negligence for e-commerce practitioners and academics when ignoring the specific geography of the online shopping adoption.

1. Introduction

Twenty-three years after the founding of Amazon, business-to-consumer electronic commerce (B2C e-commerce) keeps rising. Over the last six years, Europe experienced an average annual increase in online sales turnover of 16% (E-commerce Europe, 2016b). Despite persistent growth rates around the EU however, levels of e-commerce adoption vary greatly among different member states. While only 16% of the population of Western Europe never shopped online in 2016, this number rises to 40% for Eastern Europe (Eurostat, 2017).

The popularity of internet shopping provides retailers the opportunity of opening an additional distribution channel. At the same moment however it offers customers a wider market to choose from, hence increasing local competition and consumer power (Boschma and Weltevreden, 2008; Weltevreden, 2007). In response, retailers are required to be present online and ideally consider effective integration of their online and offline channels (Rimmer and Kam, 2018). In addition, in order to stand out brand creation and marketing have now become even more important, both for traditional and virtual merchants (Doherty and Ellis-Chadwick, 2010).

In parallel, the demand for home delivery services is growing at similar pace, resulting in a fragmentation of the goods distribution flows to the extent of a single item per delivery (Gevaers et al., 2011; Hesse, 2002). In consequence, recent literature estimates that up to 75% of the delivery costs originate from the last part of the distribution chain, i.e. the last mile (Gevaers et al., 2014). This not only puts pressure on traditional delivery models but threatens the livability of some urban areas, due to the increase in light good vehicles delivering parcels associated with the growth in online sales (Anderson and Leinbach, 2007; Browne, 2001; Cherrett et al., 2012).

Resultantly, local administrators are facing pressure to accommodate sustainable growth in online shipments while simultaneously attempting to prevent the disintegration of retail areas within their jurisdiction (Browne and Allen, 1999). The former has resulted in an increased awareness for freight planning within cities' administrations, a topic which up to now has always been overshadowed by the focus on passenger transport (Kiba-Janiak, 2017; Lindholm and Behrends, 2012). Furthermore, attempts are taken to improve the shopping experience within retail areas to limit the substitution of traditional purchases by online orders. Examples include investments in cycle hubs to store purchased goods or *facility areas* with pick-up points in main shopping streets as to attract online shoppers in nearby stores (Gómez

* Corresponding author.
E-mail addresses: joris.beckers@uantwerpen.be (J. Beckers), ivandario.cardenasbarbosa@uantwerpen.be (I. Cárdenas), ann.verhetsel@uantwerpen.be (A. Verhetsel).

https://doi.org/10.1016/j.jretconser.2018.08.006
Received 27 April 2018; Received in revised form 7 August 2018; Accepted 10 August 2018
0969-6989/

et al., 2017).

As has been proven in multiple studies over the past two decades, the continuing relevance of geography for all stakeholders involved should not be underestimated (Anderson et al., 2003; Boschma and Weltevreden, 2008; Couclelis, 2004). Besides international differences, socio-economic and geographical factors do differentiate the online shopping behaviour within a single country (Clarke et al., 2015; Farag et al., 2006b). Whether these regional differences in online shopping are due to accessibility issues or varying openness to innovation remains under study (Motte-baumvol et al., 2017).

While there is agreement that knowledge on regional variations in online shopping holds great marketing value, these also imply that the demand for e-commerce goods can change from one neighbourhood to another. Given the forecast of continuing e-commerce growth, knowledge concerning the e-shopper's profile and his related delivery preferences or failed delivery rates may help the courier, express and parcel (CEP) industry to balance cost efficiency and environmental sustainability in the last mile of e-commerce. In addition, it may provide local authorities the possibility to more efficiently implement sustainable urban logistics planning initiatives or e-resilience measures. Finally, it can support retailers to further improve their offline and online integration.

Given the value of understanding the geography of the online shopper, we analyse in this paper the spatial pattern of e-commerce demand using data on Belgium. With nearly 60% of the population buying online (Eurostat, 2017) and growth rates of over 10% for the B2C online turnover (E-commerce Europe, 2016a), currently observed impacts will be fast growing in the coming years. As a result, there is an urge for more efficient and environmental friendly logistic solutions, for conscious local authorities and for well-informed retailers that take into account the geography of the demand of e-commerce.

This motivation leads to the construction of two research questions we target here. After the introduction of our methodology, we test in the first part whether previous findings of socio-economic factors impacting online shopping behaviour currently hold in Belgium. In other words: who is the Belgian e-shopper? Second, we advance the current literature on the topic, knowing the characteristics of the Belgian e-shopper and given the socio-economic profiles of each Belgian neighbourhood, by answering the question: where does the e-shopper live? The answers to these research questions should help to identify the extent to which the specific geography of e-commerce matters for the stakeholders involved. We conclude our work in the last section and suggest paths for further research.

2. Literature review

Already at the end of the nineties, at the height of the dot-com bubble, researchers in the field of marketing were studying the relation between socio-economic characteristics and the demand for e-commerce. The main goal of these studies was to predict the chances of online shopping based on behavioural characteristics. Early publications identified the better-educated males from 26 to 35 year old with high incomes as the earliest cybershoppers (Donthu and Garcia, 1999; Kau et al., 2003; Sim and Koi, 2002; Vrechopoulos et al., 2001). These findings were later confirmed by similar studies in the field of economic geography. First by Farag et al. (2006a, 2006b) in their study of e-shopping in the Netherlands, who found a nonlinear relationship between age and buying online: up to the age of 33, the likelihood of buying increases after which it decreases again. Nine years later Clarke et al. (2015) came to similar conclusions for e-shoppers in the UK, although the maximum frequency of online shopping fell in the age category 35–39. Like the previous studies, also in the UK a strong positive correlation seemed to exist between household income and online shopping frequency.

Later, researchers adopted spatial components like urban-rural differences in similar studies starting with Anderson et al. (2003). He formulated two possible but opposite hypotheses concerning the spatial diffusion of e-commerce adoption. First, the *efficiency* hypothesis states a fast penetration of online shopping can be expected in more remote areas where e-commerce improves retail accessibility. Contrary the *innovation-diffusion* hypothesis expects early e-shopping to be limited to urban areas because new technologies are assumed to start in these centres of innovation, after which they diffuse to other regions.

Despite their contrasting nature, both hypotheses proved not to be mutually exclusive. Studies in various countries indeed find shoppers in metropolitan areas to be more prone to use the online channel and also observe the diffusion of online shopping to more rural areas over time, supporting Anderson's *innovation-diffusion* (Clarke et al., 2015; Farag et al., 2006a, 2006b; Kirby-Hawkins et al., 2018; Motte-baumvol et al., 2017; Zhou and Wang, 2014). However, others did find proof for the *efficiency* hypothesis as well, with higher frequencies for the online shoppers in less accessible areas and the identification of accessibility issues as a major driver for e-commerce (Cao et al., 2013; Farag et al., 2006b; Kirby-Hawkins et al., 2018; Motte-baumvol et al., 2017)

The work on the socio-economic and spatial characteristics of the online shopper implies that the e-commerce demand is not uniformly distributed over the population. As mentioned before, this has been acknowledged by various authors already. Examples include amongst others the assessment of infrastructure decisions in an e-commerce context. As such, in an attempt to improve the efficiency of the last mile of urban goods flows, Ducret et al. (2016) use population density and income figures to list recommendations for logistics infrastructure. This way the authors highlight the value of spatial data for urban freight modelling. In other works authors adopt a retailer perspective and use the socio demographic aspects of online revenue data to asses online and offline strategies (Birkin et al., 2017; Kirby-Hawkins et al., 2018).

The wider city logistics literature on the contrary, concerned with the trade-off between efficient and sustainable distribution in and around the city has paid little to none attention to the importance of geographical variables when studying (the distribution of) e-commerce. With the exception of a small set of works based on empirical data (e.g. Ducret et al., 2016; Weltevreden and van Rietbergen, 2007; Zhou and Wang, 2014), the majority of the B2C city logistics literature assumes the demand is uniformly distributed over the population. Nonetheless, the initial distribution of this demand is the starting point for most of the quantitative analyses, like the assessment of the potential of green transportation vehicles or the reduction of social costs through collaboration among logistics players (Arvidsson and Pazirandeh, 2017; Gonzalez-Feliu et al., 2012; Lin et al., 2016; McLeod et al., 2006; Park et al., 2016). This may imply significant differences, especially when working with detailed data.

In this paper, we first test whether previous findings concerning the role of socio-economic and spatial variables in the variability of e-commerce demand also hold in Belgium. This is done by identifying the relevant explaining variables of the probability a Belgian shops online. In a second step, we will predict for each spatial unit the number of expected online purchases. Finally, we confront this modelled demand with the population distribution to better quantify the variation in e-commerce demand. We then conclude by assessing the value of knowing the spatial variation for the different stakeholders involved.

3. Data and methodology

We make use of the *E-commerce in Belgium 2016* questionnaire organised by the Belgian retail federation Comeos to identify the online shopper. This online survey was conducted from the beginning of March to the end of April 2016 and polls the online shopping frequency of over 1500 Belgians for different categories (e.g. toys, food, …). The respondents were selected representatively according to the age, gender and regional quota of the country from a database of over 200.000 regular participants (www.futuretalkers.com). Although participation is to be decided by the respondent, response behaviour and time was

Table 1
Variable levels in the survey used in the statistical analysis. The representativeness of the survey is indicated between brackets: (% of respondents in that level | % of Belgians in that level).

	Gender	Number of children	Education	Age	Household income (net €/month)
L0 (reference)	Female *(51|51)*	0 *(71|65)*	Lower high school *(15|40)*	18–29 *(19|16)*	< 500 *(2|10)*
L1	Male *(49|49)*	1–2 *(25|28)*	High school *(62|32)*	30–39 *(20|17)*	500–1249 *(19|17)*
L2		3–4 *(4|6)*	Higher education *(20|27)*	40–49 *(22|19)*	1250–1749 *(30|17)*
L3		4+ *(0.2|0.1)*	Post-master *(2|0.6)*	50–59 *(22|17)*	1750–2499 *(26|19)*
L4				60–69 *(17|13)*	2499–3249 *(14|12)*
L5				70+ *(0.7|17)*	3250–4249 *(7|9)*
L6					> 4250 *(2|15)*

levelled out due to tracking of previous online polling behaviour.

In addition to a range of e-commerce questions, the survey collected the socio-economic variables listed in Table 1. Further, we are able to include the morphology (urban/suburban/rural) of the zip code where the individual resides. The morphology is identified through the work of Luyten and Van Hecke (2007) and serves to test the hypothesis proposed by Anderson et al. (2003) on online differences between urban and rural areas. Despite their careful delineation, one has to remain prudent when comparing this morphological subdivision with other countries. Belgium, and especially its northern part, has a particular nebular (sub)urbanized landscape (Van Meeteren et al., 2016). This results from a continuous fragmentation of the countryside by urbanisation related processes (Antrop, 2004). As a consequence, rurality in Belgium, with the exception of some areas in the south of the country, is not as remote as it can be in other countries.

The sample of respondents ranges from 18 to 70-year-olds, has an equal gender balance and is representative of Belgium's language structure (60/40 ratio Dutch/French speaking). However, there is an overrepresentation of childless households, individuals holding only a high school diploma and mid-level incomes (Table 1). The lowest educated, the eldest and the lowest income households, on the contrary, are underrepresented in the sample. General shopping characteristics of different socio-demographic groups can be found in the yearly e-commerce report of the retail federation (Comeos, 2016). However, a comprehensive analysis that crosschecks the entire sample is missing up to now. In this paper, we provide such an analysis by looking for the socio-economic characteristics that have a significant influence on the probability that one shops online.

To answer our first research question ‘who is the online shopper?’, we make use of the questioned online shopping frequency. In the original survey, each respondent could select one of seven categories to answer the question “How frequently do you buy something via the Internet (for personal purposes)?” (i.e. never; less than once a year; every 6–12 months; every 3–6 months; every 1–3 months; monthly: weekly). Of the total sample, 42% never bought online before. Resultantly, any study analysing shopping frequency or spending would have to work with less than 1000 data points. Since this would restrict the statistical significance too much, we chose to calculate the significance of the influence that different socio-economic characteristics have on the probability one shopped at least once online during the last twelve months. Since the dependent variable becomes dichotomous (Yes/No), the use of a logistic regression is advised. To calculate a binary prediction from any type of predicting variable, the logistic regression models the probability p of the occurrence of an event through the logarithm of the odds ratio in a logit function:

$$logit\,(p) = ln\frac{p}{1-p} \tag{1}$$

This logit-transformed probability p is a continuous function which can be used to predict parameters β for predictors x_1,x_2... and to test their significance via the following linear regression:

$$logit\,(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... \tag{2}$$

The odds one bought online at least once during the last twelve months then becomes:

$$\frac{p}{1-p} = e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ...} \tag{3}$$

Assume x_1 variable 1 in Table 1, Eq. (3) implies the odds a male bought online over the odds a female bought online, increases with e^{β_1}, keeping all other variables constant.

For studying our second research question ‘where lives the online shopper?’, we map the e-commerce demand by estimating the potential e-shoppers in each neighbourhood (census statistical sector) in the country. The statistical sectors are delineated at detailed geographical level, resulting in almost 20,000 sectors averaging 1.54 km^2 (Jamagne, 2001). The geographical extrapolation of the statistical outcome of the first analysis is possible because of the socio-economic information provided by the Census 2011 and the yearly publication of various tax statistics at the same geographical level (Statistics Belgium, 2014, 2018).

Except for the tax statistics, which have to be redistributed, all variable levels are equivalent in both the survey and the statistical sector units, which facilitates the extrapolation. The tax statistics do include the median, the interquartile range and an interquartile coefficient for each statistical sector and the cumulative density function of the income distribution per 1000€ (yearly) for the whole population, which allows the calculation of the shares of each sector's population within the income classes defined in the survey. This permits the prediction of an estimated total amount of potential online shoppers in each neighbourhood through the statistical model from the first part, which will be visualised geographically.

Finally, to test the extent to which the e-commerce demand is proportional to the distribution of the population, the number of modelled online shoppers is confronted with an assumed a constant penetration rate of online shopping for each sector. The applied rate equals the average amount of buyers resulting from the model to assure the total sum of buyers remains the same during the comparison.

4. Results

4.1. Who is the online shopper?

According to the results of the Comeos survey, 57% of the Belgians at least once bought online in the period April 2015–April 2016, which is the same number as published in the independent Eurostat survey on the Digital economy and society (Eurostat, 2017). Further, similar to the findings of other academic research, the highest percentages of online shoppers are found in age groups 25 till 40 years old (Fig. 1A and B). In comparison to the work of Clarke et al. (2015) in the UK however, online shopping percentages are still lower for each category with exception of the 60–64 interval. Inter-categorical differences are lower in Belgium, shown by the gradual slope of the overall graph, implying a quite homogeneous e-shopping behaviour over different age classes.

The first run of the logistic regression includes all variables listed in Table 1 since a variance inflation test did not identify any correlation among them. The maximum likelihood test of this regression can be found in Table 2. The large decrease in residual deviance given the degrees of freedom proves the significance of the model under the χ^2

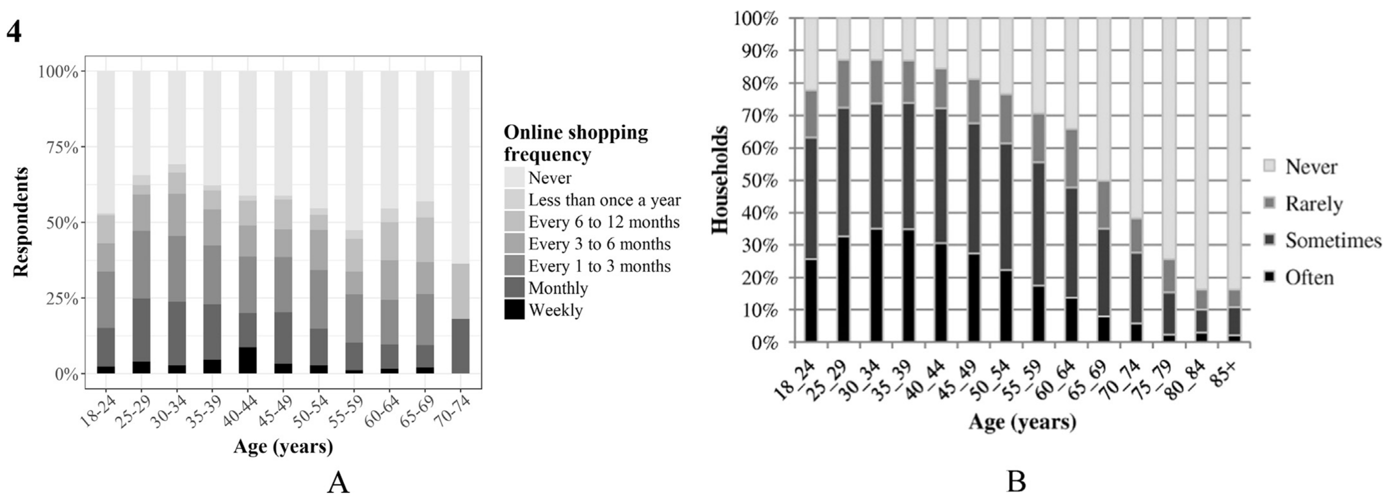


Fig. 1. Online shopping frequency over different age group in the Comeos survey (A) and the work by Clarke et al. (2015) (B). Remark when comparing the absence of respondents over 74 years in figure A.

Table 2
Maximum likelihood test.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(> Chi)	
NULL			1585	2159.7		
Income	6	87.456	1579	2072.3	< 2.2e−16	***
Gender	1	7.478	1578	2064.8	0.006244	**
Age	5	12.898	1573	2051.9	0.024350	*
Number of children	3	8.257	1568	2043.6	0.142654	
Education	3	28.146	1565	2015.5	3.385e−06	***
Morphology	2	2.124	1563	2013.4	0.35787	

Significance: '***' 0.001 '**' 0.01 '*' 0.05.

Table 3
Logit regression results.

Predictor	Dependent variable: **Bought online (Y/N)**
Income 500–1249	0.05 (0.45)
Income 1250–1749	0.68 (0.45)
Income 1750–2499	0.68 (0.45)
Income 2500–3249	1.25*** (0.46)
Income 3250–4250	1.30*** (0.50)
Income > 4250	0.95 (0.59)
Gender male	0.32*** (0.11)
Age 30–39	0.28 (0.18)
Age 40–49	0.04 (0.17)
Age 50–59	− 0.11 (0.17)
Age 60–69	0.08 (0.18)
Age = 70	− 0.98 (0.66)
High school	0.54*** (0.15)
Higher education	0.98*** (0.19)
Post-university	1.27** (0.50)
Constant	− 1.10** (0.45)
Observations	1587
Log Likelihood	− 1012.51
Akaike Inf. Crit.	2057.02

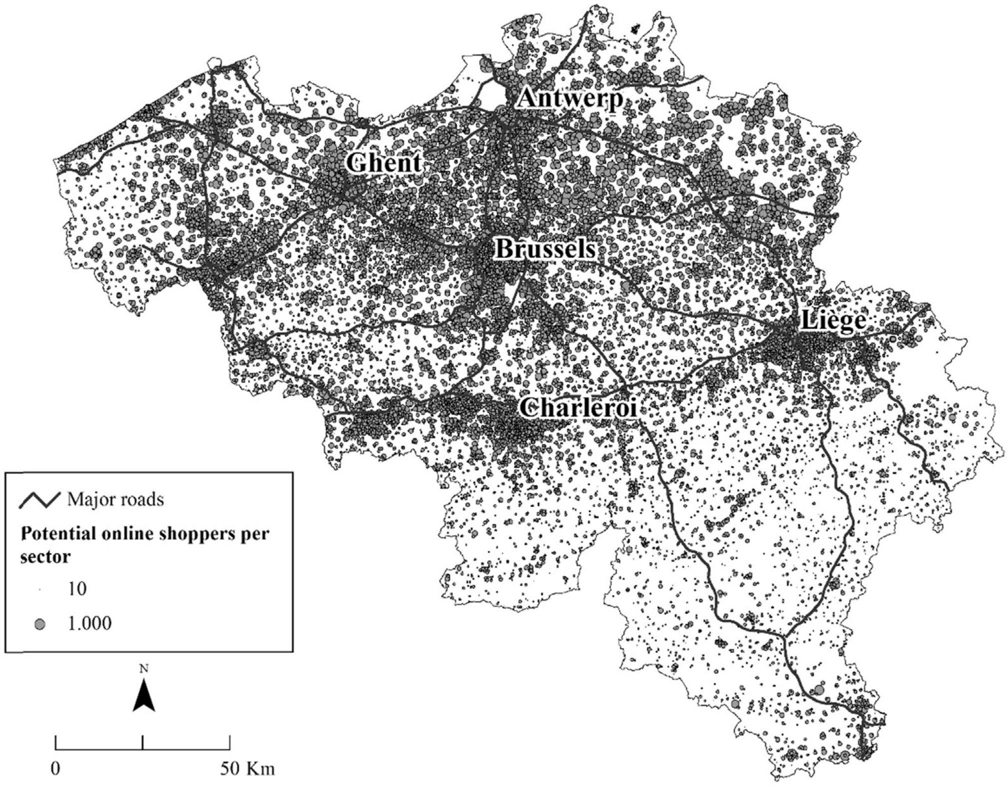
Significance: ***0.01, ** 0.05.

distribution. Further, the analysis of variance identifies only the variables income, gender, age and education as significant at the 5% level. This means that the number of children does not seem to have an effect on the probability to buy online. Next, the morphology neither has a significant influence on the online shopping behaviour. Urban, suburban and rural Belgians have similar chances of buying online, which is similar to the results found by Clarke et al. (2015). It is hard to pinpoint the exact origin of this insignificance but assuming accessibility does influence online shopping probability (cf. efficiency hypothesis), a first possible explanation can be the lack of remote rural areas like they exist in countries with really peripheral and low accessible regions. This is due to a historical lack of urban planning in Belgium, with a fragmented landscape characterised by a fairly homogeneous distribution of shop density outside the city centres as a consequence (Verhetsel et al., 2010). Resultantly, shop accessibility in many suburban and rural regions is sufficiently high to prevent a turn towards the online retail channel. A second explanation can be found in Anderson's adoption theory, assuming Belgium's progress on the e-commerce adoption timeline is quite advanced. This would mean e-commerce adoption grew beyond the early urban adopters to an almost equal share around the country. Although this conclusion seems doubtful given the country's rather low percentage of online shoppers compared to its neighbours (Eurostat, 2017, cf. Fig. 1), it can only be confirmed through a future longitudinal study.

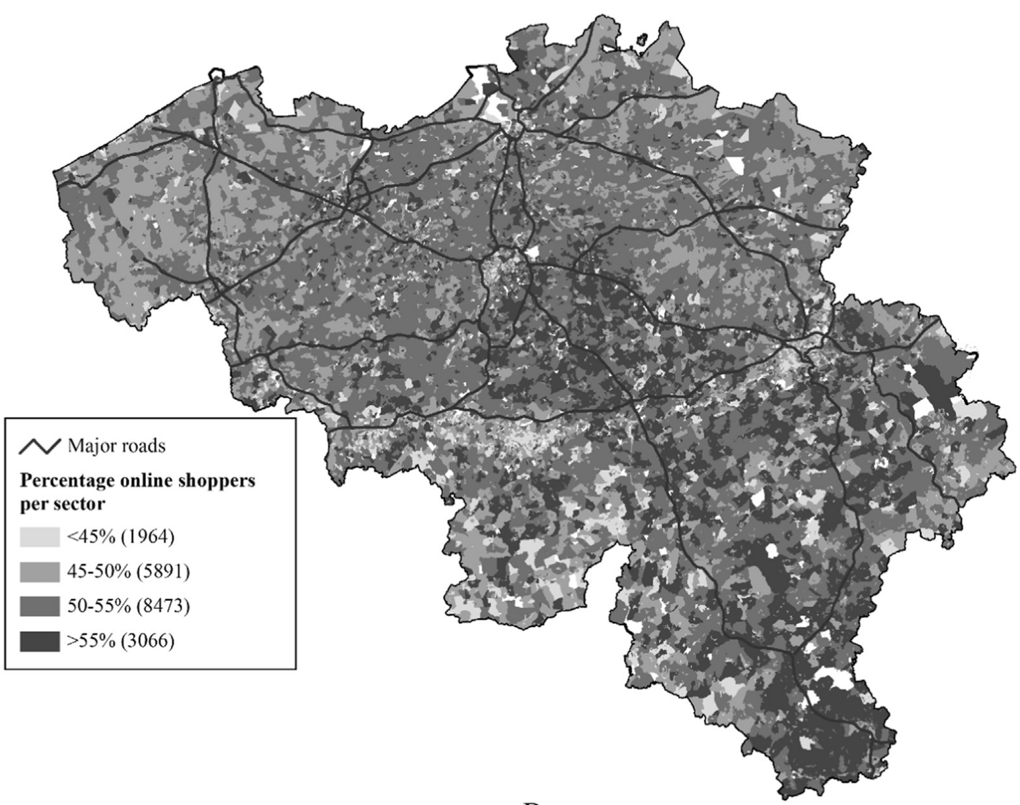
Only the significant variables from Table 2 are retained in the second logit model of which the resulting estimates with their standard error are displayed in Table 3. Concerning the individual variables, higher income categories seem to have higher probabilities to shop online compared to the base category of incomes lower than 500€. Given the close relationship between the levels, it makes sense only the higher categories display significance. Besides the income, gender is another important differentiator with the odds of men shopping online being $e^{0.32}$ or 1.4 times higher than for women. The analysis of variance pointed to a lesser significance of the variable age, which returns here with no significant differences among the various levels. Although internet users in their 30ies have higher probabilities of shopping online compared to those in their 20ies, the sign of the estimated coefficient varies among the older age groups. The inconsistency of the age variable was predictable given its distribution in Fig. 1A, i.e. with small differences between the various categories. Finally, the level of education strongly influences the online shopping odds in a positive direction: the higher the education level, the higher the probability of being an online shopper. One can conclude from the analysis that the well-educated male with good income in his 30ies has the highest odds of shopping online, which makes the Belgian online shopper similar to the profiles identified in earlier works elsewhere.

4.2. Where lives the online shopper

The final model including all variables in Table 3 considers all Belgians older than 20 years old and predicts a buying percentage of 50% when estimating the absolute number of online shoppers in each neighbourhood. This is lower than the 57% of both the original survey



A



B

Fig. 2. Extrapolation of the online buyers over Belgium. (A): Potential absolute amount of online shoppers per sector. (B): Percentage of potential online shoppers per sector.

and the Eurostat study. Although no information is present on the Eurostat data, the difference with the retail federation's survey may be explained by the uneven distribution of its target population over the variable levels (Table 1). First, since the original survey is conducted through a web platform it only reaches people with internet access. Although the internet penetration in Belgium is high (87%, Eurostat, 2017), the online panellist may be more acquitted to internet usage compared to the average Belgian and may thus have a higher probability to shop online. Further, as mentioned in the data section, we observe a considerable underrepresentation of the eldest age group, of the lowest and highest income groups, and of the lowest educational level in the survey. Except for the highest incomes, the three other classes have significantly lower probabilities for online shopping, which means their underrepresentation results in an overestimation of the total online shopping percentage in the survey. From this point of view, we may conclude that the real amount of Belgian shopping online lies somewhere between our model's estimation that 50% and the 57% derived from the survey.

The resulting geographical extrapolation is mapped in Fig. 2A. The model predicts more buyers in the northern part of the country, in and around the large cities. Further, the traditional axis of industrial cities from Charleroi to Liège appears. Next to the few remaining rural areas in the northern part of Belgium, especially the southern, rural part of the country (the Ardennes) has fewer potential online shoppers. In addition, Fig. 2B displays the percentage of online shoppers per statistical sector. This map demonstrates the interaction between the different impacting socio-economic variables which for example in Brussels results in lower percentages mainly in the low-income neighbourhoods in the west of the agglomeration. A more elaborate discussion of these results is presented in the next section.

5. Discussion

Because the logistic regression proves various socio-economic variables have a significant impact on the probability one buys online and assuming each online purchase is delivered at the shopper's residence, the demand for e-commerce deliveries is not evenly distributed over the population. For example, demand can vary drastically between two neighbourhoods with similar population sizes but with contrasting income levels. Given that in Belgium relatively more lower income neighbourhoods can be found in cities that have on average higher population densities, the urban e-commerce demand density might exhibit an inverse relation with the population density.

In other neighbourhoods the socio-economic profile may enforce the differences in population density. For example, rich and densely populated areas with a large share of middle-aged highly educated families could yield very high concentrations of potential e-commerce demand. This is observed in the wealthy neighbourhoods in and around Brussels and Antwerp. On the contrary rural areas with a poorer, less educated and elder population may show very low demand.

To assess the implication of the significance of some socio-economic parameters, we overlay the modelled e-shopper with a null model in Fig. 3. This null model is generated by assuming that 50% of the population in each statistical sector shops online, as predicted in step 2. The homogeneous distribution over the population represents the perspective of many stakeholders on the geographical distribution of the demand. Dark dots identify areas where our model predicts less than 50% of the population buying online, i.e. where traditional models overestimate the number of buyers. In Belgium, this is in northwest Brussels, around Antwerp, at the coast and in large parts of the Walloon industrial axis. While the large populations there ensure the absolute number of e-shoppers remains high (see Fig. 2A), the density of e-commerce demand will not reach the levels the population density promises due to the lower socio-economic status of its residents. In some places, we observe dissimilarities of over 50% between both models.

The many light grey dots in more suburban areas point to an underestimation of the potential number of e-shoppers in the null model. Although these dots occur in the majority of the country, the relative underestimation remains quite low with values between 10% and 25%.

Finally, in some parts of the most rural areas in the west and south of the country, we observe fewer orders in our model compared to the null model. This is also observed in Fig. 2B with lower shares for these regions. The combination of a sparse population with less than average e-shopping probabilities results in a very low e-commerce demand.

The differences between our model and the null model in Fig. 3 imply that, when distributing the e-shoppers homogeneously over the population, there will be under- and overrepresentations of the real e-commerce demand. These findings have significant value for practitioners and academics involved working with e-commerce market analyses. For example for logistics carriers concerned with the implementation of collection-and-delivery points (CDP). Originally installed to redirect not-at-home deliveries, these locations evolved to a means to consolidate shipments, hence reducing the delivery costs of the last mile (McLeod et al., 2006; Morganti et al., 2014; Weltevreden, 2008). In addition, within the current competitive logistics sector, the size of the CDP network now has become an important service factor for logistics companies. Because of this, many carriers prioritise customer service and prefer a network with equal distances to their customers, or with a certain demand within their catchment area. However, since many carriers have difficulties to separate data on business-to-business (B2B) and business-to-consumers (B2C) deliveries because they are transported and handled together, historical delivery data does not satisfy the requirements for this rather predictive analysis. Hence, carriers are obliged to turn to population statistics when implementing CDPs. Given the under- and overestimations of e-commerce demand depicted in Fig. 3 however, these statistics may result in suboptimal location choices. While ignorance of the geographical variation in such decision models will not cause the bankruptcy of a carrier, its inclusion may provide a slight competitive advantage in this highly competitive market with very small margins.

A similar example can be provided for city administrators. Concerned with the negative impacts of the large amount of vans delivering freight in the city, a measure to regulate urban freight gaining popularity is the organisation of cargo bike deliveries from micro-consolidation centres in the city (Gómez et al., 2017; Janjevic et al., 2013). Given the limited capacity of these bikes however, the key for success for these facilities is the minimisation of the distance to the final customer. As numerous examples of bigger urban distribution centres proved however, the financial viability of these projects often turns out to be negative, mostly due to insufficient throughput volumes (Allen et al., 2007; Janjevic and Ndiaye, 2017). In this context, local variances in estimated demand, especially when being of a magnitude close to 50%, can significantly impact the business models of these hubs and may be the difference between an effective measure or a waste of public funding.

The value of information on the socio-economic profile of the e-shopper for retailers has already been mentioned in the first two sections. The knowledge that middle aged higher income Belgians have higher probability to shop online can help marketing division to better target their efforts. In the discussion of infrastructure investments, Kirby-Hawkins et al. (2018) provide the example of click-and-collect stores, which, similar to the previous two examples, can be better planned when including geographical variability in online shopping.

In this discussion, we have to remark not all purchases are delivered at the home address of the shopper. For Belgium, it was calculated that 75% of the people prefer home deliveries, which means the remaining 25% choose work locations, drop-off points, pick-up points and brick and mortar stores for their deliveries (Comeos, 2016). Because most stores and offices are located within urban areas, this implies a larger share of urban deliveries compared to urban purchases and the opposite in rural locations. However, since this shift will be similar for both the

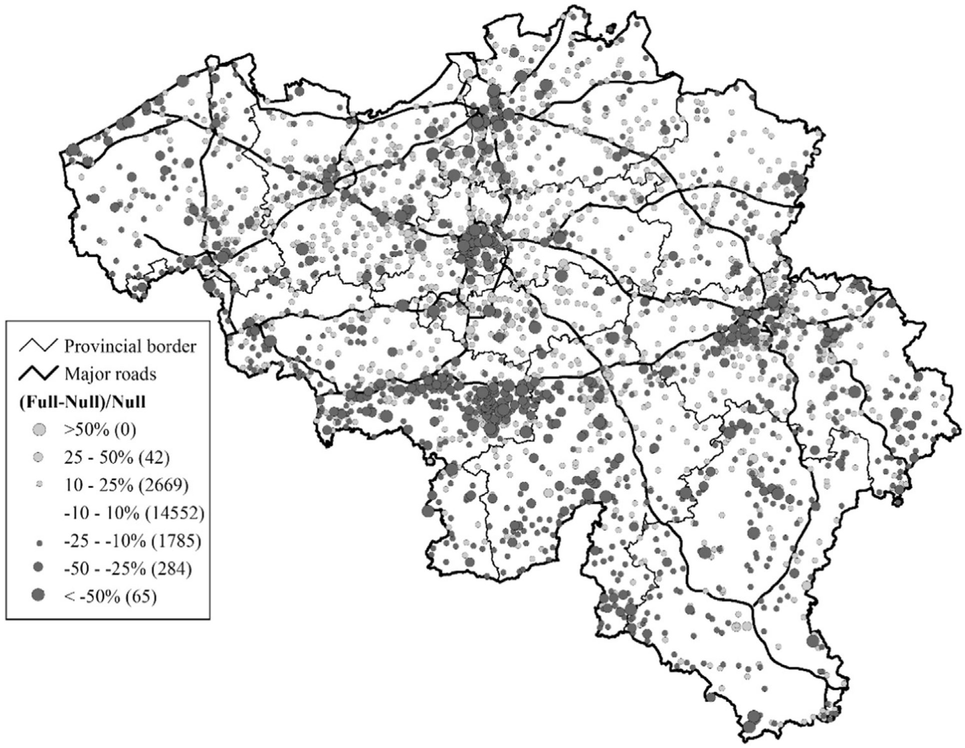


Fig. 3. Modelled order density vs population density.

modelled shoppers as well as the population, the presented results and discussion remain valid.

6. Conclusions

In this paper, we seek to provide insights into the spatial distribution of the demand for B2C e-commerce and its implications for sustainability analyses. Two research questions arise. Who is the online shopper? And where does he or she live? This learns us the extent to which the demand for e-commerce is proportional to the distribution of the population, as is often assumed. The case study area is Belgium.

We are able to provide an answer to the two research questions by analysing the dataset resulting from an online survey of over 1500 respondents concerning online shopping conducted by the Belgian retail federation (Comeos, 2016). First, the application of a logit regression allows the identification of the socio-economic variables that have a significant influence on the probability one shops online in Belgium. Similar to other studies in the Netherlands and the UK, the well-educated man in his thirties with a well-paid job has the highest probability to shop online. Whether this profile's shopping probability is due to time constraints, as proposed by Farag et al. (2007), or other incentives remains under study. Further, we found the urbanisation level of an area does not have a significant impact on the shopping probability. This could be due to the historical lack of urban planning in the country, resulting in relatively high shopping accessibility throughout Belgium, which turns the efficiency hypothesis (Anderson et al., 2003) invalid in our case.

Second, the extrapolation of the statistical model allows the calculation of the number of shoppers in each neighbourhood (statistical sector), thereby providing an answer to our second research question. Understandably most e-shoppers can be found in the dense urban areas, but the comparison of the absolute and relative numbers of shoppers displays clear discrepancies. As such, due to on average higher population densities in lower income neighbourhoods, and the opposite in higher income urban residential areas, the urban demand density might exhibit an inverse relation with the population density. Yet, while it is the case in this particular study, this hypothesis may not hold in other contexts, for example in world cities with high-end high-rise buildings to accommodate their wealthiest residents in the city centre. In rural areas, where the lowest population densities often point to lower socio-economic classes, we observe the contrary, i.e. with income and education enforcing existing variations in e-commerce demand.

Finally, the confrontation of the extrapolated number of potential buyers with a null model, i.e. where we assume that in each neighbourhood 50% of the population shops online, independent of their socio-economic profile, indicates that the demand for e-commerce goods is not evenly distributed over the population, as many involved stakeholders assume. On the contrary, large variations in e-commerce adoption exist among richer and poorer urban and rural areas. Considering the importance of understanding the geography of the demand for online shopping in both academic and professional environments, one should at least acknowledge the role of geographical variation in this story.

Granted, this paper is an attempt to raise awareness on the importance of the customer in e-commerce analyses. Although the sampling used to select respondents ensures validity of the results, the analysis here remains a snapshot of the situation at the moment of the data gathering. This implies it remains difficult to conclude on the hypothesis presented by Anderson et al. in 2003. Since they are in essence related to an evolutionary process, a temporal analysis would be a more appropriate methodology. Even now, with continuous e-commerce growth over the past years, research on the evolution of e-commerce adoption would be relevant. Not only because of the large international variance in e-commerce adoption rates, but also since the popularity of the online channel varies greatly from one product to another.

Building on the latter point, another point of concern stems from the lack of differentiation among products bought. Because of this, this study sketches a profile of the general online shopper and equals it to the demand for goods. Nonetheless, the freight volumes and flows eventually depend on the type of good. Hence a more in-depth analysis of this, which could potentially fall in the class of freight trip generation models, would only further enhance demand estimations.

Furthermore, it is important to note that this study is based on a single dataset concerning one aspect of the e-commerce story. Another interesting path for further research would be a cross-check with retail data, like in the case study of Leeds (Kirby-Hawkins et al., 2018), this would allow a better assessment of the impacts of the growth in e-commerce on traditional retail numbers and could help answering questions like: Does a higher probability of shopping implies a substitution effect in wealthier neighbourhoods?

In addition, it would be interesting if follow-up work focuses on quantifying the extent to which differences among customers define home delivery impacts and alternative delivery solutions. This may help to better estimate the size of the absolute error that is made by assuming a homogeneous distribution of the demand in urban freight modelling exercises. The calculation thereof should be able to support the findings in this paper.

Finally, the question may arise whether stakeholders need this kind of modelled demand since carriers know where the deliveries are going to. However, as has been mentioned in the discussion section, it is often difficult for these companies to separate B2B and B2C data since these goods are mostly handled and delivered together. While it is true to operational models could be built on top of this combined data, carriers often resort to population statistics for B2C market analyses, e.g. when implementing CDP networks. In addition, within this competitive environment data sharing is more exception than rule.

Because of the latter, the work in this paper can be of even more use for non-carriers involved in the topic. Given the major role (local) governments play in mitigating the negative impacts of B2C home deliveries, they too should be aware of the importance of socio-economic factors that can result in major demand variation at the neighbourhood level. With only 50–57% of the Belgians shopping online, clearly not everybody is enjoying the wide range of opportunities e-commerce offers. Furthermore, as demonstrated in the discussion section, the spatial variations may influence the decision on the implementation of CDPs or the use of greener vehicles. While local policymakers can have a significant influence on the former, through infrastructure and urban planning, the latter too can be influenced by for example the delineation of low emission zones. In addition, coupled with freight trip generation models mentioned above, the analysis presented here can help to better understand traffic flows, potential vehicle types, loading space requirements and other freight parameters and thus provide a useful tool for local infrastructure design and mobility planning.

Hence, this study may help the different stakeholders involved to grasp the geography of the Belgian online shopper and its impact on related freight flows.

Acknowledgments

The author gratefully acknowledge the retail federation Comeos for providing data.

References

Allen, J., Thorne, G., Browne, M., 2007. Good practice guide on urban freight transport. BESTUFS. Retrieved from 〈www.bestufs.net〉.

Anderson, W.P., Chatterjee, L., Lakshmanan, T.R., 2003. E-commerce, transportation, and economic geography. Growth Change 34 (4), 415–432. https://doi.org/10.1046/j.0017-4815.2003.00228.x.

Anderson, W.P., Leinbach, T.R., 2007. E-commerce, logistics and the future of globalized freight. In: Leinbach, T.R., Capineri, C. (Eds.), Globalized Freight Transport. Edward Elgar Publishing, Cheltenham, UK, pp. 287.

Antrop, M., 2004. Landscape change and the urbanization process in Europe. Landsc. Urban Plan. 67 (1–4), 9–26. https://doi.org/10.1016/S0169-2046(03)00026-4.

Arvidsson, N., Pazirandeh, A., 2017. An ex ante evaluation of mobile depots in cities: a sustainability perspective. Int. J. Sustain. Transp. 11 (8), 623–632. https://doi.org/10.1080/15568318.2017.1294717.

Birkin, M., Clarke, G., Clarke, M., 2017. Retail Location Planning in an Era of Multi-Channel Growth. Routledge, Abingdon, UK.

Boschma, R.A., Weltevreden, J.W.J., 2008. An evolutionary perspective on Internet adoption by retailers in the Netherlands. Environ. Plan. A 40 (9), 2222–2237. https://doi.org/10.1068/a3972.

Browne, M., 2001. The impact of e-commerce on transport. In: Joint OECD/ECMT Seminar. Retrieved from 〈https://doaj.org/article/92cf5b03a89e48779ecab0d54de2c824〉.

Browne, M., Allen, J., 1999. The impact of sustainability policies on urban freight transport and logistics systems. In: Meersman, H., Van De Voorde, E., Winkelmans, W. (Eds.), World Transport Research, Vol 1: Transport Modes and Systems. Elsevier, Oxford, pp. 505–518.

Cao, X., Chen, Q., Choos, S., 2013. Geographic distribution of e-shopping. Transp. Res. Rec.: J. Transp. Res. Board 2383, 18–26.

Cherrett, T., Allen, J., McLeod, F., Maynard, S., Hickford, A., Browne, M., 2012. Understanding urban freight activity - key issues for freight planning. J. Transp. Geogr. 24, 22–32. https://doi.org/10.1016/j.jtrangeo.2012.05.008.

Clarke, G., Thompson, C., Birkin, M., 2015. The emerging geography of e-commerce in British retailing. Reg. Stud. Reg. Sci. 2 (1), 370–390. https://doi.org/10.1080/21681376.2015.1054420.

Comeos, 2016. E-commerce Belgium 2016. Retrieved from 〈www.comeos.be〉.

Couclelis, H., 2004. Pizza over the Internet: e-commerce, the fragmentation of activity and the tyranny of the region. Entrep. Reg. Dev. 16, 41–54. https://doi.org/10.1080/0898562042000205027.

Doherty, N.F., Ellis-Chadwick, F., 2010. Internet retailing: the past, the present and the future. Int. J. Retail Distrib. Manag. 38 (11/12), 943–965. https://doi.org/10.1108/09590551011086000.

Donthu, N., Garcia, A., 1999. The internet shoper. J. Advert. Res. 39 (3), 52.

Ducret, R., Lemarié, B., Roset, A., 2016. Cluster analysis and spatial modeling for urban freight. Identifying homogeneous urban zones based on urban form and logistics characteristics. Transp. Res. Procedia 12, 301–313. https://doi.org/10.1016/j.trpro.2016.02.067.

E-commerce Europe, 2016a. Country Report Belgium 2016. Retrieved from 〈https://www.ecommercewiki.org/reports/61/belgium-b2c-ecommerce-country-report-2016〉.

E-commerce Europe, 2016b. European B2C E-commerce Report 2016. Retrieved from 〈https://www.ecommerce-europe.eu/cms/showpage.aspx?Id=2049〉.

Eurostat, 2017. E-commerce statistics for individuals. Retrieved from 〈http://ec.europa.eu/eurostat/statistics-explained/index.php/E-commerce_statistics_for_individuals〉.

Farag, S., Krizek, K.J., Dijst, M., 2006a. E-Shopping and its Relationship with In-store Shopping: empirical Evidence from the Netherlands and the USA. Transp. Rev. 26 (1), 43–61. https://doi.org/10.1080/01441640500158496.

Farag, S., Schwanen, T., Dijst, M., Faber, J., 2007. Shopping online and/or in-store? A structural equation model of the relationships between e-shopping and in-store shopping. Transp. Res. Part A: Policy Pract. 41 (2), 125–141. https://doi.org/10.1016/j.tra.2006.02.003.

Farag, S., Weltevreden, J.W.J., van Rietbergen, T., Dijst, M., van Oort, F., 2006b. E-shopping in the Netherlands: does geography matter? Environ. Plan. B: Plan. Des. 33 (1), 59–74. https://doi.org/10.1068/b31083.

Gevaers, R., Van de Voorde, E., Vanelslander, T., 2011. Characteristics and typology of last-mile logistics from an innovation perspective in an urban context. In: City Distribution and Urban Freight Transport: Multiple Perspectives. Edward Elgar Publishing, Cheltenham, UK, pp. 56–71.

Gevaers, R., Van de Voorde, E., Vanelslander, T., 2014. Cost modelling and simulation of last-mile characteristics in an innovative B2C supply chain environment with implications on urban areas and cities. Procedia - Soc. Behav. Sci. 125, 398–411. https://doi.org/10.1016/j.sbspro.2014.01.1483.

Gómez, T., Demeyer, V., Manville, S., 2017. D3.4 Set of Recommendations for follower cities related to infrastructural and regulative measures enabling a step-wise introduction of near zero emissions urban logistics. Cyclelogistics Ahead – a Key Step towards Zero Emission Logistics in Cities. Retrieved from 〈http://cyclelogistics.eu/index.php?Id=39&folder_id=119〉.

Gonzalez-Feliu, J., Ambrosini, C., Routhier, J.-L., 2012. New trends on urban goods movement: modelling and simulation of e-commerce distribution. Eur. Transp. 50.

Hesse, M., 2002. Shipping news: the implications of electronic commerce for logistics and freight transport. Resour. Conserv. Recycl. 36 (3), 211–240. https://doi.org/10.1016/S0921-3449(02)00083-6.

Jamagne, P., 2001. Vademecum Statistische sectoren, pp. 1–82. Retrieved from 〈https://statbel.fgov.be/sites/default/files/files/opendata/Statistischesectoren/Secteurs.stat-NL_tcm325-174181.pdf〉.

Janjevic, M., Kaminsky, P., Ndiaye, A.B., 2013. Downscaling the consolidation of goods – state of the art and transferability of micro - consolidation initiatives. Eur. Transp./Trasp. Eur. 54 (4), 1–23.

Janjevic, M., Ndiaye, A., 2017. Investigating the financial viability of urban consolidation centre projects. Res. Transp. Bus. Manag. 24, 101–113. https://doi.org/10.1016/j.rtbm.2017.05.001.

Kau, A.K., Tang, Y.E., Ghose, S., 2003. Typology of online shoppers. J. Consum. Mark. 20 (2), 139–156. https://doi.org/10.1108/07363760310464604.

Kiba-Janiak, M., 2017. Urban freight transport in city strategic planning. Res. Transp. Bus. Manag. 24, 4–16. https://doi.org/10.1016/j.rtbm.2017.05.003.

Kirby-Hawkins, E., Birkin, M., Clarke, G., 2018. An investigation into the geography of corporate e-commerce sales in the UK grocery market. Environ. Plan. B: Urban Anal. City Sci. https://doi.org/10.1177/2399808318755147.

Lin, J., Chen, Q., Kawamura, K., 2016. Sustainability SI: logistics cost and environmental impact analyses of urban delivery consolidation strategies. Netw. Spat. Econ. 16 (1), 227–253. https://doi.org/10.1007/s11067-014-9235-9.

Lindholm, M., Behrends, S., 2012. Challenges in urban freight transport planning - a review in the Baltic Sea Region. J. Transp. Geogr. 22, 129–136. https://doi.org/10.1016/j.jtrangeo.2012.01.001.

Luyten, S., Van Hecke, E., 2007. De Belgische stadsgewesten 2001. Retrieved from ⟨http://www.briobrussel.be/assets/onderzoeksprojecten/p009n014_nl.pdf⟩.
McLeod, F., Cherrett, T., Song, L., 2006. Transport impacts of local collection/delivery points. Int. J. Logist. Res. Appl. 9 (3), 307–317. https://doi.org/10.1080/13675560600859565.
Morganti, E., Dablanc, L., Fortin, F., 2014. Final deliveries for online shopping: the deployment of pickup point networks in urban and suburban areas. Res. Transp. Bus. Manag. 11, 23–31. https://doi.org/10.1016/j.rtbm.2014.03.002.
Motte-baumvol, B., Belton-chevallier, L., Dablanc, L., Morganti, E., Belin-Munier, C., 2017. Spatial dimensions of E-shopping in France. Asian Transp. Stud. 4 (3), 585–600.
Park, H., Park, D., Jeong, I.J., 2016. An effects analysis of logistics collaboration in last-mile networks for CEP delivery services. Transp. Policy 50, 115–125. https://doi.org/10.1016/j.tranpol.2016.05.009.
Rimmer, P.J., Kam, B.H., 2018. Consumer Logistics. Surfing the Digital Wave. Edward Elgar Publishing, Cheltenham, UK.
Sim, L.L., Koi, S.M., 2002. Singapore's internet shoppers and their impact on traditional shopping patterns. J. Retail. Consum. Serv. 9 (2), 115–124. https://doi.org/10.1016/S0969-6989(01)00029-7.
Statistics Belgium, 2014. Census 2011 - België. Retrieved from ⟨http://census2011.fgov.be/⟩.
Statistics Belgium, 2018. Fiscale inkomens. Retrieved from ⟨https://statbel.fgov.be/nl/themas/huishoudens/fiscale-inkomens/plus⟩.
Van Meeteren, M., Boussauw, K., Derudder, B., Witlox, F., 2016. Flemish diamond or ABC-Axis. The spatial structure of the Belgian metropolitan area. Eur. Plan. Stud. 24 (5), 974–995.
Verhetsel, A., Thomas, I., Beelen, M., 2010. Commuting in Belgian metropolitan areas. J. Transp. Land Use 2 (3), 109–131.
Vrechopoulos, A.P., Siomkos, G.J., Doukidis, G.I., 2001. Internet shopping adoption by Greek consumers. Eur. J. Innov. Manag. 4 (3), 142–153. https://doi.org/10.1108/14601060110399306.
Weltevreden, J.W.J., 2007. Substitution or complementarity? How the Internet changes city centre shopping. J. Retail. Consum. Serv. 14 (3), 192–207. https://doi.org/10.1016/j.jretconser.2006.09.001.
Weltevreden, J.W.J., 2008. B2c e-commerce logistics: the rise of collection-and-delivery points in The Netherlands. Int. J. Retail Distrib. Manag. 36 (8), 638–660. https://doi.org/10.1108/09590550810883487.
Weltevreden, J.W.J., van Rietbergen, T., 2007. The Role of Perceived City Centre Attractiveness. Tijdschr. Voor Econ. En. Soc. Geogr. 98 (1), 68–85. https://doi.org/10.1111/j.1467-9663.2007.00377.x.
Zhou, Y., Wang, X. (Cara), 2014. Explore the relationship between online shopping and shopping trips: an analysis with the 2009 NHTS data. Transp. Res. Part A: Policy Pract. 70, 1–9. https://doi.org/10.1016/j.tra.2014.09.014.