DISCUSSION PAPER SERIES

DP12216

SHOPPING EXTERNALITIES AND RETAIL CONCENTRATION: EVIDENCE FROM DUTCH SHOPPING STREETS

Hans Koster, Ilias Pasidis and Jos van Ommeren

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

ISSN 0265-8003

SHOPPING EXTERNALITIES AND RETAIL CONCENTRATION: EVIDENCE FROM DUTCH SHOPPING STREETS

Hans Koster, Ilias Pasidis and Jos van Ommeren

Discussion Paper DP12216
Published 11 August 2017
Submitted 11 August 2017

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

SHOPPING EXTERNALITIES AND RETAIL CONCENTRATION: EVIDENCE FROM DUTCH SHOPPING STREETS

Abstract

Why do shops cluster in shopping streets? According to theory, retail firms benefit from shopping externalities. We identify these externalities for the main shopping streets in the Netherlands by estimating the effect of footfall – the number of pedestrians that pass by – on store owner's rental income, which is a composite of the effects of footfall on shop rent and on vacancy rates. We address endogeneity issues by exploiting spatial variation between intersecting streets. Our estimates imply an elasticity of rental income with respect to footfall of 0.25. We find that a shop's marginal benefit of a passing pedestrian is € 0.005. It follows that subsidies to retail firms that increase with the levels of footfall generated by these shops are welfare improving. The optimal subsidy to store owners is, on average, 10 percent of the rent, but is higher for retail firms that generate high levels of footfall. Although explicit subsidies are controversial and difficult to implement, our results seem to justify current policy practices which cluster shops by pedestrianisation of shopping streets or by providing subsidised parking for shoppers.

JEL Classification: R30, R33

Keywords: retail, shopping externalities, rents, Vacancies, agglomeration economies, footfall.

Hans Koster - h.koster@vu.nl
Vrije Universiteit Amsterdam and CEPR

Ilias Pasidis - ipasidis@gmail.com
Universitat de Barcelona

Jos van Ommeren - jos.van.ommeren@vu.nl
Vrije Universiteit Amsterdam

Acknowledgements
This work has benefited from a VENI research grant from the Netherlands Organisation for Scientific Research. We thank Sergej Gubin, Jordi Jofre-Monseny, Albert Sollé-Ollé, Jacques Thisse and Elisabet Viladecans-Marsal for useful comments on this paper, as well as participants of seminars at the Baltic International Centre for Economic Policy Studies, the University of Barcelona, the Autonomous University of Barcelona, IDE-JETRO in Chiba, the London School of Economics and participants of the Urban Economics Association conference in Copenhagen.

Shopping externalities and retail concentration: Evidence from Dutch shopping streets*

By HANS R.A. KOSTER,[a] ILIAS PASIDIS[b] *and* JOS VAN OMMEREN[c]

Version: 10 August 2017

SUMMARY — Why do shops cluster in shopping streets? According to theory, retail firms benefit from shopping externalities. We identify these externalities for the main shopping streets in the Netherlands by estimating the effect of footfall – the number of pedestrians that pass by – on store owner's rental income, which is a composite of the effects of footfall on shop rent and on vacancy rates. We address endogeneity issues by exploiting spatial variation between intersecting streets. Our estimates imply an elasticity of rental income with respect to footfall of 0.25. We find that a shop's marginal benefit of a passing pedestrian is € 0.005. It follows that subsidies to retail firms that increase with the levels of footfall generated by these shops are welfare improving. The optimal subsidy to store owners is, on average, 10 percent of the rent, but is higher for retail firms that generate high levels of footfall. Although explicit subsidies are controversial and difficult to implement, our results seem to justify current policy practices which cluster shops by pedestrianisation of shopping streets or by providing subsidised parking for shoppers.

JEL-code — R30, R33

Keywords — retail, shopping externalities, rents, vacancies, agglomeration economies, footfall.

---

* This work has benefited from a VENI research grant from the Netherlands Organisation for Scientific Research. We thank Sergej Gubin, Jordi Jofre-Monseny, Albert Sollé-Ollé, Jacques Thisse and Elisabet Viladecans-Marsal for useful comments on this paper, as well as participants of seminars at the Baltic International Centre for Economic Policy Studies, the University of Barcelona, the Autonomous University of Barcelona, IDE-JETRO in Chiba, the London School of Economics and participants of the Urban Economics Association conference in Copenhagen.

[a] Corresponding author. Department of Spatial Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105 1081 HV Amsterdam, e-mail: h.koster@vu.nl. The author is also affiliated with the Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam.

[b] Department of Economics, Universitat de Barcelona, Avinguda Diagonal 690, 08034 Barcelona, Spain, e-mail: ipasidis@gmail.com; ipasidis@ub.edu. The author is also affiliated with the Barcelona Institute of Economics (IEB), c/ John Maynard Keynes, 1-11, 08034, Barcelona.

[c] Department of Spatial Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105 1081 HV Amsterdam, e-mail: jos.van.ommeren@vu.nl. The author is also affiliated with the Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam.

I. Introduction

One of the main reasons that people choose to live in the city is the presence of a rich variety of consumer goods and services offered by the retail sector (Glaeser et al., 2001). Shops tend to be concentrated either in pedestrianised shopping streets and shopping districts, often located in city centres, or in shopping malls. As an illustration, walking is so important for shopping that the majority of all Dutch pedestrian movements occur while shopping.[1] In Europe, shops are mostly concentrated in pedestrianised shopping streets.

Arguably, the most important reason for retail firms to cluster is the presence of *shopping externalities*, which are generated by consumers' 'trip-chaining' behaviour. Shopping externalities have a simple logic. In retail markets, transportation costs are usually paid by customers and incurred on a shopping trip basis (Claycombe, 1991). Consumers who visit several shops benefit from reductions in transport and search costs. In the context of shopping streets, a retail firm's productivity function depends on local footfall, which captures the number of pedestrians that pass a shop. Footfall tends to be higher in areas with more shops, since pedestrians tend to visit multiple shops in order to find the best shopping options. Hence, the associated reductions in costs for consumers imply a shopping externality *for retail firms*, which is enhanced when multiple retail firms are located in close proximity (Eaton and Lipsey, 1982; Claycombe, 1991; Schulz and Stahl, 1996).[2] Similar to other agglomeration advantages, these shopping externalities are expected to capitalise into store owners' rental income, defined as the rent paid to store owners multiplied by the share of the time that the shop is occupied.[3]

In the empirical literature, however, only limited attention has been given to the importance of shopping externalities. Of course, we are not the first to argue that the most important reason for shops to cluster is the presence of shopping externalities, but it is the first paper that quantifies these externalities. We contribute to the literature in the following ways.

First, we introduce a measure of shopping externalities, *footfall.*[4] We argue and demonstrate that footfall is a superior measure of shopping externalities compared to the number of shops in the vicinity of a retail firm. The number of shops is an alternative measure that underestimates the presence of shopping externalities because retail firms

[1] We use data from Statistics Netherlands. We exclude hiking and recreational walking activities.

[2] Externalities arise when a sufficient number of pedestrians are involved in multipurpose shopping trips. If there is substantial heterogeneity between shops in generating footfall, the number of shops is a poor proxy for externalities. For instance, a popular clothing store is likely to generate substantial footfall, whereas a fast food store may not generate much footfall, but will benefit from footfall created by other shops.

[3] In retail markets, agglomeration economies occur very locally, so capitalisation into wages must be negligible, because differences in commuting time between competing shops within the same shopping district are negligible. In non-retail markets, agglomeration advantages mainly capitalise into wages (e.g., Arzaghi and Henderson, 2008).

[4] In the retail industry, footfall is a standard measure to explain the attractiveness of a shopping location.

strongly vary in the amount of footfall they generate, so the correlation between footfall and number of shops in the vicinity is not very high. We demonstrate that the elasticity of rental income with respect to footfall is about twice the elasticity of rental income with respect to number of shops.[5] We provide a number of arguments why footfall captures shopping *externalities*, and not simply captures local variation in shopping demand (e.g., our measure of footfall predominantly includes shoppers who visit several shops). In contrast to the extensive retail literature which focuses on US shopping malls, we focus on the *full* population of main *shopping streets* of one country: the Netherlands.[6]

An important feature of shopping streets is that they are dominated by two sectors: clothing and cafés/restaurants. The main strategy followed by the retail firms in these sectors is to differentiate themselves by supplying heterogeneous products. This is in sharp contrast to other retail sectors that are examined in the economic literature, which offer homogeneous products and where spatial differentiation is the main strategy (e.g. movie theatres, gas stations, or video retailers, see Davis, 2006; Netz and Taylor, 2002; Seim, 2006).

It is important to note that shopping streets are characterised by a very different form of retail organisation than shopping malls. In contrast to the evidence for shopping malls, we will show that property ownership in shopping streets is very fragmented. As a consequence, internalisation of shopping externalities does not occur in shopping streets.[7] Thus, policies that foster retail concentration by providing subsidies are potentially welfare improving.[8] We then make a distinction between subsidies given to *store owners* (independent of the generated footfall*)*, and subsidies to *retail firms* based on the generated footfall. The former type of subsidy may lead to more shops in the shopping street, whereas the second type of subsidy will predominantly attract footfall-generating retail firms.

The second, and main, contribution of the current paper is the identification of shopping externalities by estimating the causal effect of footfall on the rental income of store owners, which depends on the rent paid by tenants as well as the probability that a property lies empty. As has been widely discussed in the agglomeration literature, proxies for spatial concentration, such as footfall, tend to be endogenous because they are correlated to

---

[5] Similarly, in the labour market, it is common to measure the number of jobs as a measure of agglomeration and not the number of firms, because firms strongly differ in their workforce size.

[6] Shopping mall floor space per person is more than tenfold in the US compared to Europe (2,150m^2 per thousand people in the US compared to 182m^2 per thousand people in Europe VS in 2011 (Cushman and Cushman, 2011).

[7] In shopping malls, store owners set the rent based on shop turnover, so shopping externalities are internalised. Thus, they charge lower rents to footfall-generating shops (or 'anchor stores') (Brueckner, 1993; Pashigian and Gould, 1998; Konishi and Sandfort, 2003), which could be regarded as a first-best subsidy.

[8] Many examples of such policies can be given. For many European countries, in particular Germany, it could be argued that pedestrianised areas subsidise local store owners, as the advantages are local whereas the disadvantages of prohibiting car use fall on other agents. Subsidies to park-and-ride facilities, including free public transport towards city centres is another similar example.

unobserved location characteristics. We address this issue by focusing on shops that are *very close* to each other (within 50m) but on different *intersecting* streets, controlling for an extensive set of shop and street characteristics.[9] By using spatial variation in footfall between intersecting streets, we control for unobserved locational endowments that attract both shoppers and retail firms (e.g. free parking etc.). Our identifying assumption is that retail firms with similar preferences for location characteristics will locate in close proximity to each other while they will sort themselves into lower and higher footfall streets, depending on their preferences for footfall. Retail firms that benefit strongly from footfall (e.g. mainstream clothing shops) will sort into high-footfall streets and pay higher rents.

We show that footfall has a strong positive effect on rental income with an elasticity of approximately 0.25, whereas the elasticity of rental income with respect to the number of retail firms is 0.10. Thus, there are substantial external benefits from fostering footfall and retail concentration. Based on these estimates, the optimal subsidy that should be given to store owners amounts to about 10 percent of their rent, on average. However, for retail firms that generate above-average footfall for surrounding shops, the optimal subsidy should be substantially higher. We subject our results to a wide range of robustness checks and ancillary regressions, for example by exploiting temporal rather than spatial variation in footfall and by investigating differential effects of footfall on 'anchor' or chain-stores. Furthermore, we do not find evidence of negative external effects of footfall on house prices.

Our paper links and contributes to three strands of literature. First, it relates to the literature on spatial competition and product differentiation (d'Aspremont et al., 1979; Osborne and Pitchik, 1987). Davis (2011) focuses on movie theatres, and evaluates consumers' transport costs, the effect of geographic differentiation, and the extent of market power among other things. Seim (2006) shows that there are significant returns to product (or spatial) differentiation and illustrates that markets with more scope for differentiation support greater entry. Jia (2008) and Arcidiacono et al. (2016) study the impact of Wal-Mart on the retail market, among others on incumbent (discount) supermarkets and small grocery stores. Zhou (2014) shows that multiproduct search, which is important when consumers buy multiple products in one shopping trip, can significantly influence retail firms' pricing decisions. Johansen and Nilssen (2016) investigate the conditions under which one-stop shopping causes the formation of big stores. Clapp et al. (2016) focus on openings and closings of multiline department stores and find evidence for strong negative competitive effects within the own branch. However, many of these papers ignore that shops may benefit from each other when located close to each other, in particular when offering not too differentiated products.

[9] Because people follow certain routes for their shopping trips, footfall strongly differs between intersecting streets. On average, the high-footfall street is roughly twice as busy as its intersecting low-footfall street.

It also relates to a second literature that explicitly focuses on the benefits of agglomeration for firms. There is ample evidence that firms that locate close together benefit through input- and output sharing, labour market pooling and knowledge spillovers Marshall (1920). Current evidence suggests that the elasticity of productivity with respect to density is around 0.05 (see e.g. meta-analysis by Melo et al., 2009). This literature typically focuses on the manufacturing industry. Recent evidence by Koster et al. (2014) shows however that agglomeration economies are substantially more important for retail firms.

Our findings contribute also to a more policy-oriented literature studying the effectiveness of retail policies, in particular towards the effects of the opening of large 'big-box' retailers near the urban fringe (Sanchez-Vidal, 2016). Some studies demonstrate that the welfare effects of current planning policies that hinder entry in retail markets, and particularly of large retailers, are negative. Several studies have shown that regulation policies reduce retail productivity and job growth and increase market power of incumbent stores (Bertrand and Kramarz, 2002; Schivardi and Viviano 2011; Haskel and Sadun, 2012; Cheshire et al., 2015). By contrast, our study shows that policies in the form of subsidies for footfall-generating retail firms may be welfare improving. This conclusion is consistent with the shopping mall literature which shows that the provision of substantial rent discounts to 'anchor' stores in US shopping malls to internalise externalities is common practice (Pashigian and Gould, 1998; Gould et al., 2005). Hence, because monetary subsidies to store owners and shops footfall-generating firms are difficult to implement for a combination of reasons, shopping externalities may justify a range of public policies such as the creation of pedestrianised shopping streets, which effectively cluster shops and internalise external benefits.

The remainder of this paper is structured as follows. In Section II we discuss the theoretical framework that guides the empirical results. Section III introduces the econometric framework, the data and reports the descriptive statistics. In Section IV, we present and discuss our results, including the estimates of the optimal subsidy. Section V summarises a couple of extensions, including the effects of footfall on residential house prices, and discusses the main the sensitivity analysis. These are described in more detail in the Appendix. We draw conclusions in Section VII.

II. Theoretical framework

A. Rental income, rents and vacancies

We aim to measure the presence of shopping externalities by estimating the effect of footfall on (expected) rental income of store owners, denoted by I.[10] Hence, in what follows, we

[10] As an alternative, one may estimate the effect on transaction prices of shops. There are two reasons we prefer to focus on rental income. First, transaction prices reflect expectations about future rents and therefore future levels of footfall. Second, sales transactions are rare relative to rent transactions. In our data, only 10 percent of the observations refer to sales transactions.

make a distinction between *store owners* that own shops and *retail firms* that rent shops. Shops can be either occupied by *retail firms* or are vacant. Store owners of vacant shops need advertising services to find a new tenant, which is costly. Given rent p and vacancy rate v, rental income of a shop is given by:

(1) $$I = p(1 - v) - cv.$$

where $p(1 - v)$ is rental income when the property is let to a retail firm and cv is the advertising costs. It seems reasonable to assume that, at least in the long run, the advertising costs c are proportional to p, so $c = \kappa p$, where $\kappa > 0$. Because vacancy rates tend to be small, $\log(1 - (1 + \kappa)v) \approx -(1 + \kappa)v$. Hence, the *logarithm* of rental income, $\log I$, is then (approximately) equal to $\log p - (1 + \kappa)v$.

If footfall has an effect on the rent and vacancy rate, it follows that the effect of log footfall on log rental income can be written as the sum of the marginal effect of log footfall on log rent and the marginal effect of log footfall on the level of the vacancy rate:

(2) $$\frac{\partial \log I}{\partial \log f} = \frac{\partial \log p}{\partial \log f} - (1 + \kappa)\frac{\partial v}{\partial \log f}.$$

In our econometric framework, we will estimate the marginal effects of footfall on the logarithm of rent, as well as on the level of the vacancy rate. We will demonstrate theoretically and empirically that the effect of footfall on log rent is positive and the effect on the vacancy rate is negative. This implies that for κ equals 0, we get the lower bound estimate of the effect of footfall on log rental income. Moreover, we calculate the effect of footfall on rental income for different values of κ which will be approximated using information about advertising costs for Dutch retail properties.

A standard hedonic model cannot be used to predict $\partial \log p / \partial \log f$ and $\partial v / \partial \log f$ because vacancies are not incorporated in such a framework. In Appendix A.1 we therefore set up a search and matching framework showing that $\partial p / \partial f > 0$ and that $\partial v / \partial f < 0$. Hence, our prior is that $\partial \log I / \partial \log f > 0$. Moreover, when we endogenise footfall in Appendix A.2, so that it is dependent on vacancy rates, i.e. $f = (1 - v)\bar{f}$, where $\bar{f}$ is the footfall generated when all shops are occupied by retail firms. We show that the impact on prices and vacancy rates will be even stronger $\partial p / \partial \bar{f} > \partial p / \partial f$ and $\partial v / \partial \bar{f} < \partial v / \partial f$.

B. Welfare and retail policies

Let us now focus on welfare effects of retail policies. Intuition suggest that there may be room for retail policies when retail firms have an influence on the level of footfall in a shopping street, potentially creating shopping externalities. One may distinguish between three behavioural margins. First, the footfall in a shopping street depends negatively on the vacancy rate. The decision (not) to open a shop ignores that there is a positive shopping externality on other retail firms. Although interesting, we will ignore this issue as the size of

this externality is numerically small given that vacancy rates are low.[11] Second, shops may increase footfall in the street for example by advertising for their shop. This is an interesting external effect which we do not analyse in the current paper because it requires information on the footfall and advertising expenditures by each shop. Third, and importantly for retail policy, store owners endogenously decide to build a store, so the number of shops in the shopping street N is endogenous. This is relevant because footfall depends positively on the number of shops.[12] This issue is essential for policies that influence the number of shops and therefore the level of footfall. An important policy question is whether an unregulated market leads to the optimal concentration of retail firms in a shopping street.

Suppose therefore that the number of shops N is endogenously determined in a competitive market. Let us further assume that the per-period marginal construction and maintenance costs are equal to $\pi(N)$, which is an increasing and convex function of N. Furthermore, suppose that footfall is an increasing function of the number of shops N in the shopping street, so $\partial f/\partial N > 0$.

The number of shops is then determined by the following equation, which states that the marginal expected income of opening shop N is equal to the marginal cost of opening a shop:

$$I(N) = \pi'(N),$$

(3)

$$p\big(f(N)\big)\Big(1 - v\big(f(N)\big)\Big) - cv\big(f(N)\big) = \pi'(N).$$

Note that the marginal benefit of opening a shop – i.e. the left-hand side of the above equation – is an increasing function of N. This is intuitive, because each shop that opens up in the shopping street benefits from the footfall generated by nearby retail firms. This implies that the cost of opening a shop – i.e. $\pi(N)$ – must be strongly convex, so that the second-order condition holds.[13]

To calculate the optimal subsidy to store owners is straightforward. The welfare generated in a shopping street is equal to $NI - \pi$. Maximisation of welfare with respect to *the number of shops* implies that $I(N) - \pi'(N) + N(\partial I/\partial N) = 0$, where the marginal store owner will ignore the last term when considering to open (or close) a store. Hence, the marginal external benefit of opening up a new shop is equal to:

(4)

$$N\frac{\partial I}{\partial N} = I \cdot \varepsilon_{I,N} > 0,$$

[11] Note that it is in general not clear whether store owners with vacancies choose the optimal advertising expenditure. See, for example, Hosios (1990) for an analysis of optimal advertising expenditure in the labour market. We therefore leave this question for further research.

[12] This has no direct consequences for our empirical investigation of the effects of footfall on rental income.

[13] The latter is likely true because in shopping streets, shops almost always solely occupy the ground level floor of a building. This assumption seems also supported by the observation that in many shopping streets all ground floor levels are used for shopping, so the number of shops is restricted by the length of the street.

where $\varepsilon_{I,N}$ denotes the elasticity of rental income with respect to the number of shops. Importantly, we will assume that this shopping externality is *not* internalised by store owners. This seems a reasonable assumption for shopping streets due to the dispersed ownership of shops, which is supported by our data.[14] Given this assumption, the Pigouvian subsidy to the marginal store owner in the first-best optimum is $\varepsilon_{I,N}$ times the rental income of a shop. In our empirical analysis, we are able to calculate this subsidy as we will observe rental income and estimate $\varepsilon_{I,N}$.[15] As we will demonstrate that $\varepsilon_{I,N}$ is positive and increasing in N , it appears that for shopping streets that are unregulated, the calculated subsidy will be an underestimate as the subsidy in the optimum will be higher.

The above expression does not make clear whether shops in large shopping streets should receive higher levels of subsidies than shops in small ones. Hence let us analyse the difference in optimal subsidy level between two shopping streets, labelled 1 and 2, which can be written as:

$$N^1 \frac{\partial I^1}{\partial N^1} - N^0 \frac{\partial I^0}{\partial N^0} = I^1 \cdot \varepsilon_{I,N}^1 - I^0 \varepsilon_{I,N}^0. \tag{5}$$

Now suppose that street 1 contains more shops than street 2. This equation shows that larger shopping streets must receive larger subsidies as long as the elasticity of rental income with respect to number of shops is positive (so, $I^1 > I^0$) and non-decreasing in the number of shops (so, $\varepsilon_{I,N}^1 > \varepsilon_{I,N}^0$). We will see that both conditions are fulfilled. This implies that subsidies to retail firms in large shopping streets must exceed those in small shopping streets.

We emphasise that the above analysis ignores that shops are heterogeneous in the amount of footfall they generate. Clearly, a policy which provides different levels of subsidy to specific retail firms is welfare enhancing compared to the policy which does not, as it changes the type of retail firms. Such a policy is particularly relevant to understand in the context where new retail firms apply for (implicit) subsidies to local governments at the time they open a store based on the argument that they (substantially) increase footfall for other firms in their vicinity. To take heterogeneity between shops into account, we redefine I as the *average* rental income in a shopping street. It is then useful to rewrite (4) as:

$$N \frac{\partial I}{\partial N} = N \frac{\partial I}{\partial f} \frac{\partial f}{\partial N} = I \cdot \varepsilon_{I,f} \cdot \varepsilon_{f,N} > 0. \tag{6}$$

[14] In contrast, in shopping malls, developers will internalise these externalities by determining the optimal number of stores and by charging lower rents to footfall-generating shops (or 'anchor stores') (Brueckner, 1993; Pashigian and Gould, 1998; Konishi and Sandfort, 2003). Therefore, in a shopping mall, the developer is able to provide first-best subsidies, based on the amount of footfall generated by each store, and maximize a mall's welfare.

[15] To evaluate the exact general-equilibrium welfare improvements associated with such a subsidy is out of the scope of the current paper and depends on, among others, to what extent such a policy leads to openings of new retail firms in the economy, and to what extent existing shops move to other shopping streets due to the policy. Given that the demand for street shopping is likely far from inelastic, it is plausible that the welfare increase of the subsidy is not negligible.

Hence, $\varepsilon_{I,f}$ has been written as the product of the elasticity of the average rental income with respect to footfall, $\varepsilon_{I,f}$, and the elasticity of footfall with respect to the number of firms, $\varepsilon_{f,N}$, where the latter is *firm specific*.[16]

In our empirical application, we observe (average) rental income I and will estimate the elasticity $\varepsilon_{I,f}$.[17] Note that $\varepsilon_{f,N}$ is not observed by us as we do not have plausibly exogenous variation to estimate the effect of number of shops on footfall. We therefore proceed by assuming specific values for $\varepsilon_{f,N}$, and then calculate the optimal subsidy to attract a marginal retail firm N which generates footfall *f*. Similar to our conclusions based on (5), it is straightforward to show that, conditional on $\varepsilon_{f,N}$, shops in streets with larger footfall levels should receive higher subsidies provided that the elasticity of rental income with footfall is non-decreasing in footfall. We will show that this is the case.

III. Econometric framework, data and descriptive statistics

A. Econometric framework

We first focus on the estimation of the effect of shopping externalities on rents of retail establishments. Let p_{ijt} be the rent paid by retail firm i at location j in year t. Furthermore, let f_{jt} be the footfall at each shop location j within a shopping street (defined later on) and z_{ijt} other shop and location characteristics (e.g. shop size, construction year, historic district etc.). The basic equation to be estimated yields:

$$\log p_{ijt} = \alpha \log f_{jt} + \gamma z_{ijt} + \theta_t + \epsilon_{ijt}, \tag{7}$$

where α and γ are parameters to be estimated, θ_t are year fixed effects and ϵ_{ijt} is an identically and independently distributed error term.

There are four major concerns when interpreting α as a causal estimate of shopping externalities. The *first* concern is that the estimated effect of footfall is causal, but that a location may also attract pedestrians that use the shopping street with no interest in shopping (non-shoppers). In particular, footfall levels are usually higher close to railway stations, because workers who commute by train may walk from the railway station to their work/home. Hence, if footfall is measured with error, it may not necessarily capture shopping externalities. This concern turns out to be minor because we use observations of footfall, which were collected on Saturdays for the main shopping streets of the Netherlands. For this sample of observations, almost all pedestrian movements are attributed to shopping. It is therefore very unlikely that any measurement error is

[16] Retail firms may also differ in the extent that they benefit from footfall $\partial I/\partial f$. However, the latter margin is not external to the decision to open a shop, so here we assume that firms are only heterogeneous with respect to the amount of footfall they generate for other firms.

[17] This estimated elasticity allows us to derive $\partial I/\partial f$, hence we can calculate the marginal benefit to a shop of a passing pedestrian.

substantial or systematic.[18] Even in the case of a non-systematic measurement error, the bias in our estimates is expected to be limited, as non-shoppers aim to avoid crowded shopping streets.[19]

The *second* concern is that the estimated effect of footfall is causal, but a location may also attract shoppers that use the shopping street to visit *one specific shop* with no interest in visiting other shops on the same street, so-called 'one-stop shoppers'. One-stop shoppers do not generate any shopping externality, although they may be included in our measure of footfall. Our identification strategy, which focuses on differences of footfall within very small areas addresses this issue. Any spatial difference in the share of one-stop shoppers would most likely lead to a bias in the estimated effect of the shopping externalities because of measurement error. We use an example to show that this measurement error is expected to be small even if the share of one-stop shoppers is substantial, because the probability that a one-stop shopper is included in our measure of footfall is (approximately) proportional to the number of shops visited. For example, if 25 percent of footfall were one-stop shoppers, and the other 75 percent visit four shops, then the proportion of one-stop shoppers would be only 7.8 percent.[20] Consequently, any measurement error because of one-stop shoppers is expected to be limited.

The *third* concern may rise because footfall data are collected only on two Saturdays per year as we will explain in detail later on. Thus, the annual measure of footfall may suffer from measurement error due to the random variation between different Saturdays of each year. Annual variation in our measure of footfall at the same location is thus likely to be substantial, even when actual annual variation in footfall is absent. Identification based on annual differences would lead to a downward bias in the estimated effect of footfall if this is the case. In contrast, spatial variation in our measure of footfall due to random sampling error is likely minimal, because different locations in close proximity are measured on the same day. Hence, identifying the effect of footfall using spatial variation in local footfall addresses such measurement error concerns.[21]

The *fourth* and main concern refers to the presence of unobserved location characteristics that are correlated with footfall. For example, building quality may be important for profits. When building quality is non-randomly distributed over space (e.g.

---

[18] On average, about 60 percent of all pedestrian movements in cities are attributed to shopping (and, for example, only 7.5 percent to commuting) (Statistics Netherlands). By focusing on Saturdays, our measure of pedestrians hardly includes any commuters.

[19] In a robustness check, we will show that by excluding observations close to train stations, our results remain robust.

[20] Furthermore, it is plausible that one-stop shoppers aim to avoid walking through busy shopping streets, and do not enter the shopping street at a random location, but from a side road which is close to the shop they want to visit. This makes it even more likely that one-stop shoppers are less than proportionally included in our measure of footfall.

[21] In a sensitivity analysis, we use the annual average of footfall, as well as the footfall of the previous year, as the main variable of interest. We obtain similar results.

nicer buildings in areas with more footfall) and customers value building quality, a naïve hedonic regression will suffer from bias. Also, zoning and other regulations may force retail firms to locate at more expensive locations with more footfall (Cheshire et al., 2015). When one does not account for characteristics that cause omitted variable bias, one is likely to overestimate the importance of shopping externalities.[22]

To control for unobserved locational endowments, we take a number of steps. First, we include shopping district or shopping street fixed effects, implying that we identify the differences in footfall within the shopping district or the shopping street, respectively. This approach mitigates the problem of unobserved endowments, but may not solve the problem entirely because shopping streets may be quite long (up to 1,269m). We therefore also propose another identification strategy using spatial variation in *local* footfall between intersecting streets.

Our idea is to compare shops that are very close to street intersections (e.g. within 100 or 50m). Locations close to intersections are arguably identical in unobserved spatial components, such as local policies, nearby parking etc. Let d_{jn} be the distance of shop at location j to the nearest intersection n and ϕ_n captures a set of intersection fixed effects, i.e. dummies that equal one when j is within $\bar{d}$ distance of intersection n. We then estimate:

$$\log p_{ijt} = \alpha \log f_{jt} + \gamma z_{ijt} + \phi_n + \theta_t + \epsilon_{ijt}, \quad \text{if } d_{jn} < \bar{d}. \tag{8}$$

One may argue that the estimate of footfall based on (8) may still suffer from omitted-variable bias, because intersecting streets may have different characteristics which are relevant for both footfall and rent. Hence, we have constructed a range of street and shop characteristics that we denote as x_{ij}. In particular, street width seems relevant, because smaller streets may restrict footfall and imply less visibility. We therefore calculate for each shop the distance to the opposite side of the street and we also include a dummy indicating whether a street is pedestrianised.

Another potential issue could be that corner shops have two shop windows in two different streets, therefore benefiting from pedestrians passing in both streets, whereas our measures refer to one street only.[23] Furthermore, shops located inside a shopping mall are expected to have different footfall and pay different rent than the shops located on the street.[24] Finally, one might also argue that a retail firm would pay a higher rent to be on the sunny side of the street while sun also attracts more pedestrians. We thus created dummy variables for corner shops, for shops inside a mall and for shops located on the sunny side of

[22] Different solutions have been proposed to address these endogeneity issues of agglomeration. Many studies use long-lagged instruments (Ciccone and Hall, 1996; Melo et al., 2009). However, there is extreme persistence of shopping streets over time. This makes it plausible that unobserved endowments that were important a century ago are still affecting current rents of shops. Hence, long-lagged instruments may be invalid in this setting.

[23] We also expect some measurement error in footfall at shopping street intersections.

[24] About 4 percent of our shop observations are 'inside malls', defined in the next Section. The results are identical when we exclude these observations (see Appendix C.5).

the street and include these additional shopping street and shop characteristics x_{ij} in the regression:

(9) $$\log p_{ijt} = \alpha \log f_{jt} + \beta x_{ij} + \gamma z_{ijt} + \phi_n + \theta_t + \epsilon_{ijt}, \quad \text{if } d_{jn} < \bar{d}.$$

where β are additional parameters to be estimated.

In the current paper, we will not only estimate the effect of log footfall on log rent, but also on whether the shop is vacant, indicated by v_{ijt}. We will then use the same approach as described above to address endogeneity issues. However, one may argue that there is reverse causality because footfall may be dependent on the vacancy rate in a neighbourhood. To address this issue, we use an insight provided by our theoretical framework (see Appendix A.2) and write $f_{jt} = (1 - v_{jt})\bar{f}_{jt}$, where v_{jt} is the vacancy rate in location j in year t and $\bar{f}_{jt}$ is the footfall generated if all shops in location j in year t are occupied. Therefore, in the preferred specification:

(10) $$\begin{aligned} v_{ijt} &= \alpha \log\left((1 - v_{jt})\bar{f}_{jt}\right) + \beta x_{ij} + \gamma z_{ijt} + \phi_n + \theta_t + \epsilon_{ijt}. \\ &= \alpha \log(1 - v_{jt}) + \alpha \log \bar{f}_{jt} + \beta x_{ij} + \gamma z_{ijt} + \phi_n + \theta_t + \epsilon_{ijt}, \quad \text{if } d_{jn} < \bar{d}. \end{aligned}$$

If shops within location j are identical and because v_{jt} is small, it holds that $\log(1 - v_{jt}) \approx -v_{ijt}$. This implies that:

(11) $$v_{ijt} = \frac{\alpha}{1 + \alpha} \log \bar{f}_{jt} + \beta x_{ij} + \gamma z_{ijt} + \phi_n + \theta_t + \epsilon_{ijt}, \quad \text{if } d_{jn} < \bar{d}.$$

When α is small (which appears to be the case), one immediately observes that $\alpha \approx \alpha/(1 + \alpha)$, so the problem of reserve causality does not seem to be important here.[25]

We will also estimate the effect of log number of shops in the shopping street – $\log N_{jt}$ – on $\log p_{ijt}$ and v_{ijt}. We will use an identical setup for the estimation of the effect of number of shops as for the estimation of footfall. We will note here a few differences. Given heterogeneity in the amount of footfall generated by shops, the rental income elasticity with respect to shops provides a downward bias of shopping externalities, but it is useful as an input to calculate the optimal subsidy to store owners (see (4)).

One subtle difference is that the estimation strategy with shopping street fixed effects cannot be applied because the number of shops does not vary within the shopping street. This is not problematic, because this is not our preferred estimation strategy. Note further that the number of shops is *not* subject to the first three concerns that apply to footfall discussed above, so that measurement error in number of shops seems less of an issue than measurement error in footfall.

B. Data

We base our empirical analysis on six datasets. The first one is obtained from *Strabo*, a consultancy firm that gathers commercial property data. It comprises transactions of

[25] We also address this issue by using footfall in the previous year(s) as the main variable of interest in the sensitivity analysis in Appendix C.5. The results do not change.

commercial properties provided by real estate agents from 2003 to 2015. The dataset contains information about annual rents and rental property attributes, such as address, size (gross floor area in m^2) and whether the building is newly constructed or renovated. From the Strabo dataset, we exclude observations for which no rent is reported. These observations comprise 27.8 percent of all shops in the Strabo dataset. The rental transactions are then matched to data from the *Administration of Buildings and Addresses*, which provides the exact location and construction year for all buildings in the Netherlands. Using a 5m distance threshold, we matched 72.9 percent of the Strabo shops. The distance between a shop location and the nearest building is zero for 90 percent of the matched shops. Based on the *Listed Building Register*, we have added information on whether the rental property is in an area that is assigned as a historic district. The latter is relevant since historic districts may attract tourists that are (not) interested in shopping. The dataset is also merged with detailed land use data from *Statistics Netherlands*. The latter data enable the calculation of distance to the nearest water body and to the nearest railway station.[26]

The fifth dataset is a retail dataset obtained from *Locatus*, which contains the *entire population* of retail establishments. For each retail establishment, we know whether the shop is vacant or occupied and the retail sector (when occupied), and whether a retail firm occupying a shop is part of a chain.

The *Locatus* dataset also provides 3,936 counts of footfall in all *main shopping streets* of the Netherlands from 2003 to 2015 (these shopping streets contain about 13.4 percent of all shops in the Netherlands). The annual footfall data, provided by *Locatus*, is the average footfall collected on two 'regular' Saturdays in Spring and two Saturdays in Autumn at four different hours of the day at many different locations close to shops in all main shopping streets of the Netherlands.[27] Using these measurements, *Locatus* calculates the average footfall per day, which represents the average number of shoppers per day. The footfall data are matched to all shops in the previously-defined shopping streets. Within each shopping street, the average distance between footfall measures is approximately 45m.

We have defined a *shopping street* as a continuous straight street (or slightly curved) based on manually created GIS polylines for all streets for which there is at least one location where footfall data are available. Using this definition, based on the above-discussed *Administration of Buildings and Addresses* dataset, we define 1,160 unique shopping streets.

Given the points of intersection between shopping streets, we calculated the distance from each shop to its *closest* shopping street intersection. We have also used information

[26] Water bodies in inner cities in the Netherlands are mainly (historic) canals, which often offer aesthetic value. Therefore, it is important to control for the attractiveness of such locations.

[27] 'Regular Saturdays' do not coincide with holiday periods (e.g. Easter) and are not preceded by bank holidays. Furthermore, on these days there is no heavy rainfall or other extreme weather conditions. The average distance of each shop to a measurement point is approximately 29m.

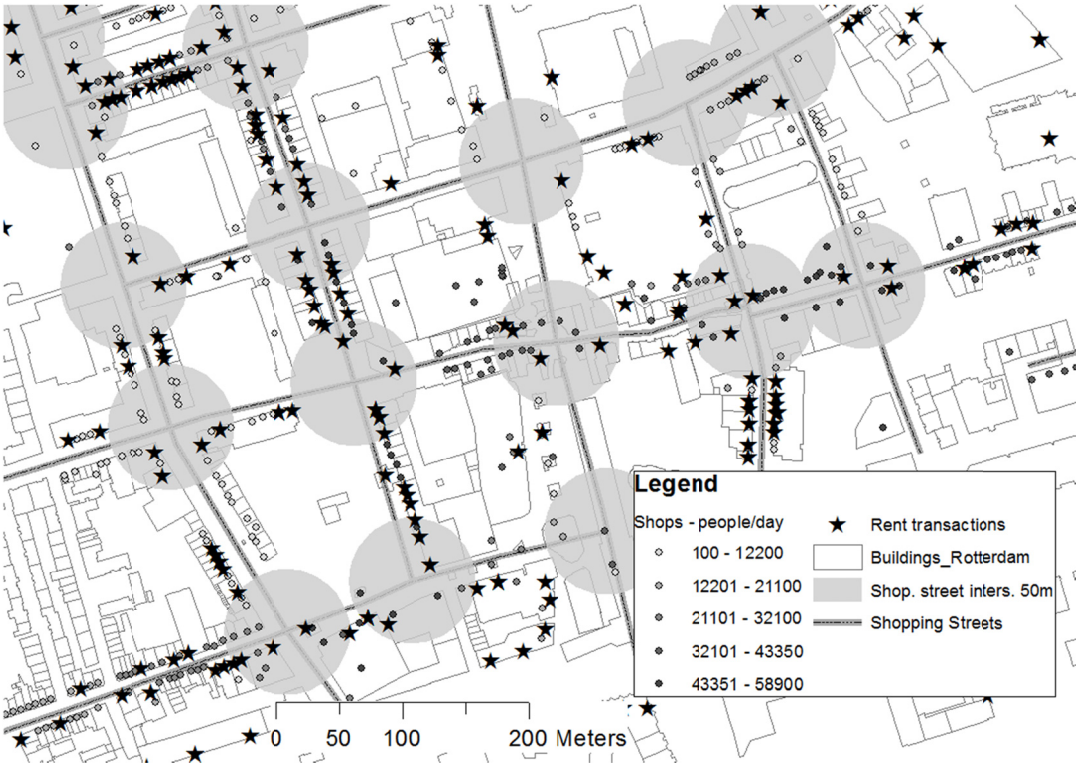

FIGURE 1 — SAMPLE MAP FOR THE ROTTERDAM CITY CENTRE

from *OpenStreetMap* in order to determine if a shopping street is pedestrianised. We also determine the street width, which is calculated using the average distance to the four closest non-contiguous buildings from the building in which each shop is located. We have set the minimum width at 3m, which applies to a few small alleys in historic districts. We also created a dummy variable for the shops located inside a mall, defined here as the shops which are in the interior of buildings. In addition, we have constructed a corner shop dummy variable for shops located within 10m from an intersection and a sunny side of the street dummy variable if the orientation of a shop is towards the south.[28] Finally, we used a distance threshold of 25m to match each Strabo shop to the nearest shopping street, as defined above.

We have also matched each rent transaction in the *Strabo* dataset to a shop in the *Locatus* dataset in order to recover more information on retail firm characteristics. Specifically, the matching is based on different combinations of building identifiers, the full shop names or the first letters of the names, the address numbers or the postal codes. For details see

[28] For corner shops, we also used alternative distance thresholds of 25m and 50m and the results are virtually unchanged.

Appendix B. It should be mentioned that our main results are not sensitive to this matching process.

We illustrate the data and identification strategy in Figure 1 based on a sample of our data for the city centre of Rotterdam. As it can be seen by the level of footfall in different locations, there is substantial spatial variation in the annual average of footfall both *within* shopping streets and *between* intersecting shopping streets. Moreover, rent transactions (the stars in the map) are numerous and cover almost the whole area that we have information on footfall.

C. Descriptive statistics

In this subsection, we present the descriptive statistics for the main variables that we include in our analysis. Our main dependent variable is the annual rental price. Table 1 summarises the descriptive statistics for the *Strabo* dataset. We have 3,102 rental transactions located on 682 different shopping streets with 831 shopping street intersections. We show that the rental price has a mean of €51,449. Footfall also exhibits substantial variation, which ranges from 200 to 79,000 pedestrians passing by a certain point each day. The mean daily footfall is 13,552 people with a standard deviation of 10,724. The majority of shops are relatively small, with a mean of 190m^2 and a median of 135m^2. Few shops are located inside a mall (3.7 percent) or on the corner of two shopping streets (3 percent) while about half the shops of our sample (48 percent) are located on the sunny side of the street. We also have information on the total building surface area and other building characteristics such as the construction year. About one percent of shops are either new or renovated when the rental transaction took place. 78 percent of shops are located in pedestrianised streets, about half the shops are in buildings constructed before the Second World War and a similar share is located within historic districts.

The average distance to the nearest train station is 1.2km, but the median distance only 747m (hence, there is good railway accessibility). Shopping street length ranges from approximately 44 to 1,270m with a mean and median of 431 and 359m, respectively. Street width is on average 8.1m. In each shopping street, there are about 70 non-vacant shops on average. In our data, we will also distinguish between 153 shopping districts (a shopping district contains about 260 shops, on average).

A substantial proportion of shopping districts (about 45 percent) are not within 5 km of the centre of a city.[29] Hence, in terms of shopping districts, we have a good representation of non-city centre shopping districts. However, the proportion of *shops* not within 5 km of the centre of a city is much smaller and only 23 percent because suburban shopping centres tend to be smaller.

[29] We define centres of all cities in the Netherlands with at least 50,000 inhabitants.

TABLE 1 — DESCRIPTIVE STATISTICS OF STRABO DATASET

	mean	sd	min	max
Rent *(€/year)*	51,449	73,163	4,800	2,700,000
Rent per m²	322.8	220.5	30	3,000
Footfall *(potential shoppers per day)*	13,552	10,274	200	79,000
Size of shop *(in m²)*	190.4	206.1	25	4,000
Building surface area *(in m²)*	1,275	4,490	20.44	86,771
Building – new	0.00387			
Building – renovated	0.00645			
Sublet property	0.00613			
Construction year < 1940	0.565			
Construction year 1940-1949	0.0193			
Construction year 1950-1959	0.0883			
Construction year 1960-1969	0.0516			
Construction year 1970-1979	0.0609			
Construction year 1980-1989	0.0645			
Construction year 1990-1999	0.0758			
Construction year ≥2000	0.0587			
Construction year missing	0.0161			
Mall	0.0374			
Corner shop	0.0297			
Sunny side of street	0.481			
Pedestrianised street	0.7795			
Shopping street length *(in m)*	430.8	270.5	43.92	1,269
Shopping street width *(in m)*	8.116	5.605	3	38.44
Number of (non-empty) shops in shopping street	70.705	51.521	2	227
Distance to nearest intersection *(in m)*	67.03	87.21	0.684	988.3
Water within 50m	0.0461			
Water 50-100m	0.0687			
In historic district	0.478			
Distance to station *(in m)*	1,206	1,928	65.97	18,280

Note: The number of observations is 3,102.

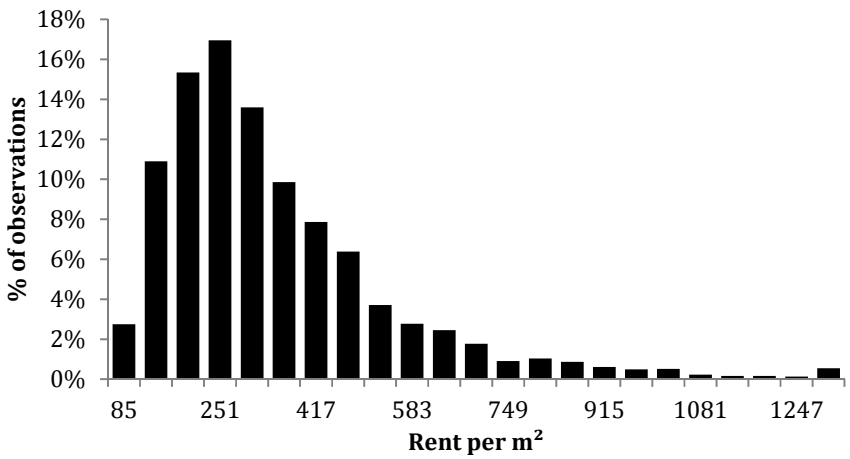

FIGURE 2 — RENT HISTOGRAM

TABLE 2 — DESCRIPTIVE STATISTICS OF LOCATUS DATASET

	mean	sd	min	max
Shop is vacant	0.0625			
Footfall	12,334	10,645	100	102,600
Size of shop *(in m²)*	175.3	1,078	1.637	27,694
Construction year < 1940	0.179			
Construction year 1940-1949	0.0136			
Construction year 1950-1959	0.0859			
Construction year 1960-1969	0.155			
Construction year 1970-1979	0.183			
Construction year 1980-1989	0.112			
Construction year 1990-1999	0.125			
Construction year ≥2000	0.147			
Mall	0.0612			
Corner shop	0.0249			
Sunny side of street	0.498			
Pedestrianised street	0.7296			
Shopping street length	403.7	244.9	21.09	1,269
Shopping street width *(in m)*	12.99	10.94	3	50
Number of (non-empty) shops in shopping street	110.5064	88	0	572
Distance to intersection *(in m)*	87.59	170.7	2.097	3,808
Water within 50m	0.0509			
Water in 50-100m	0.0743			
In historical district	0.393			
Distance to station *(in m)*	1,583	2,754	1.980	18,534

Note: The number of observations is 416,242.

Figure 2 shows an histogram of rent per m². About 57 percent of the observations are in the range €82-307, while the distribution of observations suggests that a logarithmic specification of rent is appropriate. Figure B1 and Figure B2 in Appendix B present the distribution of footfall and the number of shops per street, respectively.

In Table 2, we report descriptive statistics for shops in the *Locatus* dataset. We have 416,675 observations of shops in 161 shopping districts, 1,160 shopping streets and near 1,395 shopping street intersections. About 6 percent of the shops are vacant. The *Strabo* dataset contains a considerably higher share of shops in older buildings (particularly constructed before 1940) than the full population. The main explanation is likely to be that the *Strabo* dataset is based on rental *transactions*. Therefore, it is not a random sample of the population of shops because owned shops are not included and shops with long rental contracts are underrepresented., which suggests that our results for footfall may not extend to owned shops. The descriptive statistics of the location variables are however comparable to the descriptive statistics for the *Locatus* data.

Our sample of shops is clearly not a random sample of shops nationwide, as we focus on shops in shopping streets that presumably aim to benefit from footfall. Most shops in our

sample are clothing shops (29 percent) and therefore strongly overrepresented compared to the national average (about 8 percent). However, the share of restaurants and cafes, which is the second most common branch in our sample is fully representative for the full population of shops (16 percent in both our sample and in the whole population). Each of these branches typically comprises shops that sell close substitutes (while the two branches are complementary) although the degree of product differentiation in these branches is arguably high.

In Europe, shopping districts usually exhibit a pattern of mixed land uses. In line with this, using information from the *Administration of Buildings and Addresses* dataset for buildings within 25m of a shopping street, it appears that almost 50 percent of the properties is residential, about 25 percent for shopping and 25 percent for other purposes (e.g. offices, public services).

We mentioned in the introduction that store ownership (and therefore land ownership) of shops is fragmented. This observation is based on the *Strabo* dataset for which the store owner type is reported. We know the store owner's name for about one third of the observations. It appears that on average only 18 percent of the shops belong to store owners who own multiple properties in the same shopping street.[30] This evidence indicates that it is highly unlikely that the shopping externality that we measure is internalised. There is also information about store owner *type* that is available for about two thirds of the same sample. Store owner types are private-store owners, real estate agencies, pension funds, construction companies etc. We will use all this information in the sensitivity analysis.

Our main identification strategy is based on spatial differences between shopping street intersections. Our basic assumption is that in close proximity to an intersection, shops located on intersecting shopping streets have common unobservable characteristics (e.g. local amenities, accessibility to public transport, parking spaces etc.). Therefore, conditional on shop, location and street characteristics, we may identify the causal effect of shopping externalities.

We present here some graphical evidence. In essence, we show that in intersecting streets, footfall and rents depend on distance to the intersection *in a systematic and very similar way*, whereas e.g. the size of the shop – which is one of the main observed determinants of retail rent – does not systematically depend on this distance. We constructed 25m bins for the distance between each shop and the nearest intersection.

Negative distances denote shops located at low-footfall streets and zero distance denotes the intersection. We emphasise here that this is *not* a Regression-Discontinuity Design,

[30] Given that this dataset only contains rental transactions, it is likely that the percentage of multi-store owners is overrepresented in our sample, because this percentage is likely lower for owned shops. Moreover, only 32 percent of shops are owned by companies, so the large majority of shops are owned by individual private investors.

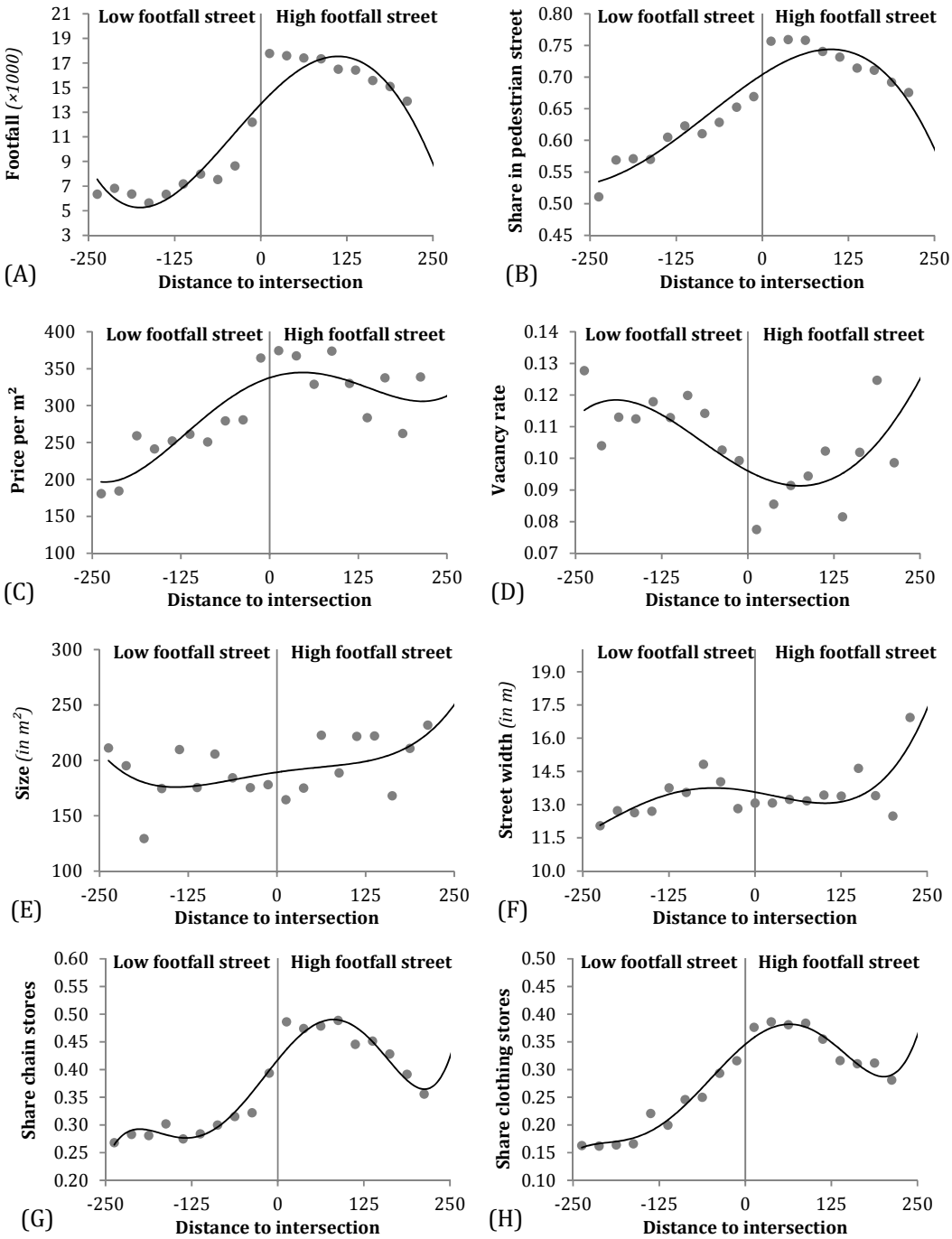

FIGURE 3 — VARIATION NEAR INTERSECTIONS

Notes: In Panels C and E we use *Strabo* data. In the rest of the panels, we use *Locatus* data. The spatial trend is estimated by a third-order polynomial of the variable of interest on the distance to the closest intersection.

because we do *not* need a discrete jump in footfall around street intersections but merely exploit the local variation in footfall close to these intersections. Hence, there should be considerable variation at the local level in the variables of interest. To construct these graphs, we exclude intersections where the difference in average footfall between two intersecting streets was minimal, i.e. below the first quartile of these differences (see similarly, Bayer et al. 2007). We then regressed footfall, shop size, retail rents per m^2 and other measures, on 25m bin dummies for observations within 250m of the intersection and a spatial trend, while including intersection fixed effects to control for unobserved characteristics that are common to the intersecting streets. These dummy variables can be interpreted as conditional means. These graphs also allow us to investigate whether shopping streets with high-footfall are distinctively different from low-footfall streets.

Figure 3 reports the results. In Panel A, it is shown that, by construction, footfall is considerably lower at the low-footfall street close to the intersection distance. Footfall is already higher close to intersections in the low-footfall street, which may be due to corner shops that have access to both streets. In Panel B, it can be seen that the share of pedestrianised streets is highly correlated with footfall, which is not too surprising. Later, we will show that if we control for pedestrianised streets, the impact of footfall is hardly affected.

In Panel C, Figure 3, we observe that there is also considerable variation in prices close to the intersection. For example, the annual price per m^2 is about € 280 in the low-footfall street, while in the high-footfall street is it about € 365, which is a considerable increase. The variation in vacancy rates is less clear-cut (Panel D), but we can still observe a lower average vacancy rate in the high-footfall street.

One may argue that high-footfall streets are distinctively different from low-footfall streets. We do not find strong evidence for this claim; both shop size and the street width do not show substantial variation around the intersection point (see Panels E and F). On the other hand, we find evidence for sorting; it seems that chain stores, which are often clothing stores, are located at the high-footfall streets (see Panels G and H). As we mentioned in the introduction, clothing shops are expected to benefit strongly from footfall as people searching for clothes often browse through shops and engage in trip-chaining. Therefore, clothing shops are expected to sort into high-footfall streets and pay a higher rent. Nonetheless, in a sensitivity analysis (Appendix C.4) we show that chain and non-chain stores have an identical preference for footfall. Moreover, the elasticity of rents with respect to footfall is unaffected once we control for the share of chain stores and share of clothing stores. Hence, this sorting is unlikely to drive our results.

IV. Results

The results section is structured as follows. We first discuss the effects of footfall and the number of shops in the shopping street on rents. In the second part, we focus on the effects of footfall on the probability of a shop to be vacant. We take these results together to

estimate the effects of footfall and number of shops on rental income. We close this section by discussing the welfare implications and derive the optimal subsidy.

A. Effects of footfall on rents

Table 3 reports the results of our baseline regressions. The specification in column (1) is an ordinary least squares (OLS) regression of the log rental price on log footfall, log size of the shop, building and location characteristics, in line with equation (7) in Section II. The elasticity of footfall with respect to rental price is 0.32. The coefficients related to shop attributes have the expected signs and magnitudes.[31]

The specification in column (1) might suffer from omitted variable bias due to the omission of unobserved features of a shop location that are correlated with footfall. For example, some shopping areas are more attractive due to their proximity to a museum, school or other neighbourhood-specific amenities. The relevance of such factors is suggested by the positive (and marginally statistically significant) coefficient of the historic district dummy. A partial solution to this problem is the inclusion of shopping district fixed effects in column (2), which may mitigate this endogeneity issue. Although the coefficient of footfall remains virtually unchanged, unobservable characteristics at a smaller spatial scale might still cause omitted variable bias. For this reason, we also include shopping street fixed effects in column (3), Table 3. In this specification, we essentially exploit the variation in footfall *within* shopping streets. Although the estimated standard error of the footfall coefficient in column (3) is substantially larger, the estimated coefficient is again essentially identical.

In our sample, the shopping street length is about 400m, on average, but can be more than 1km, suggesting that there may still be unobservable factors that vary within shopping streets that induce endogeneity issues. Examples include a small square with a fountain, a sculpture located in the middle of a street or a nice view. In order to deal with such factors, columns (4)-(6) exploit variation between shops that are within a given distance of the same intersection. In Column (4), we include fixed effects for the shop locations that are within a distance of 100m from an intersection, while in Column (5), we reduce the distance to the nearest intersection to 50m. It is plausible that shops located very close to an intersection are essentially identical, when we control for shop characteristics. Columns (4) and (5) show that when we use this identification strategy, the estimated coefficient for footfall is reduced to 0.23 and 0.22, respectively.

Finally, one could argue that even when we compare the rental prices of shops on the

[31] There is one exception; distance to the nearest station has a counterintuitive sign, because it is positively correlated with attractive unobserved features of location such as city size. Indeed, it becomes negative (but statistically significant) in the more believable specifications.

TABLE 3 — REGRESSION RESULTS FOR RETAIL RENTS
(Dependent variable: the log of rent)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Footfall *(log)*	0.322***	0.307***	0.306***	0.230***	0.215***	0.213***
	(0.0217)	(0.0181)	(0.0300)	(0.0270)	(0.0349)	(0.0350)
Size of shop in m^2 *(log)*	0.588***	0.607***	0.608***	0.628***	0.622***	0.622***
	(0.0186)	(0.0138)	(0.0177)	(0.0172)	(0.0243)	(0.0244)
Building surface area in m^2 *(log)*	0.0422***	0.0332***	0.0351***	0.0517***	0.0571***	0.0596***
	(0.00930)	(0.00879)	(0.0126)	(0.0127)	(0.0162)	(0.0196)
Building – new	0.00320	-0.0613	0.0617	0.0614	-0.0793	-0.0758
	(0.161)	(0.142)	(0.111)	(0.137)	(0.282)	(0.287)
Building – renovated	0.417***	0.334***	0.246***	0.209*	0.115	0.114
	(0.1000)	(0.0764)	(0.0750)	(0.107)	(0.0911)	(0.0910)
Sublet property	-0.0205	-0.00777	-0.0633	-0.105	-0.156	-0.150
	(0.0870)	(0.0760)	(0.0980)	(0.0991)	(0.0982)	(0.0966)
Shop is in mall						-0.0392
						(0.128)
Shop on the corner						0.0672
						(0.0536)
Shop is on sunny side of street						-0.0185
						(0.0261)
Shopping street width in m *(log)*						-0.00793
						(0.0454)
Pedestrianised street	0.0469	0.118***		0.0125	0.0519	0.0552
	(0.0358)	(0.0319)		(0.0501)	(0.0700)	(0.0701)
Water within 50m	-0.0323	-0.158***	-0.127*	-0.0122	0.0280	0.0265
	(0.0507)	(0.0525)	(0.0684)	(0.116)	(0.109)	(0.112)
Water 50-100m	0.0356	-0.0613*	-0.105**	0.0394	0.0368	0.0438
	(0.0553)	(0.0368)	(0.0424)	(0.0606)	(0.0831)	(0.0861)
In historic district	0.0676*	-0.0513	0.0613	0.0623	0.0305	0.0387
	(0.0366)	(0.0752)	(0.102)	(0.0695)	(0.0952)	(0.0967)
Distance to station *(log)*	0.0602***	0.0382	0.0193	-0.0373	-0.0363	-0.0378
	(0.0183)	(0.0445)	(0.0681)	(0.0598)	(0.0495)	(0.0477)
Construction year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—	—
Shopping street fixed effects	No	No	Yes	No	No	No
Intersection fixed effects	No	No	No	Yes	Yes	Yes
Observations	3,102	3,102	3,102	2,629	1,870	1,870
R^2	0.582	0.711	0.809	0.848	0.871	0.872

Notes: Footfall is measured as the number of shoppers per day. In column (4), we include observations within 100m of a shopping street interaction. In columns (5) and (6), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

different intersecting streets, street width may be an important omitted variable which could affect both footfall and rental prices.[32] The relationship between shopping street width and footfall may be mechanical, because street width puts an upper bound on footfall.

[32] Street width is shop-specific and therefore not captured by shopping street fixed effects.

TABLE 4 — REGRESSION RESULTS FOR RETAIL RENTS: NUMBER OF SHOPS
(Dependent variable: the log of rent)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS
Number of shops in street (*log*)	0.0646*** (0.0241)	0.0789*** (0.0171)	0.115*** (0.0223)	0.0981*** (0.0236)	0.100*** (0.0234)
Observations	3,102	3,102	2,629	1,870	1,870
R^2	0.466	0.637	0.836	0.863	0.864
Shop characteristics	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes
Shopping street characteristics	No	No	No	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—
Intersection fixed effects	No	No	Yes	Yes	Yes

Notes: The number of shops in street (*log*) is the logarithm of the number of non-vacant shops on the same street and in the same year that the rent transaction took place. In column (3) we include observations within 100m of a shopping street interaction. In columns (4) and (5) we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Moreover, shopping street width might affect the visibility of a shop, the supply of stock material, or the noise caused by pedestrians and cars in some cases. In column (6), Table 3, we include in addition to 50m intersection fixed effects, shopping width, a dummy if a shop is located on a corner, a dummy if the shop is on the sunny side of the street, the logarithm of shopping street width, and another dummy if a location is inside a mall. The estimated coefficient for footfall is highly statistically significant and its elasticity is 0.21, virtually the same as in column (5). This implies that if we increase log footfall by one standard deviation, the increase in rent is then roughly 12 percent (0.56×0.21).

Table 4 reports a similar set of specifications as in the previous table, while including the log number of shops in the shopping street and in the same year that the rent transaction took place.[33] Column (1) is the most parsimonious specification, column (2) includes shopping district fixed effects and column (3) includes 100m intersection fixed effects. We do not use street fixed effects because the number of shops in street variable does, by construction, not exhibit spatial variation within the shopping street. In columns (4) and (5), we restrict the sample to 50m from an intersections. The coefficient of log number of shops is statistically significant in all specifications. The coefficient of the log number of shops in column (5), which also includes shopping street characteristics, is 0.1. Thus, a 10 percent increase in the number of shops in a street causes a 1 percent increase in the rent of an

[33] In Appendix C.6, we estimate models were jointly include footfall and the number of shops in a street. Our results for both rents and vacancies indicate that conditional on footfall the effect of number of shops is negligible.

average shop. Although this number is substantial, it is less than half of the elasticity of rents with respect to footfall.

B. Effects of footfall on vacancy rates

In Table 5, we report the results for the incidence of a shop being vacant using a standard Logit model based on a similar set of specifications as in the previous subsection. The estimated marginal effect is shop-specific, so we report average marginal effects.[34] Column (1) is a naïve regression of a dummy variable indicating if a shop is vacant on log footfall, the log surface area of the building, construction year dummies, location attributes and year fixed effects. The average marginal effect of log footfall is −0.027.

The estimated effect is slightly higher (in absolute value) when we include shopping district or shopping street fixed effects in columns (2) and (3), respectively. In the last three columns, we focus on our preferred identification strategy where we only include observations close to intersections of shopping streets. In column (4), Table 5, we show that if we include fixed effects for shops within 100m from an intersection, the impact of footfall on vacancy rates is very similar. This effect is virtually the same once we reduce the distance bandwidth of the fixed effects to 50m from an intersection in column (5) and also essentially the same when we include shop and street characteristics in column (6).

Column (6), Table 5, is our preferred specification, which suggests that doubling footfall leads to a 1.9 percentage point reduction in vacancies. This reduction is about one third of the average vacancy rate. Thus, the effect of footfall on vacancies is substantial. An increase of one standard deviation in footfall leads to a drop in the vacancy rate of about 1.7 percentage points, roughly a quarter of the average vacancy rate. These results confirm our retail rents results. They suggest that the most attractive locations in terms of footfall have a lower probability to be vacant, in line with the idea that the opportunity cost of having an empty shop is higher for more expensive locations. We test for other explanations in Appendix C.3. For example, we test whether the effect of footfall on vacancy rates is only relevant in times of low demand, when for certain shops the marginal costs of providing shop space may be below the marginal benefits.

Table 6 presents the same set of specifications as Table 5, but replaces footfall by the log number of shops in the same shopping street instead of log footfall. The coefficient of the log number of shops is *positive* and significant in column (1). In column (2), where we include shopping district fixed effects, the coefficient of the number of shops becomes negative, albeit not significant. Columns (3) and (4) include 100m and 50m intersection fixed effects, respectively. In both columns, the coefficient of the number of shops in street is negative and statistically significant. In column (5), we also add shopping street characteristics. The

[34] The marginal effects for the sample averages of the included explanatory variables are very similar to the average marginal effect presented here.

TABLE 5 — REGRESSION RESULTS FOR VACANT SHOPS
(Dependent variable: dummy shop is vacant)

	(1) Logit	(2) Logit	(3) Logit	(4) Logit	(5) Logit	(6) Logit
Footfall *(log)*	-0.0278***	-0.0288***	-0.0311***	-0.0277***	-0.0279***	-0.0276***
	(0.00128)	(0.00115)	(0.00151)	(0.00136)	(0.00161)	(0.00163)
Building surface area in m² *(log)*	0.00273***	-0.000478	-7.14e-05	-0.000801	0.000842	0.000870
	(0.000929)	(0.000749)	(0.000718)	(0.000748)	(0.000936)	(0.000931)
Shop is in mall						0.00612
						(0.00604)
Shop on the corner						-0.00542
						(0.00416)
Shop is on sunny side of street						-0.000769
						(0.00189)
Shopping street width in m *(log)*						-0.00726***
						(0.00236)
Pedestrianised street	0.00898***	0.00678**	0.0456***	0.00470*	0.00626**	0.00566*
	(0.00297)	(0.00266)	(0.00371)	(0.00280)	(0.00290)	(0.00293)
Water within 50m	0.00542	0.0131***	0.0134***	0.0170*	0.0119	0.0119
	(0.00893)	(0.00469)	(0.00512)	(0.00927)	(0.0115)	(0.0113)
Water 50-100m	0.00132	0.00984***	0.00906**	0.00144	0.00506	0.00493
	(0.00424)	(0.00320)	(0.00399)	(0.00544)	(0.00758)	(0.00762)
In historic district	0.00450*	0.00400	0.0118	0.0171**	0.0255**	0.0235*
	(0.00261)	(0.00613)	(0.00733)	(0.00839)	(0.0128)	(0.0127)
Distance to station *(log)*	-0.00601***	-0.00301	-0.000493	0.000561	-0.00196	-0.00209
	(0.00120)	(0.00348)	(0.00456)	(0.00449)	(0.00315)	(0.00315)
Construction year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—	—
Shopping street fixed effects	No	No	Yes	No	No	No
Intersection fixed effects	No	No	No	Yes	Yes	Yes
Log-likelihood	-94305	-92232	-89570	-69467	-44021	-44005
Observations	425,834	425,834	421,204	338,099	220,049	220,049

Notes: Reported coefficients are average marginal effects. Footfall is measured as the number of shoppers per day. In column (4), we include observations within 100m of a shopping street interaction. In columns (5) and (6), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

estimated average marginal effect suggests that doubling the number of shops in an shopping street causes a 0.4 percentage point decrease in the vacancy rate on average (approximately a 6 percent decrease of the average vacancy rate). These results confirm that the number of shops has a direct effect on vacancy rates, albeit it substantially lower than the effect of footfall on vacancy rates.

TABLE 6 — REGRESSION RESULTS FOR VACANT SHOPS: NUMBER OF SHOPS
(Dependent variable: dummy shop is vacant)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS
Number of shops in street (*log*)	0.00464**	-0.000419	-0.00545***	-0.00489***	-0.00527***
	(0.00203)	(0.00107)	(0.00113)	(0.00130)	(0.00129)
Log-likelihood	-94,041	-92,184	-69,460	-44,514	-44,482
Observations	425,783	425,783	338,070	220,020	220,020
Shop characteristics	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes
Shopping street characteristics	No	No	No	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—
Intersection fixed effects	No	No	Yes	Yes	Yes

Notes: Reported coefficients are average marginal effects. The number of shops in street (*log*) is the logarithm of the number of shops on the same street and in the same year as each shop observation. Building, location and shopping street characteristics are mentioned in Table 4. In column (3), we include observations within 100m of a shopping street interaction. In columns (4) and (5), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

C. Shopping externalities: the effect of footfall on rental income

In Section II, we argued that shopping externalities are expected to capitalise into rental incomes of shop owners. Rental incomes are defined as the shop rent paid by retail firms multiplied by the share of the time that the shop is occupied. Given the effect of footfall on retail rents and vacancies that we estimated in Sections IV.A and IV.B, Table 7 provides the estimates for the effect of log footfall on log rental income. Following equation (2), we calculate this effect assuming different values of κ:

$$\frac{\partial \log I_{ij}}{\partial \log f_j} = \frac{\partial \log p_{ij}}{\partial \log f_j} - (1+\kappa)\frac{\partial v_{ij}}{\partial \log f_j}. \tag{12}$$

κ is a positive parameter that defines the relationship between advertising cost and rental price. We guesstimate κ to be equal to 0.4135, based on the costs of letting commercial space in the Netherlands, which is about 17.5 percent of the yearly rental value (Leurs, 2017).[35] Table 7 reports the estimated elasticity of rental income with respect to footfall, denoted by $\varepsilon_{I,f}$, based on the specifications listed in Table 3 and Table 5.

Let us now calculate the marginal effect of footfall on rental income using Table 7. Recall

[35] The advertising cost that a store owner incurs to find a new tenant are given by $(cost\ share \times p)/contract\ length$, which should be equal to $cv = \kappa pv$. From a small subset of the observations in *Strabo*, we know that the average contract length is 6.77 years. Furthermore, we know that the vacancy rate is on average 0.0625. Hence, $= cost/(v \times contract\ length) = 0.175/(6.771465 \times 0.0625) = 0.4135$. When fees are, let's say, only 5 percent, $\kappa = 0.118143$, while if fees are 25 percent, $\kappa = 0.590714$

TABLE 7 — THE ELASTICITY OF RENTAL INCOME WITH RESPECT TO FOOTFALL, $\varepsilon_{I,f}$

		(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Footfall *(log)*	$\kappa = 0.4135$	0.361***	0.348***	0.350***	0.269***	0.254***	0.252***
		(0.02175)	(0.01815)	(0.03005)	(0.02705)	(0.03495)	(0.03505)
	$\kappa = 0.1181$	0.353***	0.339***	0.341***	0.261***	0.246***	0.244***
		(0.02174)	(0.01814)	(0.03004)	(0.02704)	(0.03494)	(0.03504)
	$\kappa = 0.5907$	0.366***	0.353***	0.355***	0.274***	0.259***	0.257***
		(0.02176)	(0.01816)	(0.03006)	(0.02705)	(0.03496)	(0.03506)
Shop characteristics		Yes	Yes	Yes	Yes	Yes	Yes
Location characteristics		Yes	Yes	Yes	Yes	Yes	Yes
Shopping street characteristics		No	No	No	No	No	Yes
Year fixed effects		Yes	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects		No	Yes	—	—	—	—
Shopping street fixed effects		No	No	Yes	No	No	No
Intersection fixed effects		No	No	No	Yes	Yes	Yes

Notes: Footfall is measured as the number of shoppers per day. In column (4), we include observations within 100m of a shopping street interaction. In columns (5) and (6), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

that the average footfall on a typical Saturday is around 14,000, whereas average annual rent per m² for a shop is about € 300. In general, footfall on Saturday is roughly one fifth of weekly footfall (Locatus, 2006). Let us increase footfall by one pedestrian in each day of the year. The annual increase in rental income per m² is then approximately € 0.000025.[36] Consequently, the monetary benefit of one additional pedestrian passing a shop with an average size of almost 200m² is estimated to be about € 0.005.[37]

Column (1) includes shop and location characteristics, as well as year fixed effects. The estimated effect of footfall on rental income is between 0.353 and 0.366, depending on the value of κ we assume. Adding shopping district fixed effects in columns (2) and (3), respectively, has virtually no effect on the estimated coefficient. Columns (4)-(6) in Table 7 report the estimates of our main identification strategy, using intersection fixed effects. In columns (4) and (5), which include 100m and 50m intersection fixed effects, respectively, the estimated footfall coefficient decreases to 0.269 and 0.254 based on the most realistic value of κ. Finally, when we add shop and street characteristics in column (6), the elasticity of rental income with respect to footfall is 0.253.

[36] This number is the product of the log footfall coefficient of column (6), Table 7, (0.25) and the average annual rental income per m² (€ 323 per m²) divided by the product of the mean footfall (13,552), multiplied by 5 (because footfall on Saturdays is approximately one fifth of the weekly footfall) and by the number of weeks in a year (52).

[37] The order of magnitude of this result seems to make sense. Let us suppose that one out of hundred persons who pass a certain shop also enter that shop. Furthermore, assume that 25 percent of those who enter the shop also make a purchase and the profit per purchase is equal to €2. The marginal profit of footfall for a shop is then equal to 0.01×0.25× 2 = € 0.005.

In the above analysis we have focused on the effects of footfall on rents, as to estimate the elasticity of rental income with respect to footfall, $\varepsilon_{I,f}$, in order to estimate the shopping externality. Let us now focus on the effect of number of shops on rents in order to obtain the elasticity of rental income with respect to shops $\varepsilon_{I,N}$. We have argued that when shops differ in terms of generating footfall, then $\varepsilon_{I,f}$ exceeds $\varepsilon_{I,N}$. Table 8 replicates Table 7 using the estimated effects of the number of shops in each shopping street on rents and vacancies. Column (5), which is our more conservative estimate, reports an elasticity of rental income with respect to the number of shops on the street of about 0.107. Hence, $\varepsilon_{I,f}$ exceeds $\varepsilon_{I,N}$ by a factor of two. This indicates that footfall as a proxy for shopping externalities is superior to the use of the number of shops. This will be confirmed later on in the sensitivity analysis.

D. Welfare analysis

We have shown that shopping externalities are important – the shop's marginal benefit of a passing pedestrian is about €0.005 – and argued that due to fragmented ownership of shops, it is implausible that shopping externalities are internalised. We will now examine policies that either subsidise store owners or specific retail firms.

We have demonstrated that the elasticity of rental income with respect to number of shops $\hat{\varepsilon}_{I,N}$ is about 0.1 (see Table 8). Using this elasticity, we calculate the optimal subsidy for store owners, which should be equal to the marginal *external* benefit of one additional shop. Equation (4) implies that the optimal subsidy to store owners is about 10 percent of the rental income. The optimal annual subsidy per shop is about €5 thousand. Note that larger shopping streets must receive higher levels of subsidies because there are more shops and because rental income levels are higher. In our data, it turns out the 25 largest shopping streets (of the population of about thousand shopping streets) must receive about 25 percent of all national subsidies.

In this back-on-the-envelope calculation, we assumed that $\varepsilon_{I,N}$ is a constant. This is not very realistic. To examine this further, in Appendix C.2, we estimate the effects of log footfall on log rents and vacancy rates using quadratic specifications of log number of shops to allow for non-linear effects of this variable. These estimated effects indicate that $\varepsilon_{I,N}$ is increasing in number of shops. This result suggests that optimal subsidy levels to shop owners may even be higher than suggested by the above calculation for shops located in large shopping streets because the elasticity is increasing in number of shops.

As argued in Section II, shops are heterogeneous in the amount of footfall they generate for other shops, implying that $\varepsilon_{f,N}$ differs between shops. Hence, if local governments have information about the footfall generated by a shop, they may subsidise these shops, rather than store owners. Using equation (6) and our estimate of $\hat{\varepsilon}_{I,f} = 0.25$, the subsidy given to a specific shop should be larger than 10 percent of the rental income if for this shop it is true that $\varepsilon_{f,N} > 0.4$. This result suggests that substantial subsidies for certain 'anchor' stores that

TABLE 8 — THE ELASTICITY OF RENTAL INCOME WITH RESPECT TO THE NUMBER OF SHOPS, $\varepsilon_{I,N}$

		(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS
Footfall *(log)*	$\kappa = 0.4135$	0.058***	0.079***	0.123***	0.105***	0.107***
		(0.02419)	(0.01713)	(0.02233)	(0.02364)	(0.02344)
	$\kappa = 0.1181$	0.059***	0.079***	0.121***	0.104***	0.106***
		(0.02419)	(0.01713)	(0.02233)	(0.02364)	(0.02344)
	$\kappa = 0.5907$	0.057***	0.080***	0.124***	0.106***	0.108***
		(0.02419)	(0.01713)	(0.02233)	(0.02364)	(0.02344)
Shop characteristics		Yes	Yes	Yes	Yes	Yes
Location characteristics		Yes	Yes	Yes	Yes	Yes
Shopping street characteristics		No	No	No	No	Yes
Year fixed effects		Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects		No	Yes	—	—	—
Shopping street fixed effects		No	No	No	No	No
Intersection fixed effects		No	No	Yes	Yes	Yes

Notes: Footfall is measured as the number of shoppers per day. In column (3), we include observations within 100m of a shopping street interaction. In columns (4) and (5), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

attract a lot of footfall may be welfare improving. For example, suppose that a shopping street consists of hundred small shops, and a large retailer, e.g. a warehouse, threatens to leave the shopping street, reducing footfall by 20 percent. Thus, $\varepsilon_{f,N}$ is about 1.2. In this case, it is efficient to provide subsidies of about 30 percent of rental expenditure to the shop.

V. Extensions and sensitivity

A. Extensions

Until now, we have argued that shopping externalities, as measured by footfall, have a substantial positive effect on the retail market. However, the effect of footfall might well extend beyond this market. It has been argued that retail dispersion towards the suburbs may lead to the 'hollowing-out' of city centres, where shops were traditionally concentrated (Sanchez-Vidal, 2016). It could thus be argued that footfall may increase the liveability and the attractiveness of city centres. In her magnum opus, Jane Jacobs argues that *"the sidewalk must have users on it fairly continuously, both to add to the number of effective eyes on the street and to induce the people in buildings along the street to watch the sidewalks in sufficient numbers. […] Large numbers of people entertain themselves, off and on, by watching street activity"* (Jacobs, 1961). On the other hand, we are aware of examples that residents raised opposition to new retail developments next to shopping streets, suggesting that a high retail concentration may also cause negative effects for the residents (e.g. through increased traffic or noise). In Appendix C.1, we test for the existence of such external effects by analysing the effect of footfall on residential housing prices and we find no effect of the

'people on the street'. This result suggests that the *net* effect of footfall externalities on residents is zero. More details can be found in this Appendix.

In Appendix C.2 we show the effects of log footfall on log rents and vacancy rates using quadratic specifications of log footfall to allow for non-linear effects of the logarithm of footfall. These results indicate that $\varepsilon_{I,f}$ is increasing in footfall. For example, for footfall levels three times above the average the elasticity of rental income is about 50 percent higher than the average elasticity.[38] We have also analysed the effects of log number of shops on log rents and vacancy rates using quadratic specifications of log number of shops. These results also indicate the rental income elasticity is increasing in number of shops (because the effect on log rent is increasing in log number of shops, whereas the effect on vacancies is a constant).

As mentioned in Section IV.B, there might be alternative explanations that explain the effect of footfall on retail vacancy rates. One such explanation is that the effect of footfall on vacancy rates is only relevant in times of low demand, when for certain shops, the marginal costs of providing shop space are below the marginal benefits. On the other hand, in times of high demand, for almost all retail establishments, marginal costs are lower than the marginal benefits, thus, the effect of footfall on vacancy rates could be negligible. We test this hypothesis in Appendix C.3 where we regress the dummy for a vacant shop on the interaction term between log footfall and a dummy variable for the recent boom and bust period of the Dutch economy, respectively. Our results show that the effect of footfall on vacancy rates is significantly different in the boom and bust periods. Specifically, the effect in bust periods is higher, as expected. However, the effect of footfall on vacancy rates is still economically and statistically significant during the boom years. This results confirms that higher rents increase the opportunity cost of having an empty shop, so that vacancy rates are lower in more attractive areas (i.e. those with a higher footfall).

Policies that foster retail concentration can be welfare improving only if shopping externalities are not fully internalised. We also argued that the highly fragmented property ownership that we find in our sample implies that internalisation is unlikely to occur.[39] In Appendix C.4, we test whether the effect of footfall on retail rents is capitalised differently in properties that belong to store owners who possess multiple rental properties on the same shopping street (multi-property owners). For multi-property owners, the externality seems

[38] Allowing for non-linearity of the logarithm of footfall is also important when studying other interactions. For example, using a linear specification, we obtain very different effects for central city and suburban shops. Because of large differences in footfall between the city centre and suburbs, this difference turns out to be explained by the non-linear elasticity of footfall with respect to rental income. Moreover, using a non-linear specification, we find that the effect of log footfall on retail rents is identical for pedestrianised and non-pedestrianised streets and that pedestrianised streets have no direct effect on retail rents. These results are in line with our assumption that footfall captures the economic value of the shop's location.

[39] In our sample, only 18 percent of properties belong to owners who possess multiple properties on the same street.

to capitalise in rents in the same way as for single property owners. In addition, we do not find any difference between commercial property owners (real estate companies, construction companies etc.) and private property owners. Furthermore, in Appendix C.5 we show that the effect is the same for shops that are part of a retail chain and for non-chain shops. Also, controlling for the share of chain stores or clothing stores in the same street leaves the footfall elasticity unaffected.

B. Sensitivity

In order to establish the causal relationship of footfall on retail rents, vacancies – and therefore rental income, we have estimated alternative specifications to address the main identification concerns that might disparage the validity of our results. We provide here a summary of the main analyses. More details can be found in Appendix C.

A first concern with our identification strategy is that our main identification assumption (i.e. that shops located in close proximity from two intersecting streets have similar unobserved characteristics) might not hold. If the two intersecting streets differ in unobserved characteristics that affect both footfall and retail rents, our preferred estimates would be biased. In Appendix C.5, we address this concern by focusing on local differences in footfall between neighbouring shops (within 50m from an intersection) that are located on the *same* shopping street. An alternative way to control for street-specific local endowments is to use fixed effects for the 6-digit postal code (PC6) of each shop. Both these sensitivity checks supports our main identification strategy, suggesting that it is highly unlikely that our estimates suffer from omitted variable bias. As mentioned in Section II.C, another possible concern could be reverse causality. We then use log footfall in the previous year instead. We also address reverse causality concerns by using the average of the logarithm of footfall over the time period instead of the logarithm of annual footfall. Moreover, using the logarithm of the average annual footfall, we mitigate the measurement error in footfall due to the random variation between different Saturdays of each year at the same location. Using lagged footfall, or the average footfall over the study period leads to highly statistically significant effects. The effect is actually slightly higher compared to the baseline results. Another concern we raised is that shops located close to train stations and those located inside a mall may be very different from shops located in ordinary shopping streets. We therefore exclude observation located less than 1km away from a train station and the shops that are considered to be inside a mall, respectively. The estimated coefficients are hardly different from our main estimates. We also run the same general robustness checks for the vacancy analysis and the results are roughly unchanged. Moreover, using a Linear Probability Model instead of a Logit model, the estimates of the effect of footfall on vacancy rates are very similar.

In Appendix C.6 we both include our proxies for shopping externalities: footfall and the number of shops in a shopping street. We show that conditional on footfall the effect of

number of shops is negligible, suggesting that using footfall as a proxy for shopping externalities is superior to the use of the number of shops.

We have argued that footfall can be used as a measure of shopping externalities. In Section IV we have shown that also the number of shops in a shopping street has a meaningful effect on retail rents. In Appendix C.7, we investigate the spatial scope of this question by adding the (log) number of shops that are located on the same street within different distance thresholds (e.g. 100m, 200m) leading to very similar results.

VI. Conclusions

The findings of this paper add to our understanding of retail clustering and shopping externalities. Economic theory indicates that *(i)* shop rents depend *positively* on footfall and *(ii)* vacancies depends *negatively* on footfall. Hence, the effect of footfall on rental income – the shop rent multiplied with the share of the time that the shop is occupied – is positive.

Our empirical estimates for the main shopping streets of the Netherlands show that the effects of footfall on retail rents and vacancies are substantial. We have shown that shopping externalities are important – the shop's marginal benefit of a passing pedestrian is about €0.005 – and we argued that due to fragmented ownership of shops, it is implausible that shopping externalities are internalised. Shopping externalities are therefore crucial to retail location choices, as higher footfall leads to higher rental income, with an elasticity of approximately 0.25. Such a high elasticity is consistent with the notion that the main reason for shops to cluster is the presence of positive shopping externalities. Our results are robust to different identification strategies including the use of local variation in footfall around street intersections and an extensive set of control variables.

Our analysis highlights the fundamental heterogeneity of shops in their ability to attract customers to shopping streets and therefore to generate positive shopping externalities for other shops. Consistent with that, we show that employing the number of shops in the shopping street, rather than footfall, generates a strong downward bias in the estimate of the shopping externality.

In contrast to shopping malls, we demonstrate that it is very implausible that the positive externalities in shopping streets are internalised due to high fragmented ownership of shops. To formulate our policy recommendations, we derive *(i)* the shop's marginal benefit of a pedestrian passing by, *(ii)* the optimal subsidy *to store owners* as an incentive to provide more retail space in an existing shopping street, and *(iii)* the optimal subsidy to retail firms depending on footfall generated by them.

The optimal subsidy to store owners is about 10 percent of rental expenditure, about €5,000 per year. This finding may explain for example why many cities around the world implement parking policies that explicitly subsidise short-term parking aimed at shoppers (particularly in the US and Australia). The subsidy to specific retail firms is even higher for retail firms that generate high levels of footfall for nearby shops. This finding explains why

many local governments, such as the municipality of Amsterdam, provide (financial) incentives to retail firms that are known to attract footfall (e.g., famous clothing brands, Apple Stores).

We are aware that explicit subsidies to private firms are controversial and may even be illegal. However, current policy practices in many countries around the world to pedestrianise certain streets in attractive city centres and the provision of subsidised public transport or parking space close to shopping streets are an implicit subsidy to enhance shopping externalities. From the perspective of the retail market, those policies seem welfare improving.

References

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2016). The Economics of Density: Evidence from the Berlin Wall. *Econometrica*, *83*(6), 1217–2189.

Arcidiacono, P., Bayer, P., Blevins, J. R., and Ellickson, P. B. (2016). Estimation of Dynamic Discrete Choice Models in Continuous Time with an Application to Retail Competition. *The Review of Economic Studies*, *83*(3), 889–931.

Arzaghi, M., and Henderson, J. V. (2008). Networking off Madison Avenue. *Review of Economic Studies*, *75*(4), 1011–1038.

Bayer, P., Ferreira, F., and McMillan, R. (2007). A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy*, *115*(4), 588–638.

Bertrand, M., and Kramarz, F. (2002). Does Entry Regulation Hinder Job Creation? Evidence from the French Retail Industry Does Entry Regulation Hinder Job Creation? *Quarterly Journal of Economics*, (3), 1369–1413.

Brueckner, J. K. (1993). Inter-store externalities and space allocation in shopping centers. *The Journal of Real Estate Finance and Economics*, *7*(1), 5–16.

Cheshire, P. C., Hilber, C. A. L., and Kaplanis, I. (2015). Land Use Regulation and Productivity-Land Matters: Evidence from a UK Supermarket Chain. *Journal of Economic Geography*, *15*(1), 43–73.

Cheshire, P., Hilber, C., and Kaplanis, A. . (2011). Evaluating the effects of planning policies on the retail sector: or do town centre first policies deliver the goods? *Journal of Economic Literature*, (ISSN 1362-3761), 25–28.

Ciccone, A., and Hall, R. E. (1996). Productivity and the Density of Economic Activity. *The American Economic Review*, *86*(1), 54–70.

Clapp, J. M., Ross, S. L., and Zhou, T. (2016). Retail Agglomeration and Competition Externalities: Evidence from Openings and Closings of Multiline Department Stores in the US. *Journal of Business and Economic Statistics*, Forthcoming.

Claycombe, R. J. (1991). Spatial Retail Markets. *International Journal of Industrial Organization*, *9*(2), 303–313.

Cushman and Cushman. (2011). *Upturn expected in shopping centre development in Europe.*

D'Aspremont, C., Gabszewicz, J. J., and Thisse, J. F. (1979). On Hotelling's "Stability in Competition." *Econometrica*, *47*(5), 1145–1150.

Davis, P. (2011). Spatial Competition in Retail Markets: Movie Theaters. *RAND Journal of Economics*, *37*(4), 964–982.

Eaton, B. C., and Lipsey, R. G. (1982). An Economic Theory of Central Places. *Economic Journal*, *92*(365), 56–72.

Glaeser, E. L., Kolko, J., and Saiz, A. (2001). Consumer City. *Journal of Economic Geography*, *1*(1), 27–50.

Gould, E. D., Pashigian, B. P., and Prendergast, C. J. (2005). Contracts, Externalities, and Incentives in Shopping Malls. *Review of Economics and Statistics*.

Haskel, J., and Sadun, R. (2012). Regulation and UK Retailing Productivity: Evidence from

Microdata. *Economica*, *79*(315), 425–448.
Hosios, A. J. (1990). Factor Market Search and the Structure of Simple General Equilibrium Models. *Journal of Political Economy*, *98*(2), 325.
Jacobs, J. (1961). *The Death and Life of Great American Cities* (Vol. 71). New York: Random House.
Jia, P. (2008). What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry. *Econometrica*, *76*(6), 1263–1316.
Johansen, B. O., and Nilssen, T. (2016). The Economics of Retailing Formats: Competition Versus Bargaining. *The Journal of Industrial Economics*, *64*(1), 109–134.
Konishi, H., and Sandfort, M. T. (2003). Anchor stores. *Journal of Urban Economics*, *53*(3), 413–435.
Koster, H. R. A., Van Ommeren, J. N., and Rietveld, P. (2014). Agglomeration Economies and Productivity: A Structural Estimation Approach using Commercial Rents. *Economica*, *81*(321), 63–85.
Leurs, R. (2017). *Wat Kost het Verhuren van uw Bedrijfspand?*
Locatus. (2006). *Amsterdam-centrum Kalverstraat 2006, Winkelpassantentellingen*. Woerden.
Marshall, A. (1920). *Principles of Economics*. London: MacMillan and Co.
Melo, P. C., Graham, D. J., and Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics*, *39*(3), 332–342.
Netz, J. S., and Taylor, B. a. (2002). Maximum or Minimum Differentiation? Location Patterns of Retail Outlets. *Review of Economics and Statistics*, *84*(1), 162–175. http://doi.org/10.1162/003465302317331991
Nunn, N., and Puga, D. (2012). Ruggedness: The Blessing of Bad Geography in Africa. *Review of Economics and Statistics*, *94*(1).
Osborne, M. J., and Pitchik, C. (1987). Equilibrium in Hotelling's Model of Spatial Competition. *Econometrica*, *55*(4), 911–922.
Pashigian, B. P., and Gould, E. D. (1998). Internalizing Externalities: the Pricing of Space in Shopping Malls1. *The Journal of Law and Economics*, *41*(1), 115–142.
Sanchez-Vidal, M. (2016). *Small Shops for Sale! The Effects of Big-box Openings on Grocery Stores.* (IEB working paper).
Schivardi, F., and Viviano, E. (2011). Entry Barriers in Retail Trade. *Economic Journal*, *121*(551), 145–170.
Schulz, N., and Stahl, K. (1996). Do Consumers Search for the Highest Price? Oligopoly Equilibrium and Monopoly Optimum in Differentiated-Products Markets. *RAND Journal of Economics*, *27*(3), 542–562.
Seim, K. (2006). An Empirical model of Firm Entry with Endogenous Product-type Choices. *RAND Journal of Economics*, *37*(3), 619–640.
Zhou, J. (2014). Multiproduct search and the Joint Search Effect. *The American Economic Review*, *104*(9), 2918–2939.

Appendix A. Theoretical model

A.1 A search model of a shopping street

Let us introduce a search model of a shopping street with two types of homogeneous agents. Store owners that possess properties and retail firms, which rent properties from store owners. Store owners with vacant properties and retail firms searching for space have to search for each other. Property owners set the level of advertising expenditure which determines the contact rate with retail firms. Given a contact, the agents use Nash bargaining to determine the rent level. We assume steady-state and a given number of store owners N, which possess one shop each, which they aim to rent out to retail firms for rent p. For simplicity, the revenue of a shop is equal to footfall in the street. For now, we assume that the number of shops and footfall are exogenous. The future is discounted at rate r. Owners and retail firms maximise their profits.

Retail firms go bankrupt at a given rate δ, which creates vacant properties. Owners with a vacancy and retail firms searching for shop space search for each other. The rate at which they find each other is defined by a concave matching function m. This matching function depends positively on the overall advertising expenditures, ev, i.e. the number of vacant shops v times advertising expenditure e per property. Thus, $m = m(ev)$. Vacant shops become occupied at a rate $q(v, e)$, defined by $m(ev)/v$. This rate depends negatively on v, due to the concavity assumption of the matching function. Owners with a vacancy incur advertising costs $c(e)$. Advertising costs are an increasing convex function of advertising expenditure, whereas $c(0) = 0$. When an owner with a vacancy and a searching retail firm meet each other, they bargain about the shop price p, given a bargaining parameter β, where $0 < \beta < 1$. Rental income of the property owner is equal to $p\,(1 - v) - cv$.

The market for retail firms is competitive with free entry of searching retail firms, so the expected profit of searching retail firms is equal to zero. Store owners with vacancies choose their advertising expenditure conditional on the advertising expenditure of other property owners. We consider symmetric equilibria where owners choose the same advertising expenditure. The latter implies that for the representative owner, the marginal increase in the matching rate of advertising expenditure is equal to the average rate, so $\partial m/\partial e = m/e$. Similarly, $\partial m/\partial v = m/v$.

In steady-state, the inflow rate of shops is equal to the outflow rate, implying that:

(A1) $$m(ev) = \delta\,(1 - v).$$

The present-discounted value of expected profits of a vacant shop, V, can be written as:

(A2) $$rV = -c(e) + \frac{m(ev)}{v}(R - V),$$

where R denotes the present discounted value of expected profits of a shop that is rented out. The latter can be written as:

(A3) $$rR = p + \delta(V - R).$$

The present-discounted value of expected profits for a retail firm equals:

(A4) $$rS = f - p - \delta(S - Q).$$

Retail firms that yet did not find a store to locate in have the following present-discounted profits Q:

(A5) $$rQ = -z(\eta) + \lambda(S - Q).$$

where $z(\eta)$ are search costs and η is search effort of retail firms and λ indicates the chance that a retail owner finds a store. Because of a competitive market, η is chosen optimally and Q will be equal to zero.

Nash bargaining implies that the store owners' share β of their own surplus, $R - V$, is equal to the retail firms' share, $(1 - \beta)$, of their own surplus S. Consequently:

(A6) $$(1 - \beta)S = \beta(R - V).$$

These four equations, combined with the first-order condition of (A2) that the present-discounted value of expected profits of a vacant shop is maximised with respect to advertising expenditure $c(e)$, imply that in equilibrium, p, v and e are determined by the following three equations:

(A7) $$p = \frac{f(1-\beta)(v(r+\delta) + m(ev)) - (r+\delta)v\beta c(e)}{(1-\beta)m(ev) + v(r+\delta)},$$

(A8) $$v = 1 - \frac{m(ev)}{\delta},$$

(A9) $$c'(e) = \frac{(1-\beta)\big(f + c(e)\big)m(ev)}{erv + e\delta(1-(1-v)\beta)}.$$

We are interested in the effects of footfall on prices and vacancy rates. Using (A7), it is easy to see that the partial derivative $\partial p/\partial f > 0$. Although interesting, we are mainly interested in general equilibrium effects on prices and vacancy rates, taking into account the effects through changes in advertising expenditure. We formulate the following proposition:

> PROPOSITION 1: In equilibrium, *(i)* shop price depends *positively* on footfall and *(ii)* the number of vacancies depends *negatively* on footfall.

Proof. We first derive V, R, S and p by solving the system of equations (A2), (A3), (A4) and (A6). This leads to:

(A10) $$V = \frac{(1-\beta)mf - (r+\delta)cv}{r\big((1-\beta)m + (r+\delta)v\big)},$$

(A11) $$R = \frac{f(m+rv)(1-\beta) - (r\beta+\delta)cv}{r\big((1-\beta)m + v(r+\delta)\big)},$$

(A12) $$S = \frac{(c+f)v\beta}{m(1-\beta) + v(r+\delta)},$$

(A13) $$p = \frac{f(1-\beta)(v(r+\delta)+m) - (r+\delta)v\beta c}{(1-\beta)m + v(r+\delta)}.$$

Note that $m = m(ev)$ and $c = c(e)$.

First, we are interested in the effect of footfall on rents, so $\mathrm{d}p/\mathrm{d}f$. We then use equations (A7), (A8) and (A9), and using implicit differentiation. According to Cramer's rule, $\mathrm{d}p/\mathrm{d}f = \det(Z_p)/\det(Z)$, where:

$$\text{(A14)} \qquad Z = \begin{pmatrix} 1 & \dfrac{(1-v)(1-\beta)(r+\delta)(f+c)\beta\delta v}{e\left(rv+\delta\left(1-\beta(1-v)\right)\right)^2} & 0 \\ 0 & \dfrac{m}{\delta e} & 1+\dfrac{m}{\delta v} \\ 0 & -\dfrac{(1-\beta)mc'}{erv+e\delta\left(1-\beta(1-v)\right)}+c'' & -\dfrac{(1-\beta)^2(f+c)\delta m}{ev\left(rv+\delta\left(1-\beta(1-v)\right)\right)^2} \end{pmatrix}.$$

Note that $c' = \partial c/\partial e$ and $c'' = \partial^2 c/\partial e^2$. To obtain Z_p we replace the first column of Z with:

$$\text{(A15)} \qquad z = \begin{pmatrix} \dfrac{(1-\beta)(rv+\delta)}{rv+\delta\left(1-\beta(1-v)\right)} \\ 0 \\ \dfrac{(1-\beta)m(e)}{erv+e\delta\left(1-\beta(1-v)\right)} \end{pmatrix}.$$

To obtain $\det(Z_e)$ and $\det(Z_v)$, we replace respectively the second and third column of Z with z. We then take into account that $m = \delta(1-v)$ and use (A9) to obtain:

$$\text{(A16)} \qquad \frac{\mathrm{d}p}{\mathrm{d}f} = \frac{\det(Z_p)}{\det(Z)} = \frac{(1-\beta)(rv+\delta)}{\Delta} > 0,$$

$$\text{(A17)} \qquad \frac{\mathrm{d}e}{\mathrm{d}f} = \frac{\det(Z_e)}{\det(Z)} = \frac{(1-v)(1-\beta)\delta}{\Delta e c''} > 0,$$

$$\text{(A18)} \qquad \frac{\mathrm{d}v}{\mathrm{d}f} = \frac{\det(Z_v)}{\det(Z)} = -\frac{(1-v)^2(1-\beta)\delta v}{\Delta e^2 c''} < 0.$$

where $\Delta = rv + \delta\left(1-\beta(1-v)\right)$. Because $\beta < 1$, the impact of footfall on rents is positive, $\mathrm{d}p/\mathrm{d}f > 0$. Furthermore, because the cost function is convex, we have $c'' > 0$, so that $\mathrm{d}e/\mathrm{d}f > 0$. This implies that advertising expenditure will increase when footfall is higher. Consequently, when advertising expenditure increases, the matching rate will also increase implying that $\mathrm{d}v/\mathrm{d}f < 0$ (see equation (A8)), which is confirmed by equation (A18) . ■

The model implies that the marginal effect of footfall on prices is positive, but always smaller than or equal to the marginal revenue which is equal to one (when $\beta = 0$, so when retail firms have all bargaining power, then $\mathrm{d}p/\mathrm{d}f = 1$). The intuition for the result that $\mathrm{d}v/\mathrm{d}f < 0$ is that store owners' opportunity cost of not filling a vacant shop increases with footfall.

A.2 Endogenous footfall

Until now, we assumed that footfall is exogenous. However, footfall likely depends negatively on the vacancy rate of shops in the shopping street, and is therefore endogenous. To take this feature into account, we allow footfall in the shopping street to fall with the vacancy rate of shops. More specifically, we assume that footfall is proportional to the occupancy rate of shops. Hence, $f = (1-v)\bar{f}$, where $\bar{f}$ is the footfall generated when all shops are occupied by retail firms. This assumption implies that there is a negative external effect of vacant shops, because a vacant shop reduces footfall. To investigate the effects of

f and $\bar{f}$ on prices and vacancies, we make the simplifying assumption that $c = e^2/2$ so that $c''(e) = 1$. We then formulate the following proposition:

> PROPOSITION 2: When footfall is proportional to the occupancy rate of shops, *(i)* $\mathrm{d}p/\mathrm{d}\bar{f} > \mathrm{d}p/\mathrm{d}f$ and *(ii)* $\mathrm{d}v/\mathrm{d}\bar{f} < \mathrm{d}v/\mathrm{d}f$.

Proof. Using implicit differentiation and Cramer's rule, we establish that:

(A19) $$\frac{\mathrm{d}v}{\mathrm{d}\bar{f}} = -\frac{\delta v(1-v)^3(1-\beta)}{\Delta e^2 c'' - (1-v)^2(1-\beta)\delta v\bar{f}}.$$

We also obtain the second derivative with respect to $\bar{f}$:

(A20) $$\frac{\mathrm{d}v^2}{\mathrm{d}^2\bar{f}} = -\frac{(1-v)^5 v^2(1-\beta)^2\delta^2}{\left(e^2\Delta c'' - (1-v)^2(1-\beta)\delta v\bar{f}\right)^2} < 0.$$

Using implicit differentiation, it should hold that:

(A21) $$\frac{\mathrm{d}p}{\mathrm{d}f} = \frac{\mathrm{d}p}{\mathrm{d}\bar{f}}\Big/\frac{\mathrm{d}f}{\mathrm{d}\bar{f}} \qquad \text{and} \qquad \frac{\mathrm{d}v}{\mathrm{d}f} = \frac{\mathrm{d}v}{\mathrm{d}\bar{f}}\Big/\frac{\mathrm{d}f}{\mathrm{d}\bar{f}}.$$

So if $\mathrm{d}f/\mathrm{d}\bar{f} > 1$, it holds that $\mathrm{d}p/\mathrm{d}\bar{f} > \mathrm{d}p/\mathrm{d}f$ and $\mathrm{d}v/\mathrm{d}\bar{f} < \mathrm{d}v/\mathrm{d}f$. Hence:

(A22) $$\frac{\mathrm{d}f}{\mathrm{d}\bar{f}} = -\frac{\mathrm{d}v}{\mathrm{d}\bar{f}}\bar{f} + (1-v) > 1,$$

implying that $-\mathrm{d}v/\mathrm{d}\bar{f} > v/\bar{f}$. Because $\mathrm{d}v^2/\mathrm{d}^2\bar{f} < 0$, this condition holds. ■

The main consequence for our empirical investigation is that one underestimates the effect of footfall on rent and overestimates the effect on the vacancy rate when endogeneity of footfall is ignored, so one underestimates the effect of footfall on rental income and one obtains conservative estimates. As we will show later, the underestimate is small if vacancy rates are low, which is the case in our data.

Appendix B. Data appendix

Our main analysis only requires information on footfall, which is obtained from the *Locatus* dataset using a one-to-one shop matching based on location. However, for some sensitivity analyses we need information on the type of retail firm that is occupying a shop. We therefore use a matching process to obtain the retail branch code for each shop and whether the shop is a part of a chain or not. It should be mentioned that our results are not sensitive to this matching process.

The sectoral classification of the shops in the *Strabo* dataset is not very detailed. By contrast, the *Locatus* data provide information on several shop attributes (shop name, address and whether the shop is part of a chain) and includes a detailed sectoral classification up to the branch level (e.g. it distinguishes between male and female clothing, shoes etc.). The matching process between the *Strabo* and *Locatus* datasets is based on shops that are in the *Locatus* dataset the same year or up to two years after the rental transaction (the two-year period seems reasonable since a retail firm usually keeps the shop vacant to refurbish the establishment after renting a shop).

The most accurate way for matching is to use the exact shop name and building identifiers. In this way, we matched 5.64 percent of the Strabo data. Although the shop coordinates in the two datasets are very accurate, in some cases shops might be matched to another building close by.[40] Using the exact names, street number and the 6-digit postal codes (PC6), we matched 19.37 percent additional shops. In the Netherlands, the combination of each PC6 and each street number is unique. However, several observations have missing street numbers. On account of this fact, in a further step we use only the exact names and the PC6 codes and we match a further 1.93 percent of our sample. Hence, in total we matched 26.94 percent of the Strabo data using the complete name of each shop combined with some other exact location criteria.

Frequently, the name of the same shop does not appear identical in the two datasets. For this reason, we use the two first characters of the names in the two datasets for the rest of the shops (excluding articles such as "the" and other common words). Hence, using building identifiers, then PC6 and building numbers and finally, PC6, together with the first two letters of the shop names in the two datasets, we merged a further 5.64, 25.34 and 7 percent, respectively, adding up to a cumulative 64.92 percent of the shops in our sample.

The rest of the shops were matched based on the whole name and the 4-digit postal code (PC4), which approximately corresponds to a building block. This way we matched a further 3.19 percent of our sample and the remaining 31.88 percent was matched based on the exact names alone. We emphasise that we have double-checked manually all the matched observations and the results are very accurate.

[40] Depending on the area occupied by each building, which tends to be small in the central cities of the Netherlands.

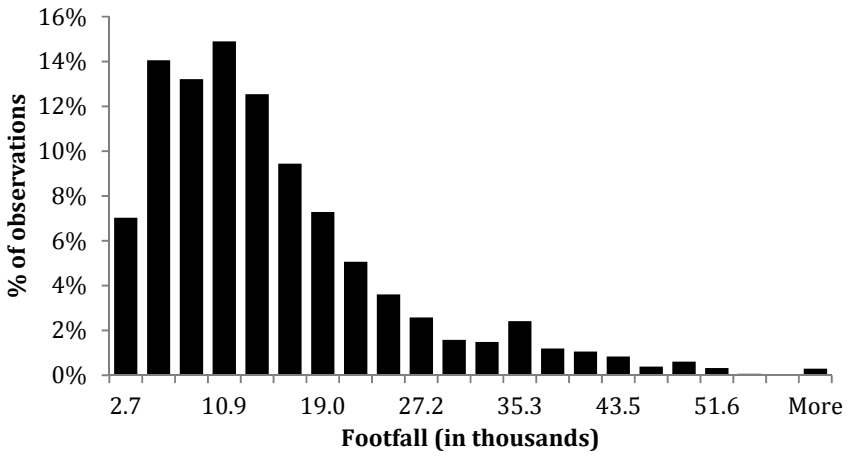

FIGURE B1— FOOTFALL HISTOGRAM

FIGURE B2— NUMBER OF SHOPS HISTOGRAM

Finally, Figure B1 and Figure B2 are the histograms of footfall and the number of shops per street, respectively. Both figures show that there are relative few observations with the lowest values of footfall and number of shops while most of our observations (approximately 52 per cent of the rent transactions) are located in relatively big streets where the number of shops is between 20 and 68 shops.

Appendix C. Extensions and sensitivity

C.1 Effects on house prices

The dataset providing information on residential housing transactions is obtained from *NVM*, the Dutch Association of Real Estate Agents. The dataset provides information on about 90 percent of transactions between 2003 and 2014. We have information on the transaction price, exact location, and a wide range of house attributes such as size (in m^2), type of house, number of rooms and construction year.[41] We merge the house price data to footfall data so that each transaction is within 25m of a shop in the *Locatus* data.[42] One might expect that shopping districts like the ones we analyse have a purely commercial use. However, we have recovered information on building use for the same area that we analyse the effect of footfall on rents, which shows that about 50 percent of the building use is residential.[43]

Table C1 reports the descriptive statistics. The average house price is about € 200 thousand and the average price per m^2 is € 2,333. As one may expect, residential properties are located in less busy shopping streets, with an average footfall of 9,256, i.e. about 30 percent less than in the *Strabo* dataset. The sample mainly includes apartments, as one expects for residential properties in shopping districts which are mainly located in the city centre. Similar to the *Strabo* dataset, about 25 percent of the properties are constructed before 1945. We now focus on the external effect of footfall through its effect on house prices. Retail concentration may have positive effects on residents through the positive effect of footfall on liveability. On the other hand, increased pedestrian traffic may generate congestion and noise, which could impose a negative external effect on residents, so the net effect is ambiguous. We employ the same identification strategy used for retail rents to estimate the external effect of footfall on residents that live in the *same* shopping streets as the ones used in our analysis of shop rents using residential house prices. Table A2 reports the results.

In column (1), we include the logarithm of footfall and we control for housing, building, location characteristics and year fixed effects. The coefficient of footfall suggests that doubling footfall leads to an increase in house prices of 2.2 percent. However, this coefficient is only marginally statistically significant. The positive effect may, however, be explained by the fact that areas with more footfall are generally located in or near the city centre. Such areas are often considered more attractive and therefore command higher housing prices. In column (2), we therefore include shopping district fixed effects, implying

[41] Transactions with prices that are above € 1.5 million or below € 25,000 or a m^2 price below € 250 or above € 5,000 are removed. Furthermore, we drop transactions that are in properties smaller than $25m^2$ or larger than $250m^2$. These selections do not influence the results.
[42] If we decrease this distance, the results are not influenced.
[43] Retail is about 25 percent and the remaining 25 percent is dedicated to other uses. This information is included in the *Administration of Buildings and Addresses* dataset.

TABLE C1 — DESCRIPTIVE STATISTICS OF NVM DATASET

	mean	sd	min	max
House price *(€)*	201,156	95,503	40,000	950,000
Footfall	9,256	7,750	100	66,100
Size of property *(in m^2)*	91.32	34.89	26	250
Number of rooms	3.137	1.155	0	13
House type – apartment	0.901	0.299	0	1
House type – terraced	0.0672	0.250	0	1
House type – semi-detached	0.0256	0.158	0	1
House type – detached	0.00664	0.0812	0	1
Garage	0.102	0.303	0	1
Maintenance state – good	0.895	0.307	0	1
Central heating	0.863	0.344	0	1
Listed building	0.0299	0.170	0	1
Construction year < 1945	0.259	0.438	0	1
Construction year 1945-1959	0.0729	0.260	0	1
Construction year 1960-1969	0.0553	0.229	0	1
Construction year 1970-1979	0.0922	0.289	0	1
Construction year 1980-1989	0.166	0.372	0	1
Construction year 1990-1999	0.164	0.371	0	1
Construction year ≥2000	0.190	0.392	0	1
Mall	0.0546	0.227	0	1
Corner shop	0.00664	0.0812	0	1
Sunny side of street	0.496	0.500	0	1
Pedestrian street	0.629	0.483	0	1
Shopping street length *(in m)*	385.3	270.2	34.68	1,269
Shopping street width *(in m)*	10.60	8.760	3	49.93
Distance to nearest intersection *(in m)*	83.02	116.3	4.241	2,961
Water within 50m	0.0687	0.253	0	1
Water 50-100m	0.0890	0.285	0	1
In historic district	0.341	0.474	0	1
Distance to station *(in m)*	1,692	2,351	35.21	18,602

Note: The number of observations is 9,947.

that we identify the effect of footfall within shopping districts. The coefficient of footfall is then very close to zero and highly insignificant. The low magnitude of the standard errors provides convincing evidence for the absence of an external effect of footfall on residents. In other words, footfall is not a determinant of house prices or alternatively, the positive and negative effects of footfall perfectly counteract each other. Specifications (3)-(6) confirm this finding. When we include street fixed effects in column (3), or 100m intersection fixed effects in column (4), the log footfall coefficient is essentially zero. The same holds in column (5), where we use 50m intersection fixed effects, and in column (6), where we additionally include street and shop characteristics in column (6).

TABLE A2 — REGRESSION RESULTS FOR THE HOUSING MARKET
(Dependent variable: the log of house price)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Footfall *(log)*	0.0237***	0.0141***	0.00326	-0.00160	-0.00120	-0.000895
	(0.00902)	(0.00478)	(0.00526)	(0.00576)	(0.00751)	(0.00748)
Property is in mall						0.0296
						(0.0223)
Property is on sunny side of street						0.0143
						(0.0103)
Property on the corner						0.0268
						(0.0344)
Shopping street width in m *(log)*						0.00494
						(0.00944)
Pedestrian street	-0.000438	-0.0101		-0.0213*	-0.0104	-0.0123
	(0.0256)	(0.0105)		(0.0116)	(0.0150)	(0.0153)
Water within 50m	0.140***	0.0618***	0.0344*	0.0405	0.0903*	0.0926*
	(0.0300)	(0.0147)	(0.0190)	(0.0286)	(0.0516)	(0.0511)
Water 50-100m	0.106***	0.0336***	0.00882	0.0170	0.0114	0.0133
	(0.0329)	(0.0111)	(0.0172)	(0.0199)	(0.0282)	(0.0278)
In historic district	0.0531*	-0.0376**	-0.00736	0.0118	0.0723	0.0685
	(0.0291)	(0.0188)	(0.0483)	(0.0300)	(0.0451)	(0.0448)
Distance to station *(log)*	0.0476***	0.00930	0.00114	0.0719	0.154	0.147
	(0.0112)	(0.0157)	(0.0241)	(0.0541)	(0.111)	(0.110)
Housing characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—	—
Shopping street fixed effects	No	No	Yes	No	No	No
Intersection fixed effects	No	No	No	Yes	Yes	Yes
Observations	9,947	9,947	9,947	7,935	4,847	4,847
R^2	0.615	0.824	0.868	0.879	0.896	0.896

Notes: Footfall is measured as the number of shoppers per day. Housing characteristics include the size of property *(in m^2)*, the number of rooms, the house type (apartment, terraced, semi-detached, detached), the maintenance state (if good), the existence of a garage and central heating, a dummy variable whether a building is listed and construction year dummies. In column (4), we include observations within 100m of a shopping street interaction. In columns (5) and (6), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

C.2 Nonlinearity

We discuss here the results given quadratic specifications in order to allow for a non-linear effect of the logarithm of footfall on the logarithm of rents. We demean the logarithm of footfall by subtracting the logarithm of average footfall over the whole sample from each observation. In Table C3, we show the results of all our main rent specifications (shown in Table 3) using the demeaned log footfall and its square, instead of the annual log footfall. In column (1), we estimated a parsimonious specification where we control for shop and location characteristics and year fixed effects. We observe that the square term is statistically significant and positive, while the coefficient of log footfall is also highly statistically significant and considerably higher than in our main results (0.322). In columns

TABLE C3 — POLYNOMIAL REGRESSIONS: FOOTFALL AND RETAIL RENTS
(Dependent variable: the log of rent)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Footfall (log) – average Footfall (log)	0.438***	0.396***	0.376***	0.302***	0.285***	0.283***
	(0.0300)	(0.0230)	(0.0332)	(0.0302)	(0.0362)	(0.0362)
(Footfall (log) – average Footfall (log))^2	0.0990***	0.0759***	0.0726***	0.0615***	0.0692***	0.0690***
	(0.0139)	(0.0108)	(0.0155)	(0.0124)	(0.0142)	(0.0143)
Shop characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Construction year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—	—
Shopping street fixed effects	No	No	Yes	No	No	No
Intersection fixed effects	No	No	No	Yes	Yes	Yes
Observations	3,102	3,102	3,102	2,629	1,870	1,870
R^2	0.606	0.723	0.814	0.851	0.875	0.875

Notes: Footfall is measured as the number of shoppers per day. Shop characteristics are dummy variables for new and renovated buildings, as well as for sublet properties, building surface area in m^2 (log) and construction year dummies. In column (4), we include observations within 100m of a shopping street interaction. In columns (5) and (6), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

(2), (3) and (4), we add shopping district, shopping street and 100m shopping intersection fixed effects, respectively. From columns (1)-(4), the coefficient of log footfall drops from 0.438 to 0.302. Adding 50m shopping street intersection fixed effects in column (5), as well as shop and street characteristics in column (6), reduces the coefficient of log footfall to 0.285 and 0.283, respectively. These results indicate that the elasticity of rents with respect to footfall is increasing in footfall.

Table C4 presents a quadratic specification for the vacancy analysis using a logistic regression. The main independent variables are the demeaned log footfall and its square and the reported coefficients are average marginal effects. Again, we start from a naïve specification in column (1), which includes shop and location characteristics and year fixed effects. In columns (2)-(6), Table C4, we gradually add shopping district, shopping street, 100m, 50m shopping intersection fixed effects and shop and street characteristics, respectively. In all specifications of Table C4, the squared term is statistically significant. The coefficient of the demeaned footfall in column (6), Table C4, is –0.0367, which is substantially higher the coefficient of log footfall in Table 5 (–0.0272).[44]

We discuss here also the results given quadratic specifications of the logarithm of the

[44] It should be mentioned that we use a non-linear specification to test whether the estimated effect is robust to such a specification. However, the focus of the paper is on the average effect. Thus, we use the linear log specification in the main results of the paper.

TABLE C4 — POLYNOMIAL REGRESSIONS: FOOTFALL AND VACANT SHOPS
(Dependent variable: dummy shop is vacant)

	(1) Logit	(2) Logit	(3) Logit	(4) Logit	(5) Logit	(6) Logit
Footfall (log) – average Footfall (log)	-0.0432***	-0.0435***	-0.0434***	-0.0398***	-0.0370***	-0.0367***
	(0.00237)	(0.00176)	(0.00189)	(0.00201)	(0.00250)	(0.00252)
(Footfall (log) – average Footfall (log))^2	-0.00767***	-0.00723***	-0.00610***	-0.00587***	-0.00445***	-0.00442***
	(0.000907)	(0.000776)	(0.000874)	(0.00100)	(0.00122)	(0.00123)
Construction year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—	—
Shopping street fixed effects	No	No	Yes	No	No	No
Intersection fixed effects	No	No	No	Yes	Yes	Yes
Log-likelihood	-93,914	-91,926	-89,420	-69,358	-43,983	-43,967
Observations	425,834	425,834	421,204	338,099	220,049	220,049

Notes: Reported coefficients are average marginal effects. Footfall is measured as the number of shoppers per day. In column (4), we include observations within 100m of a shopping street interaction. In columns (5) and (6), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

TABLE C5 — POLYNOMIAL REGRESSIONS: NUMBER OF SHOPS AND RETAIL RENTS
(Dependent variable: the log of rent)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS
Number of shops *(log)* –	0.0585	0.124***	0.190***	0.155***	0.158***
Average number of shops *(log)*	(0.0407)	(0.0280)	(0.0357)	(0.0395)	(0.0395)
(Number of shop *(log)* –	-0.00537	0.0351**	0.0540***	0.0384*	0.0393**
Average number of shops *(log)*)^2	(0.0195)	(0.0150)	(0.0172)	(0.0201)	(0.0200)
Shop characteristics	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes
Construction year dummies	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—
Shopping street fixed effects	No	No	No	No	No
Intersection fixed effects	No	No	Yes	Yes	Yes
Observations	3,102	3,102	2,629	1,870	1,870
R^2	0.466	0.638	0.838	0.864	0.864

Notes: Shop characteristics are dummy variables for new and renovated buildings, as well as for sublet properties, building surface area in m^2 (log) and construction year dummies. In column (3), we include observations within 100m of a shopping street interaction. In columns (4) and (5), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

number of shops. We demean the logarithm of number of shops by subtracting the logarithm of average number of shops of the whole sample from each observation. In Table C5 and Table C6, we show the results of all our main specifications. We observe that the

TABLE C6 — POLYNOMIAL REGRESSIONS: NUMBER OF SHOPS AND RETAIL RENTS
(Dependent variable: dummy shop is vacant)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS
Number of shops *(log)* –	-0.0116**	-0.0245***	-0.0123	-0.00469	-0.00664
Average number of shops *(log)*	(0.00526)	(0.00677)	(0.0119)	(0.0128)	(0.0127)
(Number of shop *(log)* –	0.00240	-0.00286	0.00484	0.00435	0.00391
Average number of shops *(log)*)^2	(0.00437)	(0.00440)	(0.00629)	(0.00757)	(0.00761)
Shop characteristics	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes
Construction year dummies	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—
Shopping street fixed effects	No	No	No	No	No
Intersection fixed effects	No	No	Yes	Yes	Yes
Observations	3,080	3,080	2,610	1,861	1,861
R^2	0.032	0.138	0.517	0.609	0.613

Notes: Reported coefficients are average marginal effects. In column (3), we include observations within 100m of a shopping street interaction. In columns (4) and (5), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

square term is statistically significant and positive for rent, but statistically insignificant for vacancies. Hence, the elasticity of rental income is increasing in the number of shops.

One important street characteristic that we control for in all our specifications is a dummy for whether a street is pedestrianised. In our dataset, about 80 percent of shops are located in pedestrianised streets. Not surprisingly, mean footfall is about twice as high in pedestrianised streets. It is important to point out that the dummy variable for pedestrianised streets will be endogenous when the model is misspecified (because of the strong positive correlation between pedestrianised streets and footfall). In Table C7, we have estimated a model interacting the pedestrianised and non-pedestrianised streets with the demeaned log footfall and its square, the same transformed variables we used in Table C3. Column (1), Table C7, which is a naïve specification that includes shop and location characteristics, together with year fixed effects, shows that the coefficients related to pedestrian streets and non-pedestrian streets seem to differ. However, this difference becomes statistically insignificant when we include shopping district, street and 100m intersection fixed effects in columns (2), (3) and (4), respectively. Finally, in column (5), where we include 50m intersection fixed effects, and in column (6), where we also add shop and street characteristics, we find that both the linear and the quadratic terms for pedestrian and non-pedestrianised streets are identical and that the pedestrian street dummy is equal to zero (this holds for the individual constraints, as well as for a *F*-test, which jointly tests the three constraints). These results are in line with our assumption that the logarithm of footfall fully captures shopping externalities (otherwise the dummy for pedestrianised streets would be greater than zero).

TABLE C7 —PEDESTRIANISED STREETS VERSUS NON-PEDESTRIANISED STREETS: POLYNOMIAL REGRESSIONS

(Dependent variable: the log of rent)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Pedestrian×footfall *(log demeaned)*	0.457***	0.404***	0.380***	0.292***	0.283***	0.281***
	(0.0327)	(0.0247)	(0.0356)	(0.0309)	(0.0377)	(0.0378)
Pedestrian×footfall *(log demeaned)*^{2}	0.0986***	0.0655***	0.0610***	0.0429**	0.0562**	0.0557**
	(0.0176)	(0.0133)	(0.0220)	(0.0197)	(0.0264)	(0.0267)
Non-pedestrian×footfall *(log demeaned)*	0.296***	0.326***	0.326***	0.343***	0.278***	0.282***
	(0.0485)	(0.0456)	(0.0704)	(0.0897)	(0.0906)	(0.0911)
Non-pedestrian×footfall *(log demeaned)*^{2}	0.0555***	0.0663***	0.0704***	0.0827***	0.0741***	0.0755***
	(0.0191)	(0.0170)	(0.0203)	(0.0261)	(0.0240)	(0.0240)
Shop characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Shopping street characteristics	No	No	No	No	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—	—
Shopping street fixed effects	No	No	Yes	No	No	No
Intersection fixed effects	No	No	No	Yes	Yes	Yes
Log-likelihood	-93,898	-91,906	-89,402	-69,338	-43,957	-43,942
Observations	425,834	425,834	421,204	338,099	220,049	220,049

Notes: Footfall *(log demeaned)* is calculated by subtracting the log of the annual mean of footfall from footfall (log). Footfall is measured as the number of shoppers per day. In column (4), we include observations within 100m of a shopping street interaction. In columns (5) and (6), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

C.3 Booms and busts

Here, we consider an alternative explanation for the negative effect of footfall on vacancies other than the effect of increased opportunity costs of vacant properties in expensive areas (i.e. with higher levels of footfall). In times of high demand, the marginal costs of providing shop space are likely to be above the marginal benefits for most of the shops, so footfall might not have a statistically significant effect on vacancy rates during a boom period. However, in bust times because marginal costs of providing space may be above the marginal benefits, retail space may lie empty in areas with lower rents (i.e. with lower footfall). Hence, footfall may only have an effect during busts. We test this hypothesis by regressing a dummy for a vacant shop on the interaction term between log footfall and a dummy for the recent boom (2003-2008) and bust (2009-2015) period of the Dutch economy, respectively.[45] Table C8 reports the results.

Column (1), Table C8, includes only building, location characteristics and year fixed effects, as control variables. The coefficient of log footfall in the boom and the bust period is –0.0259 and –0.0288, respectively, and they are both highly statistically significant. While

[45] The actual years of recession were 2009, 2012, 2013 and 2015. We have also performed the same exercise using the exact years that the economy was in recession. The results are virtually the same.

TABLE C8 — REGRESSIONS RESULTS FOR VACANT SHOPS: BOOMS AND BUSTS
(Dependent variable: dummy shop is vacant)

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	OLS	OLS
Boom×footfall *(log)*	-0.0259***	-0.0259***	-0.0287***	-0.0254***	-0.0243***	-0.0241***
	(0.00139)	(0.00127)	(0.00159)	(0.00149)	(0.00183)	(0.00185)
Bust×footfall *(log)*	-0.0288***	-0.0303***	-0.0323***	-0.0289***	-0.0296***	-0.0294***
	(0.00139)	(0.00127)	(0.00164)	(0.00153)	(0.00180)	(0.00182)
Shop characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Shopping street characteristics	No	No	No	No	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—	—
Shopping street fixed effects	No	No	Yes	No	No	No
Intersection fixed effects	No	No	No	Yes	Yes	Yes
Log-likelihood	-94,298	-92,216	-89,560	-69,459	-44,010	-43,994
Observations	425,834	425,834	421,204	338,099	220,049	220,049

Notes: Reported coefficients are average marginal effects. Footfall is measured as the number of shoppers per day. The boom period is 2003-2008 and the bust period is 2009-2015. Footfall is measured as the number of shoppers per day. Shop, location and shopping street characteristics are mentioned in Table C13. In column (4), we include observations within 100m of a shopping street interaction. In columns (5) and (6), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

these two coefficients are not statistically different, when we include shopping district fixed effects in column (2), this difference becomes significant. Including street fixed effects in column (3) or fixed effects for observations within 100m from their closest intersection in column (4), makes the difference between the log footfall coefficient for the boom and the bust period non-statistically significant. Finally, in column (5), where we restrict the sample to observations within 50m from an intersection, and in column (6), where we also add shop and street characteristics, the difference between the effect of shopping externalities on vacancies between the boom and the bust period becomes statistically significant again.

The difference between the two coefficients corroborates our intuition that in times of low demand, an increase in footfall raises marginal benefits above marginal costs for certain shops. Nevertheless, in times of high demand, the effect of footfall on vacancies is in line with the opportunity cost hypothesis: store owners' opportunity cost of not filling a vacant shop increases with footfall. Overall, these results suggest that the effect of footfall on vacancy rates in times of high demand is slightly lower but still highly statistically significant.

C.4 Retail chains and shop ownership

Another potential concern is that the location decisions of independent retailers and shops that are part of a retail chain could be fundamentally different. Chains may be interested in maximizing their profits 'globally' while independent retailers only focus on the local

TABLE C9 — REGRESSIONS RESULTS FOR RETAIL RENTS: CHAINS AND NON-CHAINS

(Dependent variable: the log of rent)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
			Panel A: Chains			
Footfall *(log)*	0.320***	0.266***	0.242***	0.141***	0.156***	0.162***
	(0.0318)	(0.0311)	(0.0539)	(0.0529)	(0.0603)	(0.0623)
Observations	1,118	1,118	1,118	984	715	715
R^2	0.602	0.729	0.848	0.913	0.934	0.934
			Panel B: Non-chains			
Footfall *(log)*	0.277***	0.266***	0.274***	0.193***	0.176***	0.171***
	(0.0218)	(0.0173)	(0.0339)	(0.0328)	(0.0456)	(0.0461)
Observations	1,984	1,984	1,984	1,645	1,155	1,155
R^2	0.511	0.695	0.806	0.839	0.870	0.871
Shop characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Shopping street characteristics	No	No	No	No	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—	—
Shopping street fixed effects	No	No	Yes	No	No	No
Intersection fixed effects	No	No	No	Yes	Yes	Yes

Notes: Footfall is measured as the number of shoppers per day. Shop, location and shopping street characteristics are mentioned in Table . In column (4), we include observations within 100m of a shopping street interaction. In columns (5) and (6), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

TABLE C10 — REGRESSION RESULTS FOR RETAIL RENTS: CONTROLLING FOR SHARE CHAIN AND CLOTHING STORES

(Dependent variable: the log of rent)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Footfall *(log)*	0.261***	0.258***	0.306***	0.200***	0.199***	0.198***
	(0.0222)	(0.0198)	(0.0303)	(0.0269)	(0.0373)	(0.0378)
Share chain stores in shopping street	0.395***	0.476***	0.306	0.436***	0.486***	0.485***
	(0.0938)	(0.0616)	(0.205)	(0.0837)	(0.0986)	(0.0991)
Share clothing stores in shopping street	0.327***	0.110*	-0.116	-0.116	-0.132	-0.127
	(0.0804)	(0.0591)	(0.117)	(0.0863)	(0.116)	(0.118)
Shop characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Shopping street characteristics	No	No	No	No	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—	—
Shopping street fixed effects	No	No	Yes	No	No	No
Intersection fixed effects	No	No	No	Yes	Yes	Yes
Observations	3,102	3,102	3,102	2,629	1,870	1,870
R^2	0.582	0.711	0.809	0.848	0.871	0.872

Notes: Footfall is measured as the number of shoppers per day. In column (4), we include observations within 100m of a shopping street interaction. In columns (5) and (6), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

market. If this hypothesis holds, then chains would engage in coordinated strategic location choices in order to maximize their total catchment area and they would avoid unnecessary local competition between their shops. An alternative strategy for a retail chain could be to establish various shops in close proximity to deter the entrance of possible competitors in the market.[46] Furthermore, advertising may also influence the location choices of chain shops. On the one hand, exposure to high footfall may be good advertisement that may yield popularity for the whole retail chain. On the other hand, due to the advertising campaigns of the big retail chains, the probability that a pedestrian passing by enters a shop and purchases something could be higher for chains compared to independent retail firms. Table C9 sheds light on this issue by reporting the results when splitting the sample into chain shops and non-chain shops.

Again, we follow the specifications used in Table 3. Column (1) in Table C9, shows the results of the naïve specification where we control for shop and location characteristics, as well as for year fixed effects. In columns (2), (3) and (4), we use shopping district, street and 100m intersection fixed effects, respectively. While the coefficients for chain shops and non-chain shops appear to be different in columns (1) and (4), they are not statistically different. Moreover, when we include 50m intersection fixed effects in column (5) and shop and street characteristics in column (6), the estimated coefficients of log footfall for chain shops and non-chain shops are very similar. These results suggest that both chains and non-chains value shopping externalities similarly.

We also aim to test whether the effect of footfall could be partly attributed to the presence of chain stores. That is, one could argue that other shops may be mostly interested in footfall generated by chain stores. The same hold for clothing stores. We test this in Table C10 where we directly control for the share of chain stores, as well as the closing stores in the same shopping street. It is immediately observed that the coefficient of footfall is essentially unaffected.

As we mentioned in the introduction, policy intervention fostering the concentration of footfall-generating retail activities can be welfare improving *only if* the external effect of footfall is not internalised. In the introduction we argued that internalisation is unlikely to occur in the Netherlands due to the fragmentation of shop ownership. As an empirical test for this argument, we use the information of shop owner name and shop owner type, which is available in the *Strabo* property dataset in order to test whether different ownership statuses yield different estimates of the effect of footfall on retail rents. As mentioned in Section III.B, information on shop owner name is available for about one third of the sample that we use in the rent analysis while information on shop owner type is available for about two thirds of the same sample.

[46] In reality, some big chains tend to locate many of their shops in close proximity, even within the same shopping street.

TABLE C11 — REGRESSION RESULTS FOR RETAIL RENTS: OWNERSHIP STATUS
(Dependent variable: the log of rent)

	(1) OLS	(2) OLS
Panel A: Multi vs. single-property ownership		
Single-property owners×footfall *(log)*	0.139	0.217***
	(0.0977)	(0.0741)
Multi-property owners×footfall *(log)*	0.136	0.225***
	(0.102)	(0.0767)
Observations	558	760
R^2	0.942	0.922
Panel B: Private vs. corporate ownership		
Private property owners×footfall *(log)*	0.229***	
	(0.0404)	
Real estate companies×footfall *(log)*	0.229***	
	(0.0398)	
Observations	1,458	
R^2	0.889	
Shop characteristics	Yes	Yes
Location characteristics	Yes	Yes
Year fixed effects	Yes	Yes
Shopping street characteristics	Yes	Yes
Intersection fixed effects	Yes	Yes

Notes: Multi-property (single) ownership is a dummy variable which takes the value one if a shop belongs to a property owner who owns multiple (no other) shops in the shopping street. Private property owners are those listed as private investors. Footfall is measured as the number of shoppers per day. Shop, location and shopping street characteristics are mentioned in Table . In column (1) we include observations within 50m of a shopping street interaction while in column (2) we increase this distance to 100m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

In Panel A, Table C11, we report the results using two interaction variables for the properties that belong to shop owners owning a single property in the same shopping street, multiplied by the logarithm of footfall. The second interaction term uses properties that belong to shop owners owning multiple properties in the same shopping street. In column (1), we use our preferred specification which includes the full set of controls and 50m intersection fixed effects. Given the limited information on shop owner names in our data and consequently, the low number of observations, we cannot consistently estimate the effect of log footfall on retail rents using this specification, which only includes observations within 50m of an intersection. For this reason, column (2), includes observations within 100m from an intersection with the full set of control variables. In both columns (1) and (2), the coefficients are virtually the same for single and multi-property ownership. The results of Panel A, Table C11, suggest that multi-property store owners value footfall in the same way as single property owners and thus, it might be expected that both behave in a similar manner.

Panel B in Table C11 uses again two interaction terms of the properties that belong to private-property owners (versus real estate agencies, pension funds, construction companies etc.) who are (versus not) listed as private investors, multiplied by the logarithm of footfall. The coefficients of log footfall for private and commercial property owners are exactly the same. Overall, the results in Table C11 seem to confirm that any coordination among property owners in order to attract high-footfall generating activities and fully internalise the shopping externality is very unlikely to happen in the setting of the Dutch shopping streets.

C.5 General robustness checks

In this subsection, we present additional results confirming the results obtained in in Section IV. We start by analysing the effects of footfall on rents. Our baseline specification is reported in column (6) of Table 3. In that specification, we regress log rents on log footfall controlling for shop and location characteristics, construction year dummies, year fixed effects, shopping street and other shop characteristics, as well as 50m intersection fixed effects based on the distance of each shop to its closest street intersection. The coefficient of log footfall we estimated is 0.213 and highly statistically significant.

In column (1) of Table C12 we use instead of 50m intersection fixed effects, street-specific 50m intersection fixed effects (interacting the shopping street dummy with the 50m intersection dummy). Consequently, we control for time invariant local endowments that are the same in all shops located on the same street where the distance between these shops is less than 100m. The magnitude of the estimated elasticity is lower (0.136) since these fixed effects absorb a considerable part of the identifying variation (between intersecting streets).[47] However, the effect of footfall is still highly statistically significant.

This result suggests that our findings are robust even when we control for unobservable endowments that are the same between shops located in close proximity and *in the same* shopping street. Another way to control for street-specific local endowments is to use fixed effects for the 6-digit postal code (PC6) of each shop. PC6 refers to roughly one side of a building block, approximating the street-specific 50m intersection interaction fixed effect. The main difference between the two fixed effects is that PC6 is an administrative area. Column (2) in Table C12 then includes PC6 fixed effects. The estimated log footfall effect is 0.127 and highly statistically significant.

We also estimate our preferred specification using one or two-year lags of footfall to control for reverse causality. Column (3) reports the log footfall coefficient in the previous year than the rent transaction took place. The estimated coefficient is still highly statistically

[47] The coefficient of log footfall that we obtain using our baseline specification (Column (6)) in Table 3 with the same sample as in column (1), Table , is 0.193 and highly statistically significant.

TABLE C12 — ROBUSTNESS CHECKS FOR RETAIL RENTS

(Dependent variable: the log of rent)

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	OLS	OLS
Footfall *(log)*	0.136***	0.127***			0.232***	0.223***
	(0.0474)	(0.0379)			(0.0346)	(0.0674)
Footfall *(log)*, one-year lag			0.233***			
			(0.0339)			
Footfall *(log)*, average over time				0.272***		
				(0.0398)		
Observations	1,471	1,373	1,839	1,870	1,793	425
R^2	0.888	0.887	0.873	0.874	0.873	0.904
Shop characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Shopping street characteristics	Yes	Yes	Yes	No	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
PC6 fixed effects	No	Yes	No	No	No	No
Shopping street fixed effects	Yes	No	No	No	No	No
Intersection fixed effects	Yes	No	Yes	Yes	Yes	Yes
Shopping street×intersection fixed effects	Yes	No	No	No	No	No

Notes: The coefficients of log footfall using our preferred specification (Column (6) in Table 3) but using the same sample as in columns (1) and (2), Table , is 0.193***(0.039) and 0.218***(0. 048), respectively. Footfall is measured as the number of shoppers per day. Shop characteristics are the size of shop in m^2 (log), dummy variables for new and renovated buildings, as well for sublet properties, building surface area in m^2 (log) and construction year dummies. Location characteristics are dummies for pedestrian streets, for proximity to water within 50m, or in the range 50-100m, for historic districts and for the distance to the closest station (log). Shop and street characteristics are dummies for properties in malls, on corners, on the sunny side of street, as well as the shopping street width in m (log). In columns (1)-(6), we include observations within 50m of a shopping street interaction. In column (5), we have excluded the shops inside a mall and in column (6), we have excluded the shops located further than 1 km from the closest train station. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

significant and although we lose some observations from our sample, its value is 0.233, very similar to our preferred estimate. Using a two-year lag of log footfall (not reported in Table C12), our estimated effect is 0.243 and highly statistically significant.

The third concern of measurement error in footfall that we discussed in Section III.A is related to random variation between different Saturdays of each year at the same location. Although our identification strategy is based on spatial differences in footfall and rents, in column (4), Table C12, we use the logarithm of the average footfall over time in order to fully address this concern. The estimated coefficient is 0.272 and highly statistically significant. Again, the coefficient is higher but not statistically different from our main estimate. Another concern we raised is that shops located close to train stations and those located inside a mall may be very different from shops located in ordinary shopping streets. For this reason, in columns (5) and (6), we exclude observations located less than 1km away from a train station and the shops that are inside a mall, respectively. As we can see in Table C12, the estimated coefficients are hardly different from our main estimates.

TABLE C13 — ROBUSTNESS CHECKS FOR VACANT SHOPS
(Dependent variable: dummy shop is vacant)

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	OLS	OLS
Footfall *(log)*	-0.0290***	-0.0494***			-0.0266***	-0.0272***
	(0.00217)	(0.00799)			(0.00158)	(0.00247)
Footfall *(log)*, one-year lag			-0.0254***			
			(0.00163)			
Footfall *(log)*, average over time				-0.0313***		
				(0.00201)		
Shop characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Shopping street characteristics	Yes	Yes	Yes	No	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Shop fixed effects	No	Yes	No	No	No	No
Shopping street fixed effects	Yes	No	No	No	No	No
Intersection fixed effects	Yes	No	Yes	Yes	Yes	Yes
Shopping street×intersection fixed effects	Yes	No	No	No	No	No
Log-likelihood	200,377	66,940	194,808	220,049	206,117	57,595
Observations	-42,010	-27,246	-36,870	-44,042	-41,540	-11,102

Notes: Reported coefficients are average marginal effects. Footfall is measured as the number of shoppers per day. Shop characteristics are building surface area in m^2 (log) and construction year dummies. Location characteristics are dummies for pedestrianised streets, for proximity to water within 50m, or in the range 50-100m, for historic districts and for the distance to the closest station (log). Shop and street characteristics are dummies for properties in malls, on corners, on the sunny side of street, as well as the shopping street width in m (log). In columns (1)-(6), we include observations within 50m of a shopping street interaction. In column (5), we have excluded the shops inside a mall and in column (6), we have excluded the shops further than 1 km from the closest train station. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

We run the same set of robustness checks for vacancies in Table C13. The reported results are estimated using a Logit regression based on our preferred specification used in column (6), Table 5. In column (1) of Table C13, we compare shops located less than 50m away from an intersection on the same street by including street×intersection fixed effects. The reported coefficient is –0.0290, essentially the same as in in our baseline specification. Column (2) in Table C13 is different from column (2) in Table C12, where we use PC6 fixed effects. Given that we can observe each shop on an annual basis, we use retail establishment (shop) fixed effects and we obtain identification only based on temporal variation of vacancies and footfall. The fact that our results are still highly statistically significant and in the same order of magnitude as in our main estimates is reassuring.

Columns (3) and (4) in Table C13 include a one-year lag of log footfall and the annual average of log footfall, respectively, instead of the yearly footfall as the main variable of interest. In line with the robustness checks for the rent analysis, the coefficient of lag footfall *(log)* is slightly lower while the coefficient of the annual average of log footfall is slightly

TABLE C14 — REGRESSION RESULTS FOR VACANT SHOPS: LINEAR PROBABILITY MODEL
(Dependent variable: dummy shop is vacant)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Footfall *(log)*	-0.0309***	-0.0325***	-0.0354***	-0.0315***	-0.0312***	-0.0309***
	(0.00147)	(0.00143)	(0.00200)	(0.00183)	(0.00229)	(0.00232)
Building surface area in m^2 *(log)*	0.00275***	-0.000473	4.04e-05	-0.000569	0.00102	0.00108
	(0.000956)	(0.000777)	(0.000747)	(0.000763)	(0.000927)	(0.000924)
Shop characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—	—
Shopping street fixed effects	No	No	Yes	No	No	No
Intersection fixed effects	No	No	No	Yes	Yes	Yes
Observations	425,834	425,834	425,834	338,099	220,049	220,049
R^2	0.018	0.028	0.043	0.046	0.052	0.052

Notes: Footfall is measured as the number of shoppers per day. In column (4), we include observations within 100m of a shopping street interaction. In columns (5) and (6), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

higher (in absolute terms) than the estimates using the annual log footfall. These results confirm that any bias introduced by reverse causality or by measurement error in annual footfall is not substantial. Finally, in columns (5) and (6), Table C13, we exclude shops that are inside a mall and shops that are close to a train station, respectively. The coefficient of log footfall is not significantly different from our main results, confirming that such shops do not drive our main results.

Table C14 replicates the results reported in Table 5 using a Linear Probability Model (LPM) instead of a logistic regression. The reported coefficients can be directly interpreted as semi-elasticities. These results are very similar to the average marginal effects that we discussed in Section 0.B. Column (1) is a naïve specification, which only includes shop and location characteristics, construction year dummies and year fixed effects. The coefficient in column (1) suggests that doubling footfall leads to a 2.1 percentage point reduction in vacancies. When we include shopping district fixed effects in column (2), street fixed effects in column (3) or 100m intersection fixed effects in column (4), the effect is essentially the same. Column (5), which includes 50m intersection fixed effects and column (6), which additionally includes street and shop characteristics, yields the exact same coefficient as in column (1), confirming our main results.

C.6 Combining footfall and number of shops

Here we will explore the extent to which footfall is superior to the use of the number of shops in the shopping street, as a proxy for shopping externalities. In Table C15 we add the

TABLE C15 — REGRESSION RESULTS FOR RETAIL RENTS: FOOTFALL AND NUMBER OF SHOPS
(Dependent variable: the log of rent)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS
Footfall *(log)* year average	0.374***	0.360***	0.282***	0.255***	0.252***
	(0.0227)	(0.0193)	(0.0312)	(0.0414)	(0.0417)
Number of shops in street (*log*)	-0.0285	-0.00255	0.0343*	0.0353	0.0370*
	(0.0221)	(0.0155)	(0.0195)	(0.0223)	(0.0225)
Observations	3,102	3,102	2,629	1,870	1,870
R^2	0.605	0.727	0.853	0.874	0.874
Shop characteristics	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes
Shopping street characteristics	No	No	No	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—
Intersection fixed effects	No	No	Yes	Yes	Yes

Notes: Footfall is measured as the number of shoppers per day. The number of shops in street (*log*) is the logarithm of the number of shops on the same street and in the same year that the rent transaction took place. In column (3), we include observations within 100m of a shopping street interaction. In columns (4) and (5), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

TABLE C16 — REGRESSION RESULTS FOR VACANT SHOPS: FOOTFALL AND NUMBER OF SHOPS
(Dependent variable: dummy shop is vacant)

	(1) OLS	(2) OLS	(3) OLS	(3) OLS	(5) OLS
Footfall *(log)* year average	-0.0292***	-0.0294***	-0.0277***	-0.0278***	-0.0275***
	(0.00138)	(0.00118)	(0.00140)	(0.00165)	(0.00168)
Number of shops in street (*log*)	0.00927***	0.00399***	2.33e-06	-0.000372	-0.000597
	(0.00176)	(0.00102)	(0.00107)	(0.00123)	(0.00123)
Log-likelihood	-94,041	-92,184	-69,460	-44,013	-43,997
Observations	425,783	425,783	338,070	220,020	220,020
Shop characteristics	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes
Shopping street characteristics	No	No	No	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Shopping district fixed effects	No	Yes	—	—	—
Intersection fixed effects	No	No	Yes	Yes	Yes

Notes: Reported coefficients are average marginal effects. Footfall is measured as the number of shoppers per day. The number of shops in street (*log*) is the logarithm of the number of shops on the same street and in the same year as each shop observation. In column (3), we include observations within 100m of a shopping street interaction. In columns (4) and (5), we reduce this distance to 50m. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

logarithm of the number of shops that are located at the same street as the shop where a rent transaction took place, together with the logarithm of the average annual footfall.[48] We use the logarithm of the average annual footfall instead of log (annual) footfall to mitigate reverse causality and measurement error concerns as we discuss in detail in Section III.[49] Following Nunn and Puga (2012), if the number of shops entirely accounted for the differential effect of log footfall between high and low footfall intersecting streets, the coefficient of log footfall should diminish and the effect of log number of shops should remain unaffected

In column (1) we regress the log rent on the log average annual footfall controlling for shop and location characteristics and time fixed effects. The coefficient of log average annual footfall is virtually unchanged compared to the coefficient we obtain using the same specification without including the log number of shops on the same street (0.365). Using shopping district fixed effects in column (2) does not affect our results. In column (3), where we include 100m intersection fixed effects, the coefficient of the log number of shops becomes marginally statistically significant. However, the marginal effect is relatively low while the log average annual footfall coefficient is very similar to the same specification without including the log number of shops. Finally, in columns (4) and (5), Table C15, we restrict our sample to observations within 50m from an intersection and in column (5), we also add shopping street and other shop characteristics. Again, the results are very similar, suggesting that the number of shops cannot capture the full potential of shops to generate shopping externalities. In other words, these results suggest that the potential of shops to generate footfall is heterogeneous so that the elasticity of footfall with respect to shops, $\varepsilon_{f,N}$ is heterogeneous.

In Table C16 we repeat the same exercise as above, but using a shop's incidence of being vacant, as the dependent variable. Again, we find that footfall is the main determinant of vacancies whereas the effect of shops is largely absent. Hence, these results are in line with the results reported in Table 4 and our notion that footfall is the most appropriate measure to capture the heterogeneity of shopping externalities generated by each shop.

C.7 Number of shops with different distance thresholds

In the previous section, we have estimated the effect of footfall and number of shops. However, footfall exhibits substantial local variation within the same street. Therefore, one could argue that the aggregate number of shops at the *street* level cannot capture the local

[48] We matched each rent transaction (or each shop observation in the vacancy analysis) to all shops on the same street if they were non-vacant during the same or the previous year that the rent transaction took place (or for each shop observation).

[49] Using the log average footfall in our main specification yields relatively higher coefficients compared to when we use the log footfall. The results are included in column (4), in Table C12 in Appendix C.6,

TABLE C17 — REGRESSION RESULTS FOR RETAIL RENTS: NUMBER OF SHOPS IN DISTANCE THRESHOLD
(Dependent variable: the log of rent)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS
	Panel A: Footfall and number of shops (log) (in same shopping street)				
Number *(log)* of shops within:	Baseline	50m	100m	150m	200m
Footfall *(log)* year average	0.272***	0.260***	0.246***	0.245***	0.243***
	(0.0398)	(0.0394)	(0.0406)	(0.0415)	(0.0415)
Number of shops (*log*)		0.0618*	0.0739**	0.0642**	0.0645**
		(0.0343)	(0.0336)	(0.0306)	(0.0277)
Observations	1,870	1,869	1,870	1,870	1,870
R^2	0.874	0.874	0.874	0.874	0.875
	Panel B: Number of shops (log) (in same shopping street)				
Number of shops within:	Whole street	50m	100m	150m	200m
Number of shops (*log*)	0.100***	0.126***	0.167***	0.153***	0.144***
	(0.0234)	(0.0385)	(0.0367)	(0.0319)	(0.0286)
Observations	1,870	1,869	1,870	1,870	1,870
R^2	0.864	0.862	0.864	0.865	0.865
Shop characteristics	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes
Shopping street characteristics	Yes	Yes	Yes	Yes	Yes
Intersection fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: The number of shops includes all shops within each distance threshold that are located on the same street and in the same year that the rent transaction took place. Footfall is measured as the number of shoppers per day. In columns (1)-(5), we include observations within 50m of a shopping street interaction. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

nature of shopping externalities. In Panel A, Table C17, we include both the log (average) footfall and the log number of shops for different distance thresholds.[50] Following our baseline specification using the logarithm of the average annual footfall and the full set of control variables, year and 50m intersection fixed effects (shown in column (1), Panel A, Table C17), we use four different distance thresholds for the shops that are located on the same street that the rent transaction took place. The distance thresholds range from 50m to 200m. Columns (2)-(4) present the results for each distance threshold. Regardless of the distance threshold used, the effect of the number shops is statistically significant. Nonetheless, the marginal effect of footfall is also highly statistically significant and virtually unchanged compared to the baseline specification in column (1). If the log number of shops accounted for the differential effect of log footfall between high and low footfall intersecting

[50] We matched each rent transaction (or each shop observation in the vacancy analysis) to all shops on the same street if they were non-vacant during the same or the previous year that the rent transaction took place (or for each shop observation).

TABLE C18 — REGRESSION RESULTS FOR VACANT SHOPS: NUMBER OF SHOPS IN DISTANCE THRESHOLD
(Dependent variable: dummy shop is vacant)

	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	OLS	OLS	OLS
	Panel A: Footfall and number of shops (log) (in same shopping street)				
Number of shops within:	Street *(baseline)*	50m	100m	150m	200m
Footfall *(log)*	-0.0313***	-0.0309***	-0.0307***	-0.0307***	-0.0309***
	(0.00201)	(0.00203)	(0.00207)	(0.00208)	(0.00208)
Number of shops (*log*)		-0.00441**	-0.00398**	-0.00265*	-0.00169
		(0.00173)	(0.00170)	(0.00155)	(0.00148)
Log-likelihood	-44,042	-43,915	-44,006	-44,031	-44,033
Observations	220,049	219,503	219,974	220,020	220,020
	Panel B: Number of shops (log) (in same shopping street)				
Number of shops within:	Street *(baseline)*	50m	100m	150m	200m
Number of shops (*log*)	-0.00527***	-0.00624***	-0.00926***	-0.00842***	-0.00753***
	(0.00129)	(0.00186)	(0.00181)	(0.00164)	(0.00155)
Log-likelihood	-44,482	-44,373	-44,447	-44,469	-44,473
Observations	220,020	219,503	219,974	220,020	220,020
Shop characteristics	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes
Shopping street characteristics	Yes	Yes	Yes	Yes	Yes
Intersection fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: The number of shops includes all shops within each distance threshold that are located on the same street and in the same year as each shop observation. Reported coefficients are average marginal effects. Footfall is measured as the number of shoppers per day. In columns (1)-(5), we include observations within 50m of a shopping street interaction. Robust standard errors clustered at the shopping street level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

streets, the coefficient of log footfall should diminish. Thus, it appears that regardless of the area we use for the shop density measures, the latter cannot capture the heterogeneity of shops in generating shopping externalities.

In Panel B of Table C17, we only include the log number of shops in the same shopping street for the different distance thresholds, as a proxy for shopping externalities. Column (1) shows the baseline results when we include the log number of shops at the whole street where the rent transaction took place. Columns (2)-(4) are based on the same distance thresholds used in Panel A. In all columns, the elasticity of rents with respect to the number of shops is statistically significant and comparable to the results reported in Section V.A.

We also repeat the same exercise for vacancies. Panel A, Table C17 shows the results when we include both the log average annual footfall and the log number of shops that are located on the same street as a shop that we observe (if it is vacant or not), for different distance thresholds. The coefficient of the log number of shops is statistically significant but very small between 50 and 150m, while for 200m or for the whole street, it is not even statistically significant. Moreover, the coefficient of log average footfall is essentially the

same as in the baseline specification (column (1)). Panel B, Table C17 reports the results when we only include the log number of shops in different distance thresholds. The results show that for each distance threshold chosen the elasticity of vacancies with respect to the number of shops is much smaller than the same elasticity with respect to footfall, suggesting that number of shops is an imperfect measure of shopping externalities.

In Table C18 we repeat the same exercise, but replace log rent by the incidence of a shop to be vacant. The results confirm the previous findings for rents: we find that footfall has a much stronger effect than the number of shops, irrespective of the chosen distance threshold.