Bouncing with the Joneses? A neural network approach to consumption neighborhood effects

Erik Grenestam*

April 2019

EARLY DRAFT VERSION

Abstract

If one of your neighbors decides to purchase a good, does that affect your decision to buy the same good? I study this question using household level data collected in a novel way. I use an image classification algorithm to process a large set of aerial photos in order to infer household ownership of a visible durable good, specifically a trampoline. To estimate the neighborhood effect, I use a neighborhood fixed effects approach together with exogenous variation stemming from new neighbors moving in. I find that neighborhood effects are present, but only over short distances (the five closest neighbors) and only for prior non-owners. Information transmission between neighbors seems to be an important underlying channel. A back-of-the-envelope simulation suggests that the effect is large enough to have a significant impact on demand for this particular product.

JEL classifications: D12, H23, H42

Keywords: Households, Consumption, Neighborhood, Peer effect, Machine learning

---

*Department of Economics and Center for Economic Demography, Lund University School of Economics and Management. I am grateful for financial support from Jan Wallanders and Tom Hedelius foundation and Stiftelsen för främjande av ekonomisk forskning vid Lunds universitet. I would also like to thank Mats Högström (SLU Umeå) for generously providing the images used in this paper.

1 Introduction

If one of your neighbors decides to purchase a good, does that affect your decision to buy the same good? In this paper, I study this question using household level data collected in a novel way. I examine a large set of aerial photos to infer household consumption of a good, specifically a trampoline. To collect a sufficiently large data set in a cost-effective manner, I automate the data collection process using a state-of-the-art neural network trained to identify trampolines. Neural networks are a type of non-linear classifiers[1] capable of advanced pattern recognition. They have revolutionized (Krizhevsky et al., 2012) algorithmic image classification and are now rivaling human performance in some real-world classification tasks (Haenssle et al., 2018). This study provides an example of how neural networks can be applied to accurately collect economically relevant data from an image-based source on a scale which was infeasible just a few years ago.

A common motivation for the study of peer effects [2] is the discovery of social multipliers. For example, if low-performing pupils impact their classmates negatively, addressing the needs of these students may generate additional benefits by improving the outcomes of their peers (Lavy et al., 2012). Take-up of government programs can increase without the need for expensive information campaigns provided that information about the reform is transmitted between peers (Dahl et al., 2014).

In the consumption literature, the bulk of empirical research on peer effects concerns total household consumption expenditures and it's life-cycle pattern. While total consumption expenditures is arguably the most relevant outcome when examining the macro-economic implications of social multipliers, studying demand for specific goods is necessary to gain an understanding of the nature and the underlying mechanisms of peer effects (De Giorgi et al., 2016). As Moretti (2011) notes, the limited empirical work on spillovers in the consumption of specific goods is likely caused by the lack of data. The ideal data set does not only need to record individual-level purchases of a specific good, it must also contain information on a relevant network and the purchases made by peers. Additionally, the data set needs to be sufficiently large to allow the researcher to apply the tools necessary to overcome the econometric obstacles to identifying peer effects.

In this study, I use a novel technique to construct a unique large-scale data set on

[1] In a machine learning context, a classifier refers to a model or algorithm tasked with assigning a label to e.g. an image from a predetermined set of labels.

[2] Here, I use "peer effect" and "network effect" interchangeably

household-level trampoline ownership. By observing the exact location of each residence, I am able to form a complete picture of neighborhood ownership. By using multiple consecutive observations of the same property, I am able to (roughly) identify the time of purchase as well as exploit within-household changes to implicitly account for many confounding factors.

My main identification strategy relies on the conditionally exogenous change in household exposure to neighboring trampolines that occur when an old neighbor moves out and a new one moves in. By comparing incumbent households living in the same small area, where one is treated with a new neighboring trampoline, I am able to identify the effect of an additional neighboring trampoline on the subsequent purchase decision of the household. My results indicate that a neighborhood effect is present, as an additional neighboring trampoline increases the probability of next-period ownership by about 1.4 percentage points.[3] A back-of-the-envelope simulation suggests that the neighborhood effect has an economically significant impact on demand for this particular product.

Conspicuous consumption and keeping-up-with-the-Joneses-effects are perhaps the most intuitively appealing explanations of peer effects. De Giorgi et al. (2016) find evidence of peer effects between co-workers which are broadly consistent with status-seeking explanations. However, when looking into expenditures for specific categories of goods, the authors find no evidence in favor of household demand shifting in favor of goods consumed by, or visible to, peers. However, for the latter exercise, the authors rely on a small sample consumer expenditure survey and they note that their approach is likely underpowered.

There are complementary explanations of peer effects that directly relate to the characteristics of the individual good, such as social influence and information transmission (Moretti, 2011; Grinblatt et al., 2008; Dahl et al., 2014) as well as network externalities (Gilchrist and Sands, 2016). Grinblatt et al. (2008) studies neighborhood effects in the context of Finnish car purchases. They find that neighborhood effects are present, but that information transmission between neighbors is a more likely explanation than status-seeking. Kuhn et al. (2011) studies effects on neighbors of Dutch lottery winners. Consistent with Grinblatt et al. (2008), they find neighborhood spillovers in car purchases. From a welfare perspective, it is important to understand the underlying mechanisms. If

[3] The baseline ownership rate is 13 percent.

consumer demand is distorted due to a desire to maintain even footing with a reference group, limiting peer effects may improve welfare by relaxing constraints on demand and utility. However, if peers effects reflect social learning about different products, their presence is likely beneficial to consumers.

This study complements previous work by applying a new data collection method in order to study a new type of product. Several studies have focused on cars purchases. From a macro perspective, cars represent an important consumer durable. As such, any social multiplier may have significant impact on the macro economy. However, if we want to deepen our general understanding of peer effects in consumption, it seems fruitful to broaden our focus. A new car is an outlier in terms of e.g. it's share of total consumer expenditures. As such, a car purchase is likely to be preceded by months of deliberation and information gathering. However, Grinblatt et al. (2008) finds that neighborhood effects are limited to within 10 days of a purchase. The authors speculate that a neighbor may "tip the scale" in favor of a purchase by providing information about e.g. pricing at local dealerships. For a relatively inexpensive and homogeneous good like a trampoline, a purchase is likely more spontaneous and the relevant mechanisms may be different.

While the choice to study trampolines may seem peculiar, I argue that a trampoline has a set of unique characteristics that makes it suitable in the context of neighborhood effects. First, it is a visible good designed for outdoor use.[4] As a trampoline is highly visible to neighbors, neighborhood effects can arise even under limited social interaction between neighbors. Second, it is durable, meaning that neighbors are likely exposed to it for a long time.[5] Third, it is a product used by and marketed to children. I conjecture that children are more likely to have local social network than adults, increasing the likelihood that the neighborhood is a relevant network to study. Given these characteristics, I argue that if neighborhood effects exist, a trampoline is likely to be subject to them. In addition to visibility, another key characteristic for my study is that for safety reasons, a trampoline is unlikely to be placed next to walls or trees. This makes them easy to observe in an aerial photograph.

As with any study concerning a specific good, there are idiosyncrasies that may limit the external validity of my findings. Regarding the underlying mechanisms, we have to

[4] The most popular trampolines listed on the leading Swedish price-comparison site have a diameter of 3-4 meters and a price of 2,000 to 10,000 SEK (240 to 1,200 USD).

[5] Kuhn et al. (2011) interpret their findings as visibility and durability being crucial for generating neighborhood spillovers.

consider the implications of a good that is marketed to children. To an adult, a trampoline is arguably not desirable from a status-seeking perspective. However, to the extent that the preferences of the child enter into household decision-making, through parental altruism or as outcomes of a bargaining process, status-seeking behaviour may manifest. My prior is that a neighborhood effect, if present, should be net positive. However, one could consider a trampoline an (impure) public good, meaning that a neighboring trampoline may act as a substitute to a self-owned. Ultimately, my reduced form estimates are only informative about the net marginal effect. However, I find some evidence suggesting that a neighbor's trampoline can act as a substitute in certain scenarios.

Another contribution of this paper is to serve as an example of how neural networks can be used to tap into a previously inaccessible data source. An increasing share of all data generated by humans comes in the form of images, be it aerial photos or images posted on social media. Up until very recently, interpreting and classifying images was largely a labor-intensive manual task. This has prevented all but the most well-funded researchers to transform images to a form that allows for conventional analysis. Over the last few years, neural networks have been at the foundation of large performance gains in computer vision. Their rise in popularity have had the added benefit of easy access and implementation. Expensive hardware and a degree in computer science are no longer prerequisites for their use. So far, there are few economic studies using neural networks and computer vision. Henderson et al. (2012) uses satellite imagery of night-time illumination to improve measures of growth in developing countries. However, information is inferred directly from individual pixels and no interpretation or pattern recognition is necessary. Gebru et al. (2017) uses Google Street View photos of parked cars and predict their make and model using a neural network. They use the composition of vehicles to predict the socio-economic composition of a neighborhood. They find that their derived measures of e.g. median income correlate well with survey data and that their approach has advantages in terms of cost and speed of measurement.

The paper is organized as follows. The next section presents the data along with the output and performance of the neural network (further information is presented in appendix A.1 and A.2). Section 3 discusses my empirical framework and identification. Section 4 presents my results and section 5 concludes.

2 Data

My primary data source is a set of aerial photos of Swedish neighborhoods. The photos are produced by the Swedish mapping, cadastral and land registration authority (*Lantmäteriet*). The images cover the major urban centers and surrounding areas. My specific subset of images are selected by virtue of containing a large number of single-family homes, which is my main unit of observation.[6] The photos are taken between 2006 and 2018. The observation weighted mean time between consecutive photos of a property is 2.5 years. A property is on average photographed 3.8 times during my sample period.

By slicing the raw images into smaller "chips", each containing a single property, I am able to apply the neural network to infer the presence of a trampoline on any given property (see section A.1 for details on image preprocessing). Ultimately, the neural network produces a binary vector where 1 indicates the presence of a trampoline on a particular property at the time the photo was taken. To be able to map trampoline ownership to a single household, I exclude other forms of housing tenure from my sample (e.g. apartment buildings). Around 60 percent of Swedish households with young children reside in a single-family home. For most families residing in apartment buildings, a trampoline is simply not feasible.

To the trampoline ownership data I add data on real estate transactions, including the property coordinates, price and date of sale as well as the size of the home. The coordinates and date of sale allows me to match the transaction data to observations on trampoline ownership for a specific property. By comparing the transaction date to the photo date, I infer whether or not a new neighbor has moved in just prior to the photo date. If, between two consecutive photos, no transaction is recorded, I infer that the same household resides in that home in both periods. This allows me to exploit within household changes in my model.

In addition to property level data, I have data based on two larger spatial groupings, grid square and district. I add data data on demographics and median income at a spatial grid level. The grid is the highest spatial resolution for household data available from Statistics Sweden.[7] In urban areas, the grid has a resolution of 250 by 250 meters and in rural areas the resolution is 1 by 1 km. This data is available from 2014 until 2017,

[6] table A.2 presents the distribution of property lot by year observations at the municipal level.
[7] Coordinate level data on e.g. income is typically not provided to researchers.

Table 1: Summary statistics

	Unique obs.	Mean	SD	Min	Max
Houses per district	4,277	274	223	1	2,324
District population (2017)	4,277	1,717	425	662	3,297
Share ages 0-15 years	4,277	.194	.0524	0	0
Grid square population (2017)	61,280	83.7	111	3	1,768
Share ages 0-15 years	61,280	.193	.136	0.00	1
Grid square median income**	57,918	231,783	123,294	196	15,721,554
Trampoline (T) present	4,482,548	.129	.335	0	1
Neighbors (N) with T	4,482,548	.628	.819	0	5
Nbrs moved in, past 60 days	4,482,548	.0275	.171	0	5
Nbrs moved in, past 200 days	4,482,548	.101	.335	0	5
ΔT_{nbrs} due to new nbr, 60 days*	3,186,740	-.000402	.0651	-2	2
ΔT_{nbrs} due to new nbr, 200 days*	3,186,723	.00118	.127	-4	2
Transition from T to no T	.	.057	.232	0	1
Transition from no T to T	.	.0767	.266	0	1
No. of properties	1,173,124	.	.	.	.
Property area (sqm.)	1,173,124	921	556	21.00	6,005
No. of households	1,305,929	.	.	.	.
Photos per household	1,305,929	3.43	1.27	1	7
House price (2010 SEK)	178,085	2,407,312	1,802,919	1,091	45,085,713
House area (sqm.)***	170,091	125	39	10	283

Neighborhood defined as the 5 closest neighbors. *Conditional on same owner in $t-1$ and t
**2017, in 2010 SEK ***Winsorized at the top 1 percent due to irregularities in the data.

which is the later part of my main sample period (2006-2018).

My main spatial grouping are districts. Districts[8] are constructed by Statistics Sweden and subdivides Sweden into about 6,000 areas, each containing 700 to 3,300 residents. Rather than aiming for a fixed population size, the district borders are drawn to account for local geographic features such as major roads, waterways, railroads and other infrastructure. This feature makes the districts ideal as a control for local economic conditions, advertising by local retailers and other supply side factors. My sample covers 4,277 districts and each district contains on average 274 single-family homes (see table 1 for summary statistics).

2.1 Neural network performance

I detail the image prepossessing and the training of the neural network in sections A.1 and A.2. Table 2 summarizes final model performance when tasked with finding trampolines in a manually labeled set of test images ($N = 1,819$). While the model performs well in terms of recall (the share of labeled trampolines found), the precision (the share of

[8] Officially called Demographic Statistics Zone (DeSO)

Table 2: Model performance, probability threshold = 0.5

		Label ($N = 1,819$)		
		Trampoline	No trampoline	
Prediction	Trampoline	241	34	Precision = 87.6 %
	No trampoline	7	1,537	
		Recall = 97.2%	FPR = 2.8 %	F1-score = 92.2%

predicted trampolines that are correct) is somewhat lacking. The reason for this is that trampolines are rare, and there are many negatives that the model has to get right. When inspecting images where the model produces a false positive, I find that the model is sometimes fooled by similar structures, such as a round jacuzzi or a gazebo. Since I rely on within-household variation to identify effects, a prediction error that is due to some permanent feature of the property is likely to be repeated over time meaning that the error will not contribute to the identifying variation.

The raw output of the model is a vector of probabilities, where probabilities close to 1 represent a high degree of certainty that a trampoline is present. As figure 1 illustrates, there is a trade-off between precision (type 1-errors) and recall (type 2-errors). In other words, we can make the model find fewer false trampolines by increasing the probability threshold, but in doing so I will also increase the likelihood that the model misses a few actual trampolines. As the figure shows, increasing precision entails a substantial drop in recall. In my predictions, I use the default 0.5 probability threshold to infer trampoline ownership.

It is important to relate performance to alternative approaches. Typically, consumption data is obtained through surveys which are known to be less than perfectly accurate. One Swedish study found underreporting of car purchases by up to 30 percent (?). Given that car expenditures make up a large fraction of household consumption and that a new car is highly salient, this is a sobering result. A survey of trampoline expenditures is likely to be burdened by even higher error rates.

Another relatively common source of consumption data is credit card statements (e.g. Agarwal et al., 2017a). While this data partially addresses measurement error, it typically does not cover the full universe of transactions and researchers are often limited

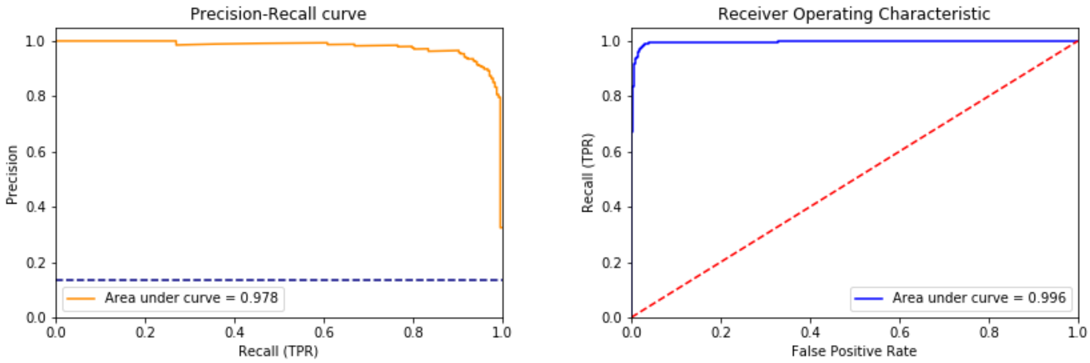

Figure 1: The two panels illustrate the trade offs between precision and recall (left) and between precision and false positives (right) by plotting the respective statistics for a number of different probability thresholds. The dashed line on the left is the maximum precision possible a model with complete overlap between the two classes (a coin-toss prediction) and is given by the ratio of *Trampoline* to *No trampoline* images in the test set (13 %). To the right, a coin-toss prediction will be located somewhere along the dashed line.

to customers of a specific financial institution. This is problematic for the study of network effects, since the researcher is not able to get a complete picture of the relevant network.

Figure 2 shows mean trampoline ownership rates as predicted by the neural network. To avoid confounding the trend with sample selection due to the rotation of houses being photographed each year, I plot the trend for the six most common photo-year sequences. All of them depict an upward trend in trampoline ownership in the sample, going from under 10 to around 16 percent. The increase in popularity is corroborated by data from Google Trends. As shown by Figure 3, the number of searches for "trampoline" has increased steadily since 2012. There is a strong seasonal component, with Sweden's cold winters freezing the public's interest between October and March. Mentions of trampolines in news media has declined since the peak in 2010.[9]

To examine the roles of income and demographics as determinants of ownership, I add grid level data on income and population. Figure 4 presents ownership rates by deciles of income and share of the population aged 0-15 years. The income gradient is overall positive, although locally declining between the 7th and 8th decile. This decline could be caused by properties in more affluent urban areas being smaller, decreasing the feasibility

[9] Data from Retriever Media Archive. Between 2000 to 2004, mentions of trampoline are much less frequent. The recent decline is likely due to the fading novelty of trampolines as a common household product. In addition, EU regulation introduced in 2014 made safety enclosures a mandatory feature of large trampolines. This has likely decreased injuries, which used to be a common theme in media coverage.

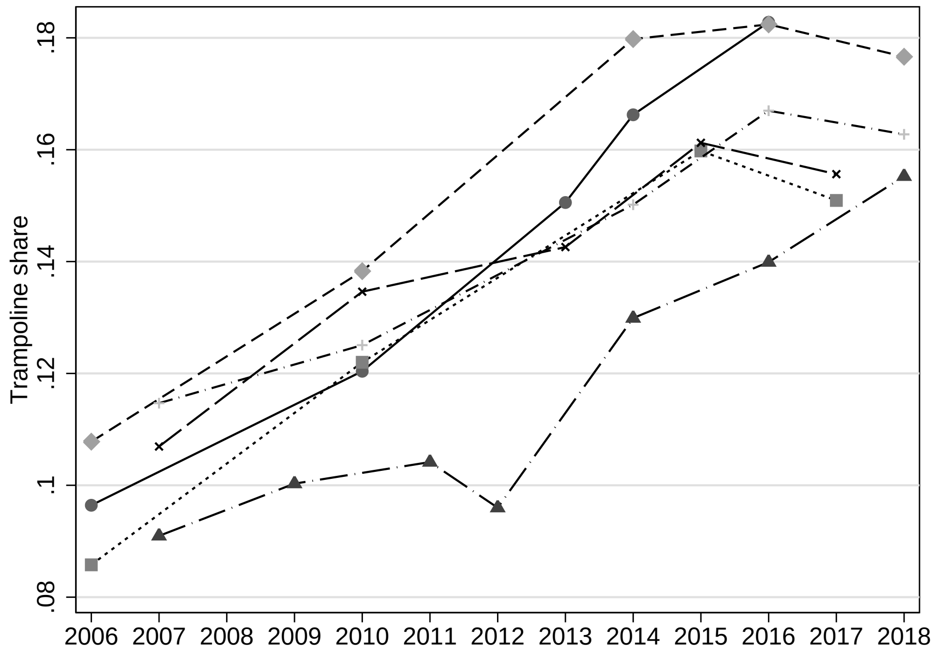

Figure 2: Mean share of homes in the sample with a trampoline, as predicted by the neural network. Lines represent the most common photo-year sequences.

of a trampoline. The apparent lack of a strong income effect is likely explained by the fact that most households we observe are relatively well-off and that a trampoline makes up a small share of the household budget. As expected, trampolines are much more common in areas where there are more children residing. The ownership rate increases threefold between the first and the tenth decile of the population share made up of ages 0-15 years. (second panel of figure 4).

3 Empirical framework

My goal is to estimate the effect of neighbors' decision to acquire a trampoline on the household ownership decision. This outcome-on-outcome effect belongs to what Manski (1993) refers to as an endogenous social effect or what is commonly referred to as a peer effect. Concerning peer effects, a large body of both theoretical and empirical work deals with the issue of identification. The identification issues stem from the fact that peer groups are rarely formed by chance alone. Formation is often heavily influenced by sorting and self-selection. The formation process therefore tends to produce groups whose members have similar characteristics and preferences and are subject to similar

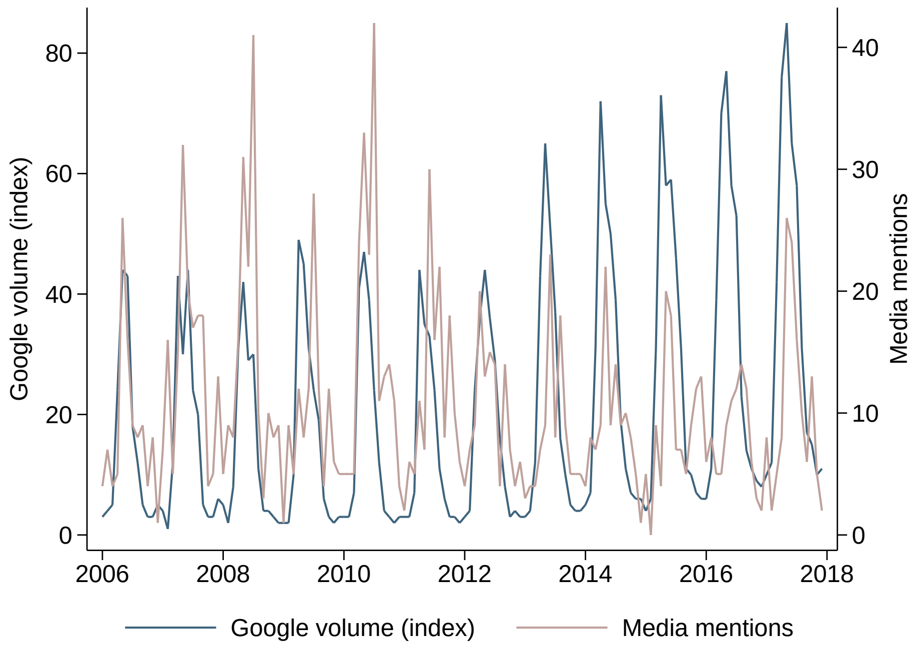

Figure 3: Trends in Google search volumes and media mentions for the keyword *studsmatta* (*trampoline*). Monthly observations.

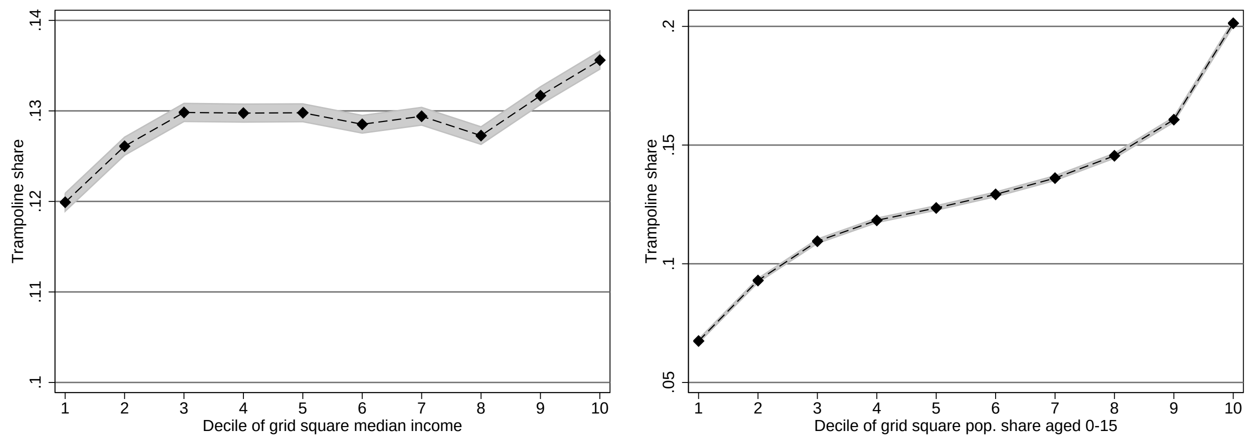

Figure 4: Deciles of grid level income (right) and share of children (left) are calculated as means during 2014-2017. Data on trampoline ownership from 2006-2018. Shaded blue areas represent the 95 % CI.

shocks.[10]

Further jeopardizing a causal interpretation of estimates is what Manski (1993) calls "the reflection problem". In a simple contemporaneous outcome-on-outcome regression, the direction of causality cannot be inferred unless the researcher can distinguish action from reaction. For example, is a student performing better in school because their peers have improved, or have peers improved because of the student's change in behaviour? Due to this simultaneity of outcomes, we cannot interpret the estimated coefficient as a social interaction effect but have to rely on e.g. instruments for the outcome of peers.

My point of departure is the model explored by Bramoullé et al. (2009) in their paper on identification of peer effects in social networks:

$$y_{it} = \alpha_i + \beta W_i \mathbf{y}_t + \gamma \mathbf{X}_{it} + \delta W_i \mathbf{X}_t + \varepsilon_{it} \tag{1}$$

Here, the outcome observed for household i in period t is a function of $\mathbf{y}_t$, the $n \times 1$ vector of contemporary outcomes for all other households. Equation (1) is inspired by the spatial econometrics literature which commonly makes use of a spatial weight matrix $\mathbf{W}$ to capture the spatial arrangement of units. $\mathbf{y}_t$ is multiplied by W_i, the i^{th} row in the $n \times n$ spatial weight matrix $\mathbf{W}$. The spatial weight matrix is partially given by the local geographic arrangement of homes, but some further modelling is required. In the next section, I explain how I construct $\mathbf{W}$. β represents the *endogenous* social interaction effect that is the focus of this paper, i.e. the effect of my neighbor getting a trampoline on my own purchase decision. The specification in Bramoullé et al. (2009) also includes the *exogenous* social effect δ, i.e. the effect of my neighbors exogenous characteristics on my own purchase decision. While there is not much intuition for this type of effect for this particular good, one could imagine e.g. that the purchase decision could be affected by the ages and number of children that live next door.

In the absence of correlated effects, i.e. innovations to ε or $\mathbf{X}$ that are common to the neighborhood, Bramoullé et al. (2009) shows that non-overlapping peer groups, formalized as the linear independence of $\mathbf{I}$, $\mathbf{W}$ and $\mathbf{W}^2$ is sufficient to be able to identify $(\alpha, \beta, \gamma, \delta)$. Another way of stating this condition is that my neighbors have to have neighbors who are not my neighbors, meaning that we can form an *intransitive triad* of households. This provides a set of exclusion restrictions, as any shock to the neighbors of my neighbors

[10] What Manski (1993) refers to as correlated effects or an exogenous peer effect.

may only affect me through my neighbors. Intransitive triads have been used to identify network effects in empirical research (De Giorgi et al., 2016). While the presence of intransitive triads is ultimately an empirical question, it is hard to infer their existence from from geographic arrangement alone. Therefore, I opt to use a source of exogenous variation in $W_i\mathbf{y}_t$ stemming changes to the neighborhood composition. Specifically, I use changes in the neighborhood trampoline composition caused by new neighbors moving in.

As pointed out by Angrist (2014), even if β is identified in an econometric sense, it can only be interpreted as a social interaction effect if the error term is not correlated within neighborhoods. To address this, Bramoullé et al. (2009) proposes the inclusion of a neighborhood fixed effect, α_i, which absorbs the effect of time-invariant unobservables. As my data includes the exact location of each household, every household forms the basis for a neighborhood. A household fixed effect α_i therefore subsumes a neighborhood fixed effect.[11] To control for area-specific trends, I include a full set of district by year fixed effects.

According to Angrist (2014), the most compelling studies of peer effects uses variation in peer group composition that is orthogonal to baseline characteristics. In the spirit of this ideal, I exploit variation stemming from new neighbors moving in as a source of plausibly exogenous change in household trampoline exposure (from the perspective of incumbent neighbors). For a discussion of the identifying assumptions behind this approach, see section 4.2. Under the assumption that a new neighbor moving in constitutes a (conditionally) exogenous change in the trampoline profile of a neighborhood, I use the following specification to estimate the neighborhood trampoline effect:

$$\Delta y_{it} = \alpha + \beta D_{it-1}\Delta\mathbf{y}_{t-1} + \gamma\mathbf{X}_{it} + \varepsilon_{it} \tag{2}$$

The key difference between equation (2) and (1) lies in how I specify treatment, $D_{it-1}\Delta\mathbf{y}_{t-1}$. The first part, D_{it-1}, is the i^{th} row in the *move* matrix $\mathbf{D}_{t-1}$. $\mathbf{D}$ is a sparser extension of $\mathbf{W}$. If we define an element in w_{ij} of $\mathbf{W}$ as taking a value of 1 if j is a neighbor of i, then an element d_{ij} of $\mathbf{D}$ takes on a value of 1 if j is a neighbor of i *and* household j moved-in just before the photo was taken. The second part, $\Delta\mathbf{y}_{t-1}$, is motivated by my assumption that for a social interaction effect to arise, we need incumbent households

[11] The two are not equivalent since the neighborhood persists through households moving in and out.

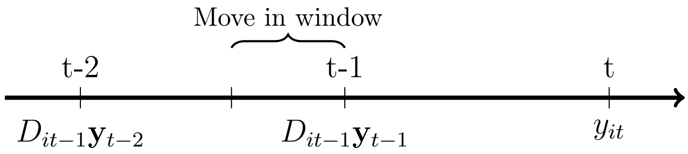

Figure 5: Timeline illustrating my empirical approach of using new neighbors to identify the neighborhood effect.

exposure to trampolines to change. I argue that this is true for both a social learning and a status-seeking channel. Consequently, if a new neighbor does not contribute to a change in the neighborhood trampoline profile, there is no treatment ($D_{it-1}\mathbf{\Delta y}_{t-1} = 0$). Thus, treatment is defined as a trampoline (no-trampoline) household moving into a home where a no trampoline (trampoline) household used to live. Therefore, I specify the equation in first differences, which implicitly accounts for a household fixed effect.

Another important aspect of (2) concerns the issue of simultaneity and timing of the effect. As I conjecture that a new neighbor constitutes a conditionally exogenous change in the neighborhood trampoline profile, specifying a contemporaneous effect of $D_{it}\Delta\mathbf{y}_t$ would implicitly assume a unidirectional effect in the sense that incumbent neighbors can be affected by new neighbors in the same period, but new neighbors are unaffected by the incumbent. By lagging treatment, I give incumbent neighbors time to react to the change in neighborhood trampoline profile. Hence, we measure the reaction of household i in period t to a change in the trampoline profile that occurs between $t-2$ and $t-1$. Grinblatt et al. (2008) presents a similar argument for dealing with the reflection problem. If treatment precedes the outcome, there is limited scope for reverse causality.

As a secondary approach, I estimate a specification similar to Grinblatt et al. (2008). By comparing effects of near and far[12] neighbors, the authors effectively difference-out unobservable trends in outcomes common to a neighborhood. While highly localized[13] *unobservable* differences may still bias the estimates, the authors argue that since *observable* differences between near and far neighbors are small in their sample (using a rich set of socio-economic variables), it is unlikely that any unobservable differences can explain a significant treatment effect. I implement this approach using the following quasi-triple-difference specification:

[12] Sorted by distance, the authors use neighbors in the intervals 1-5, 6-10, 11-50 etc.
[13] Meaning a confounder only shared between close neighbors

$$y_{it} = \alpha_i + \sum_{k \in \mathcal{K}} \beta_k W_{ki} \mathbf{y}_{t-1} + \gamma \mathbf{X}_{it} + \varepsilon_{it} \tag{3}$$

where $\mathbf{X}_{it}$ is a full set of district by year effects. The specification compares within household changes to ownership across households residing in the same district. By including subsets of neighbors (k) separately, I am able to compare the effect of close (e.g. β_{1-5}) to more distant neighbors (e.g. β_{16-20}).[14] Under the assumption that there are no unobservable differences between near and far neighbors related to trampoline demand, (3) identifies the neighborhood effects.

As in the previously described approach, while lagging the treatment variable plausibly solves the simultaneity problem, a dynamic panel data model brings it's own host of estimation issues. For a sufficiently small neighborhood, y_{it-1} may impact the aggregated neighboring trampoline composition though a neighborhood effect, meaning that $W_i \mathbf{y}_{t-1}$ is a function of y_{it-1}. The presence of a lagged dependent variable of the right-hands side of (3) along with a household fixed effect causes the error term u_{it} to be correlated with $W_i \mathbf{y}_{t-1}$, even if the latter is pre-determined. This makes OLS an inconsistent estimator of β in (3) (Nickell, 1981). In general, $\hat{\beta}_{\text{OLS}}$ will have a downward bias proportional to $1/T$. One route to a consistent estimate of β is due to Arellano and Bond (1991). Simply put, the Arellano-Bond GMM-estimator transforms the level equation (3) to first differences and uses past levels and differences (e.g. $\mathbf{y}_{t-2}$ and $\Delta \mathbf{y}_{t-2}$) to instrument $\Delta \mathbf{y}_{t-1}$.

In my main approach, rather than using all past realizations as instruments, I use a single contemporary "2SLS-style" instrument for treatment $\Delta \mathbf{y}_{t-1}$ to estimate the reduced form equation (2). Under the assumption that $D_{it-1} \Delta \mathbf{y}_{t-1}$ is uncorrelated with the transformed error term Δu_{it}, the change due to new neighbors moving in just prior to the beginning of $t-1$ is a valid instrument and my estimate of β is consistent.

Regarding the the dynamic nature of (3), I argue that under my prior of a positive neighborhood effect, the downward bias likely present in OLS estimations of β in (3) will give me conservative estimates. As a robustness check, I also estimate (3) using a theoretically consistent GMM-estimator (Arellano and Bond, 1991; Blundell and Bond, 1998; Roodman, 2009).

[14] This post-estimation comparison is the third differential, hence I call it a "quasi-triple-diff".

3.1 Modelling neighborhoods

The role of the spatial weight matrix $\mathbf{W}$ in (3) and (2) is to describe which households are to be considered neighbors in the sense that their respective trampoline ownership status is communicated (directly or indirectly). As noted by Anselin (2009), there is little in the way of general guidance in how $\mathbf{W}$ should be specified. Researchers commonly use geographic distance and contingency. Rather than starting from an arbitrary choice of $\mathbf{W}$, I opt for a more data-driven approach. Suppose that the relationship between the trampoline status of household i and its M closest neighbors can be described as:

$$y_{it} = \alpha_i + \sum_{j=1}^{M} \beta_j y_{jt} + \varepsilon_{it} \tag{4}$$

Where β_j represents the effect (in a purely correlational sense) that household j's status has on household i. To test the role of distance between neighbors in determining between household correlation, I apply the Least Absolute Shrikage and Selection Operator (LASSO) to (4). The LASSO is a commonly used prediction and model selection tool developed by Tibshirani (1996) that imposes a measure of regularization to standard OLS in order to prevent overfitting caused by superfluous variables. Formally, LASSO can be defined as the following Langrangian optimization problem:

$$\beta_L = \underset{b_L}{\arg\min} \{ \frac{1}{N} ||\mathbf{y}* - \alpha - \mathbf{Y}^* b_L||_2^2 + \lambda ||b_L||_1 \} \tag{5}$$

where $\mathbf{y}^*$ is an N-vector of household ownership status, $\mathbf{Y}^*$ is the $N \times M$ matrix describing the ownership status for the M closest neighbors to each household and b_L is an M vector of parameters to be estimated by the LASSO. Prior to the LASSO estimation, I partial out household and year fixed effects ($\mathbf{y}^*$ represents the transformed $\mathbf{y}$ vector). The second term within braces is the regularization term, operating on the L1-norm of b_L (the sum of absolute values). The Lagrange multiplier λ is a penalization coefficient, determining the relative importance of shrinking $||b_L||_1$. Because the LASSO uses L1-regularization, the solution will force a subset of parameters in β_L to be exactly equal to zero. This is the model selection property of the LASSO that allows me to determine which neighbors are the most important predictors of a within-household change in trampoline status. In the estimation procedure, λ is a free parameter to be determined by the researcher. To explore the role of geographic distance in determining the size of a neighborhood, I use a range of values for lambda. While there are several methods for finding an optimal value[15] for λ, for this exercise I believe it is more informative to determine how the β_L evolves under different levels of regularization.

I consider the 40 closest neighbors to each household in my sample ($M = 40$).[16] Figure 6 presents the LASSO results. Two noteworthy patterns emerge. First, even within this very local scope, distance seems to matter. For the largest penalty, only the closest two neighbors are selected. As λ decreases, selection seems initially follow the distance ordering. Second, as we get past the 15 or so closest neighbors, the selection of new neighbors seems less related to distance. This could indicate that neighborhood effects become less important than e.g. supply side effects common to a greater area. Based on my interpretation of figure 6, I consider $M = 5$ as my baseline neighborhood definition. However, it should be emphasized that the LASSO selection does not constitute evidence of a neighborhood effect. Figure 6 simply confirms that distance matters for the correlation of outcomes and is at best suggestive of the scope of a true neighborhood effect. As a robustness check, I explore how my results are affected by increasing the size of my neighborhoods past the five closest houses.

By setting $M = 5$, I standardize the number of non-zero elements w_{ij} in each row i of $\mathbf{W}$ to 5.[17] However, as some houses straddle the border of the area covered by the aerial photography, one or more of their neighbors may fall outside the area of coverage. As a result, some homes will have less than 5 effective neighbors.[18] To reduce measurement error, in my main analysis I exclude households with less than all neighbors covered by the aerial photo. As for the value of each element w_{ij}, I simply count the number of neighboring trampolines ($w_{ij} = 1$ if j is a neighbor of i). For convenience, I introduce the following notation for the number of neighboring trampolines to household i among the M closest neighbors:

$$C_{it}^{M}(\mathbf{y_t}) \equiv \sum_{j=1}^{M} \mathbb{I}[y_{ijt} = 1] \tag{6}$$

[15] This is typically done by minimizing prediction error or an information criteria.

[16] While the number 40 is arbitrary, what is critical is picking a number larger than the number of neighbors one could reasonably expect each household to interact with.

[17] I do not consider households to be an neighbor of themselves, $w_{ii} = 0$.

[18] 92 percent of all homes have all 5 neighbors covered

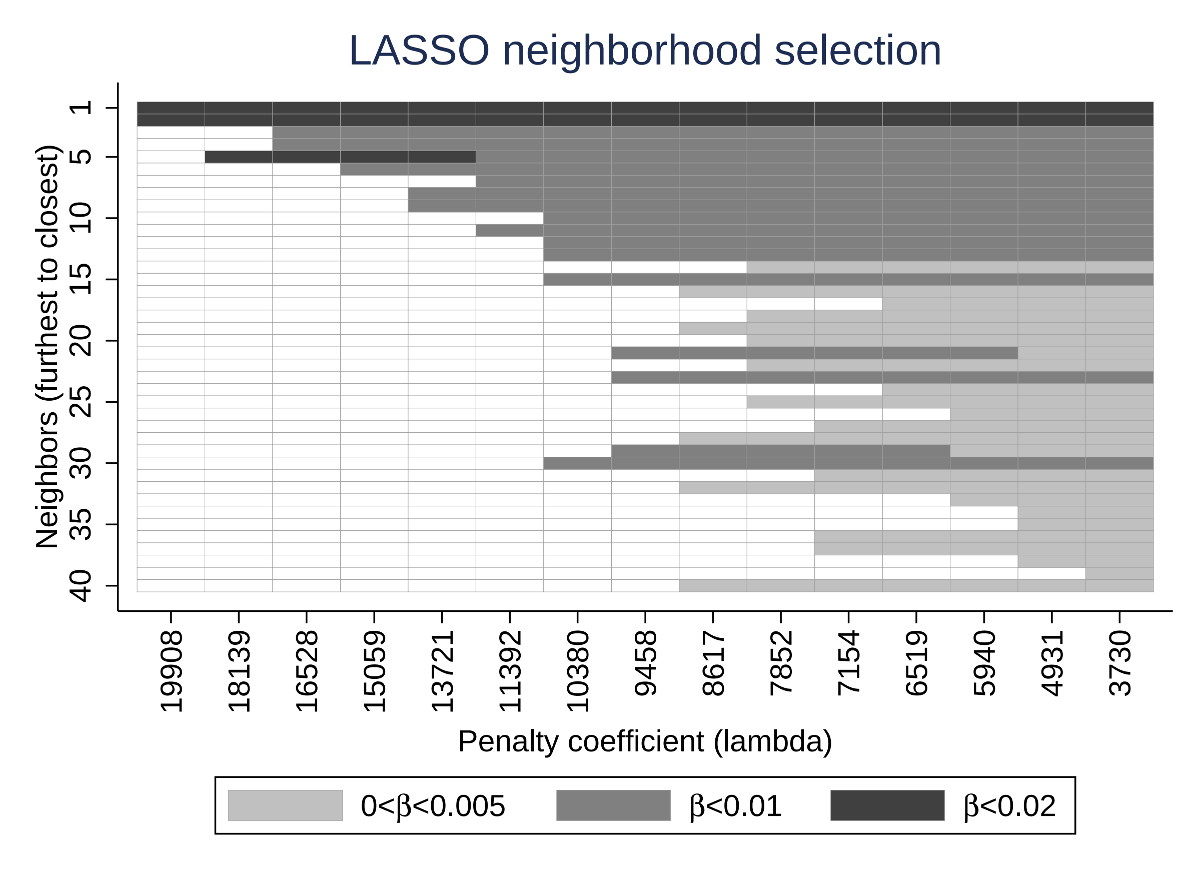

Figure 6: The figure depicts the subset of neighbors selected by estimating equation (5) using a range of different values for the penalization coefficient λ. The values used is the default sequence in the Stata package LASSOPACK. White areas indicates that the corresponding parameter is set to zero. Shaded areas indicate the value of the corresponding parameter estimate in an OLS regression on the LASSO-selected subset of neighbors.

On average, the centroid-to-centroid distance to the fifth neighbor is 61 meters (Figure A.5 describes the distance distribution to the 15 closest neighbors). My relatively narrow definition of a neighborhood has some support in the literature as other studies have found that neighborhood effects seem to transmit only over short distances. Agarwal et al. (2017b) only finds effect among residents in the same residential building and Kuhn et al. (2011) find that effects are limited to neighbors "within two doors". Grinblatt et al. (2008) uses the 10 nearest neighbors as their main definition of a neighborhood.

4 Results

I begin this section by presenting results from the Grinblatt et al. (2008) approach using (equation 3). I then discuss and test the validity of using new neighbors as a source of exogenous variation in the neighborhood trampoline profile. Next, I present results using the variation stemming from new neighbors moving in. Finally, I put the effect size into context by way of a simple comparison.

4.1 Neighborhood fixed effects

My first set of results presents estimates from a difference in differences specification similar to the approach by Grinblatt et al. (2008). In addition to a household fixed effect, I account for correlated effects and common shocks by adding a full set of district by year fixed effects. I include neighbors 1-5, 6-10, 11-15 and 16-20 separately and estimate specifications with contemporaneous as well as lagged treatment. For marginal effects to be easily comparable across groups, I define treatment simply as the number of trampolines observed among each of the four groups.

Table 3 presents the results. We note that the positive and significant marginal effect of an additional trampoline among the five closest neighbors diminishes over distance and is not significant for neighbors 16-20 for the lagged treatment specification (column 2).[19] However, 0.2 percentage point increase in ownership probability for an additional trampoline among the five closest neighbors is arguably quite small.

To address potential Nickell bias due to the dynamic nature of the panel, Column 3 is estimated using the GMM estimator described by Roodman (2009). As this estimator is computationally expensive, I am forced to use a random 10 percent subsample of districts. Effects increase by an order of magnitude, consistent with the theoretically negative bias present in column 2. While the GMM estimator is sensitive to the exact choice of moment conditions, the effect size in column 3 is comparable to estimates from using variation stemming from new neighbors moving in. (see the following section).

While the fact that the effect diminishes even across very short distances could be attributed to a true neighborhood effect, we cannot rule out unobserved heterogeneity due to e.g. highly localized sorting within districts as a source of bias. However, I believe that this is an unlikely explanation. First, the mean distance between the 5^{th} and the 20^{th} neighbor is only 75 meters. To get a better idea of how similarity between households change over this short distance, I calculate the mean absolute deviation (MAD) between

[19] The difference between neighbors 1-5 and 16-20 is statistically significant in both columns.

each household and a group of neighbors K according to

$$M(X)_K = \frac{1}{N_K} \sum_{i}^{N} \sum_{k=1}^{N_K} |x_k - x_i| \tag{7}$$

By comparing means and mean absolute deviations between groups of neighbors, I can quantify how similarity to neighbors evolve as we move further away from the household. The major caveat is that my data at the household level is limited to just a few variables. The comparisons are presented in table 4. For house price and house area, I can only observe houses who are sold at some point during 2005 to 2018. We note that all three variables have a higher MAD when comparing households to neighbors 16-20 as opposed to neighbors 1-5. However, the bulk of deviations to household i are present already among neighbors 1-5 (column 1 of table 4). Relative to the initial deviation, the increase is small. In the bottom panel, we also note that the raw means are almost equal between the two sets of neighbors. Finally, I would like to direct the reader to table 3 in Grinblatt et al. (2008). They find that even over greater distances, neighbors remain broadly similar in terms of household-level demographics and income. Given the similarities between the Finnish and Swedish context, I would likely find similar patterns, had a richer set of observable been available. Table 4 is not consistent with the idea of neighborhoods as consisting of small enclaves that are internally homogeneous. The results in table 3 are likely not fully explained by unobserved local heterogeneity and omitted variable bias.

4.2 New neighbors

Do new neighbors constitute a source of exogenous variation of the neighborhood[20] trampoline composition? This key question merits some discussion. Because new neighbors self-select into homes and/or areas of residence, it would be naive to expect that the trampoline status of a new neighbor is orthogonal to the prior trampoline profile and preferences of the neighborhood. For example, parents with young children likely sort into neighborhoods with a similar demographic profile. However, the notion that a new neighbor can provide a *conditionally* exogenous source of variation in the neighborhood trampoline profile is plausible. To clarify this idea, consider the following example: The

[20] In this section, neighborhood is defined as the 5 closest houses to each individual household unless otherwise specified.

Table 3: Household FE regressions

Number of trampolines in t			
Neighbors 1-5	0.497*** (0.0293)		
Neighbors 6-10	0.172*** (0.0288		
Neighbors 11-15	-0.0247 (0.0285)		
Neighbors 16-20	-0.0623*** (0.0295)		
Number of trampolines in $t-1$			
Neighbors 1-5		0.190*** (0.0350)	2.409*** ((0.649))
Neighbors 6-10		0.0796** (0.0344)	1.101* (0.653)
Neighbors 11-15		0.0828** (0.0343)	1.196* (0.683)
Neighbors 16-20		0.0412 (0.0353)	0.820 (0.662)
Observations	4,227,707	2,860,365	292,271
Households	1,156,289	942,969	109,426
R-squared	0.540	0.620	.
Outcome mean	12.92%	13.29%	13.34%
Fixed effect	District × Year	District × Year	District × Year

The table presents estimates of equation (5). Columns 1 and 2 includes household fixed-effects. Column 3 uses the dynamic panel GMM estimator described by Roodman (2009). All coefficients are expressed as percentage points. Standard errors clustered on households. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Differences in MADs and means

	$M(X)_{1-5}$ (SE)	$M(X)_{16-20}$ (SE)	|Diff|
House area	24.14 (0.06)	28.57 (0.06)	4.43
House price	1,112,339 (2,637)	1,202,815 (2,820)	90,476
Property area	306.93 (0.52)	389.69 (0.57)	82.77
	Mean_{1-5}	Mean_{16-20}	|Diff|
House area	127.65	127.48	0.17
House price	3,170,184	3,148,675	21,509
Property area	992.08	984.20	7.88

Smith's are looking to buy a new home, to which they will be bringing their trampoline. While the Smith's have preferences over the type of neighborhood and house they would prefer to live in, they cannot pick any home. They are limited to the homes currently available on the market. The week after they purchase their new home, a similar home in the same district is put on the market. The Smith's would have been just as happy in this new home, but due to the quasi-random timing of the supply of homes in the market, the trampoline-less Jones family moves in instead. From the perspective of incumbent neighbors, whether the Smith's or the Jones' become their new neighbors is as good as random.

Analogous to the classic diff-in-diff framework, the identifying assumption is that the neighbors of the Jones' represent the counterfactual change in trampoline ownership that the Smith's neighbors would have experienced. By only comparing incumbent households residing in the same district in the same year, this assumption is likely to hold. Because there is no pre-treatment period in which households do not receive new neighbors, we cannot observe ownership pre-trends of any meaningful length. However, I can tentatively test if a new neighbor provides a plausibly exogenous source of variation in local trampoline composition.

A closely related qualification concerns the definition of a *new* neighbor. If neighborhood effects exist, it would be wrong to assume that the newly moved-in Smiths and Joneses are immune to these effects. As time goes by, their expected trampoline status is likely to converge to that of incumbent neighbors, be it because of shared characteristics, common shocks or an endogenous neighborhood effect. Therefore, establishing a post-move-in window during which new neighbors plausibly constitute an exogenous source of variation is crucial. The type of exogenous variation that I wish to capture is represented by households who contribute to a change in the neighborhood trampoline composition because they have committed to (non-)ownership prior to moving in by e.g. bringing an already purchased trampoline to their new home. If a household moving in is immediately affected by their new neighbors, it is hard to claim that a new neighbor can constitute an exogenous shock to incumbent neighbors.

I test the validity of using new neighbors as a source of exogenous variation by examining if their ownership status is contemporaneously correlated with that of incumbent neighbors after I control for a set pre-determined covariates, presumably observable to the new neighbors prior to moving in. The underlying assumption is that correlation in observed ownership is informative of unobserved determinants of ownership. I estimate the following specification:

$$y_{itd} = \alpha + \beta C_{it}^{M}(\mathbf{y_t}) + \sum_{D \in \mathcal{D}} \gamma_D C_{it}^{M}(\mathbf{y_t}) \times \mathbb{I}(d \in D) + \sum_{D \in \mathcal{D}} \delta_D \mathbb{I}(d \in D) + \theta' \mathbf{X}_{it-1} + \epsilon_{itd} \quad (8)$$

where y_{itd} is the trampoline status of household i who moved in d days before the period t photo was taken. The number of neighboring trampolines, $C_{it}^{M}(\mathbf{y_t})$, is interacted with indicators for different move-in windows D (e.g. 60 to 120 days) collected in the set $\mathcal{D}$. $\mathbf{X}_{it-1}$ is a vector of pre-determined covariates. If the marginal effect of an additional trampoline in period t is small and insignificant, this suggests that the impact of neighboring trampolines on a newly moved in household is limited and that the exogeneity assumption is plausible. In practice, the test is likely too sensitive due to my lack of pre-determined observables at the household and neighborhood level. However, in my main analysis I am only concerned with changes within incumbent households, meaning I am implicitly accounting for a greater number of confounders. The controls included in $\mathbf{X}_{it-1}$ are a full set of district by year and month of sale fixed effects as well as decile indicators for the lagged share of the grid square population below 16 years of age and the mean area of neighboring properties.[21]

The results from estimating equation (8) is presented by figure 7 as the marginal impact of additional neighboring trampolines in period t for a household who moved in

[21] As described previously, age is important determinant of demand. By controlling for pre-determined demographics, I want to capture the latent "child-friendliness" of an area. I conjecture that property area is also important since a trampoline is quite large.

just prior to the photo was taken. It seems that the ownership status of households who move in between 10[22] and 60 days prior to the photo is uncorrelated with the number of trampolines in their new neighborhood. While the choice of the different move-in windows is arbitrary, reshuffling them has very limited impact on the results. The trade-off I ultimately face is that while a shorter window better limits the possibility of reverse causality, it also makes treatment an exceedingly rare event which gives us less variation by which we can identify the effect. To further account for highly localized confounders, I also estimate 8 with the number of trampolines in the prior period as a control. This does not does qualitatively affect the result.

The fact that it takes almost a year for any significant correlation to arise could be due to the pronounced seasonality in trampoline sales (see figure 3. Since public interest in trampolines almost non-existent during October through Mars and since photos are typically taken during spring or early summer, only households with 200-300 days or more spent living in the neighborhood would have experienced a prior trampoline season, which is when they would have had the opportunity to buy one.

My baseline estimates defines a neighborhood as the 5 closest neighbors and uses a move-in window for new neighbors set at between 10-60 days prior to the photo date. Since treatment can be both positive and negative depending on whether the new neighbor contributes to an increase or a decrease in local trampoline ownership, I create indicators by treatment type and include both types of treatment separately. I argue that the potential response could be quite different depending on if you are more or less exposed to the good, and we may have different priors depending on the type of treatment. While this grouping eliminates some information on treatment intensity, receiving any treatment is rare and receiving more treatment than the loss or gain of a single trampoline is even rarer.[23] Thus, grouping by treatment type facilitates interpretation and we can still roughly interpret estimates as the loss or gain of a single neighboring trampoline. Since treatment is influenced by the number of trampolines in $t-2$, I include a full set of indicators for the number of trampolines present in $t-2$ as controls.[24]

[22] To account for potential differences between the registered date of sale and the unobserved actual move-in date, I set the lower bound of the move-in window to 10 days.

[23] Using the 60 day window, 99.6 percent of observations have $D_{it-1}\Delta\mathbf{y}_{t-1} = 0$ and are classified as controls (the reference group). Conditional on any treatment ($|D_{it-1}\Delta\mathbf{y}_{t-1}| > 0$), only 0.2 percent receive treatment greater than $\Delta 1$.

[24] For example, a neighborhood with five out of five trampolines in $t-2$ cannot be treated with an additional trampoline

Table 5: Results, new neighbors (60 days, 5 closest neighbors)

Negative ($D_{it-1}\mathbf{\Delta y}_{t-1} < 0$)	-0.935	-0.292	0.000165
	(0.00622)	(0.00498)	(0.00496)
Positive ($D_{it-1}\mathbf{\Delta y}_{t-1} > 0$)	-0.741	1.420***	1.447***
	(0.677)	(0.539)	(0.537)
Negative $\times$ $\mathbb{I}[y_{t-1} = 1]$		-52.00***	
		(1.857)	
Control $\times$ $\mathbb{I}[y_{t-1} = 1]$		-52.29***	
		(0.111)	
Positive $\times$ $\mathbb{I}[y_{t-1} = 1]$		-54.46***	
		(2.112)	
Marginal effects for $y_{t-1} = 1$			
Negative vs Control		0.00521	
		(1.794)	
Positive vs Control		-0.754	
		(2.046)	
Observations	1,766,022	1,766,022	1,533,450
Households	872,152	872,152	800,499
R-squared	0.008	0.262	0.019
Outcome mean	13.49%	13.49%	7.06%
Fixed effect	District $\times$ Year	District $\times$ Year	District $\times$ Year

The table presents estimates of equation (2). Coefficients expressed as percentage points. The outcome is the change in trampoline ownership status, Δy_t. Treatment, $D_{it-1}\mathbf{\Delta y}_{t-1}$, is categorized as positive or negative. The second column includes interactions between initial ownership and treatment. The third column only includes initial non-owners. All regressions include a full set of indicators for the number of neighboring trampolines in $t-2$. Standard errors clustered on households. *** p<0.01, ** p<0.05, * p<0.1

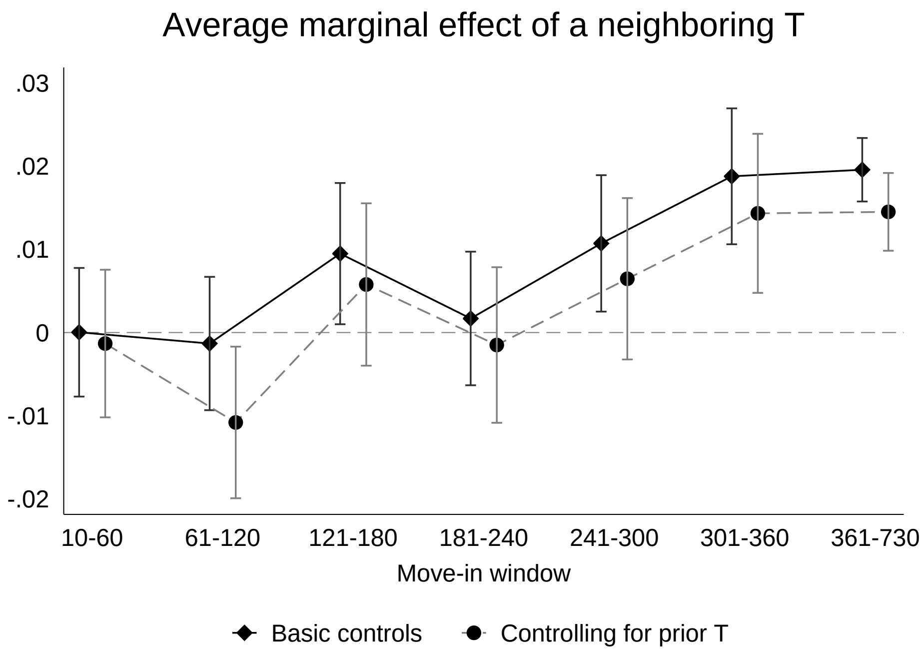

Figure 7: The figure presents marginal effects from estimating equation (8) using a range of move-in windows prior to the period t photo was taken. Controls include a full set of district by year and month of sale fixed effects as well as decile indicators for the lagged share of the grid square population below 16 years of age and the mean area of neighboring properties. The green line includes an additional control, $W_i\mathbf{y}_{t-1}$. Bars indicate 95 percent CI's.

Table 5 presents the results from my estimation of equation (2). In the first column, I only include the treatment indicators and I find no significant treatment effect. In fact, both estimates are negative. However, it is reasonable to assume that treatment affects prior owners and non-owners differently. As I only explore the extensive margin of ownership, most of the underlying mechanisms suggested by the literature only apply to non-owners. A trampoline owner is much less likely to respond to a new neighboring trampoline as the information and status-seeking mechanisms are all but rendered irrelevant as they are only deciding whether to keep the trampoline for another period or not. While it could be that neighbors can signal whether or not trampolines are still in style, such effects are likely limited. Trampoline owners are in minority in terms of observations, but they contribute dis-proportionally to the variation in the outcome since many households among the initial non-owners are likely "never-takers", i.e. they would never purchase a trampoline regardless of influence from peers.

The second column in table 5 includes a set of interactions to estimate separate treatment effects for household who are owners and non-owners in $t-1$, respectively. When I estimate separate effects for owners and non-owners, I find a positive effect of additional neighboring trampolines among non-owners. Non-owners do not respond to negative treatment, i.e. they are not less likely to purchase when there are fewer neighboring trampolines. Among prior owners, I find no significant effects when comparing marginal effects in either treatment group to the control. The impact of a neighboring trampoline on prior non-owners is also established in the third column where I only include initial non-owners in the estimation sample. Among these, mean ownership rates across all years is only 7 percent, meaning that the 1.45 percentage point increase is quantitatively quite large (20 percent of the mean). As a sensitivity check, I extend the move-in window for new neighbors to 120 days (see table A.1). Overall, results are similar but smaller in magnitude. As we increase the window, new neighbors and incumbents have more time to react to each other prior to the photo date, meaning we are more likely to underestimate the effects.

To examine the geographic scope of neighborhood effects, I reproduce columns 2 and 3 of table 5 using both the 10 and 15 closest neighbors as my neighborhood definition. As shown by table 6, effects largely disappear. The impact of positive treatment is reduced by more than half when using the 10 closest and is even smaller for the 15 closest. While I find a small and marginally significant effect of negative treatment among non-owners, this effect is not corroborated by column 2 and 4 where I only include non-owners. The seemingly narrow scope for the neighborhood effect on consumption is consistent with previous studies (Grinblatt et al., 2008; Kuhn et al., 2011; Agarwal et al., 2017b) as well as my baseline household fixed effect estimates.

Which of the mechanisms suggested by the literature can best explain the positive effect on prior non-owners? To discern the underlying mechanisms, I examine how the effect interacts with the prior trampoline composition in the neighborhood. I hypothesize that if information transmission between new neighbors is important, we should see larger effects in neighborhoods who initially have no trampolines as these households are more likely to be exposed to the product for the first time. Through the new owner, incumbent neighbors may learn about product use and characteristics, prices at local stores etc. Conversely, if effects are larger in neighborhoods where trampolines are already present, this information channel is likely less relevant and households acting to 'keep up' with a reference group is a more plausible explanation.

Table 7 presents three different specifications aimed at disentangling the effect. All

Table 6: Results, expanded neighborhoods

	10 closest		15 closest	
Negative ($D_{it-1}\mathbf{\Delta y}_{t-1} < 0$)	-0.610*	-0.362	-0.555*	-0.298
	(0.357)	(0.356)	(0.302)	(0.301)
Positive ($D_{it-1}\mathbf{\Delta y}_{t-1} > 0$)	0.561	0.626	0.394	0.440
	(0.384)	(0.382)	(0.329)	(0.326)
Negative × $\mathbb{I}[y_{t-1} = 1]$	-52.84***		-52.17***	
	(1.406)		(1.214)	
Control × $\mathbb{I}[y_{t-1} = 1]$	-52.34***		-52.33***	
	(0.116)		(0.120)	
Positive × $\mathbb{I}[y_{t-1} = 1]$	-53.06***		-50.21***	
	(1.591)		(1.345)	
Marginal effects for $y_{t-1} = 1$				
Negative vs Control	-1.110		-0.395	
	(1.369)		(1.186)	
Positive vs Control	-0.167		2.514*	
	(1.550)		(1.314)	
Observations	1,631,966	1,417,256	1,525,698	1,325,580
Households	812,094	744,673	764,695	700,681
R-squared	0.262	0.019	0.262	0.020
Outcome mean	13.47%	7.05%	13.44%	7.03%

The table presents estimates of equation (2). Coefficients expressed as percentage points. The outcome is the change in trampoline ownership status, Δy_t. Treatment, $D_{it-1}\mathbf{\Delta y}_{t-1}$, is categorized as positive or negative and is based on a move-in window of 60 days. Columns 1 and 3 includes interactions between initial ownership and treatment. Columns 2 and 4 only includes initial non-owners. All regressions include a full set of indicators for the number of neighboring trampolines in $t-2$. Standard errors clustered on households. *** p<0.01, ** p<0.05, * p<0.1

three samples only include prior non-owners. The first column treats the number of trampolines present in $t-2$ (C^5_{it-2}) as a continuous variable. The effect of receiving additional trampoline should be interpreted as the effect among households with $C^5_{it-2}=0$ and is larger than the estimate without the interaction term (column 3 of table 5). The negative sign on the interaction suggests a reduced treatment effect as C^5_{it-2} increases, although the estimate is insignificant. The second column treats C^5_{it-2} as categorical and includes a full set of indicators for each number of prior trampolines as well as treatment interactions. For clarity, table 7 only presents the marginal effect of positive treatment at different numbers of initial trampolines.[25]

As shown in column 2, there are no positive effects of additional trampolines in neighborhoods with one or more trampolines already present, although estimates are noisy. Among households with $C^5_{it-2}=0$, the effect is 60 percent larger than the unconditional estimate in column 2 of table 5. I find a significant negative effect at 4 initial trampolines. It is reasonable that adding a fifth neighboring trampoline does not increase the probability of purchase further, but a negative effect also speaks to the idea of neighboring trampolines acting as a substitute to an owned trampoline. With even more substitutes available, the probability of buying one seems to decrease. It should be noted that only a few thousand households reside in neighborhoods where 4 out 5 neighbors have trampolines. Coupled with the low probability of treatment, any interpretation of this estimate warrants caution.

For negative treatment, the reference category is $C^5_{it-2}=4$ since the loss of a neighboring trampoline is not possible for $C^5_{it-2}=0$. The positive point estimate in column 2 could also be explained by neighboring trampolines acting as substitutes, with the loss of a neighbor's forcing a purchase. The third column replicates the second by simply limiting the sample to non-owners living in neighborhood with zero trampolines in $t-2$.

To summarize the findings, there is evidence of neighborhood effects. Perhaps unsurprisingly, the only robust finding is a positive effect among prior non-owners experiencing the arrival of the first neighboring trampoline. This finding could suggest that information transmission between neighbors is an important mechanism. Although this does not rule out status-seeking effects, the fact that there are no positive effect present past the first trampoline makes a 'keeping up'-explanation less appealing. In fact, there is some

[25] There are no significant effects of negative treatment. Positive treatment is not feasible at $C^5_{it-2}=5$ since this 5 is the maximum number of trampolines possible among the 5 closest neighbors.

suggestive evidence that neighboring trampolines can act as substitutes for an owned trampoline. Effects are largely limited to within the five closest neighbors, which is consistent with prior work.

4.3 Is the effect economically significant?

To put the size of the neighborhood effect into context, we may ask how much of aggregate trampoline demand that can be explained by the propagation of a positive neighborhood effect. My data suggests that trampolines in Sweden became increasingly popular between 2006 and 2018. How much of this increase can be explained by neighborhood spillovers in ownership? By comparing a simple simulation in which the only driver of aggregate demand is the neighborhood effect to the actual trend, we can put the effect size into perspective. Let us abstract from other covariates and assume that the only determinant of y_t is the interaction effect and the prior ownership status according to:

$$y_i t = y_{it-1} + \beta W_i \mathbf{y}_{t-1} + \varepsilon_t \tag{9}$$

For this simple numerical simulation, I assume that the trampoline status of household i follows a steady state Markov chain:

$$P_i = \begin{bmatrix} 0.48 & 0.52 \\ 0.05 + \delta_i & 0.95 - \delta_i \end{bmatrix}$$

where $P_{1,1}$ is the probability of keeping your trampoline for another period. The term *delta* is the added probability of purchasing a trampoline that a non-owner faces as a result of the positive neighborhood effect. I calculate δ_i as the sum of all neighboring trampolines times the neighborhood effect β. For $\delta = 0$, the Markov chain is a steady state process that maintains the initial ownership rate which I set at 9.4 percent, the mean rate I observe in 2006. The simulation is performed as follows:

Step 0: 9.4 percent of properties are randomly allocated a trampoline.

Step 1: Non-owners have their δ_i calculated.

Step 2: Ownership status is updated for the next period according to P_i

Table 7: Results, by initial no. of trampolines (60 days, 5 closest neighbors)

Negative ($D_{it-1}\mathbf{\Delta y}_{t-1} < 0$)	-0.892	4.689	
	(1.220)	(6.146)	
Positive ($D_{it-1}\mathbf{\Delta y}_{t-1} > 0$)	1.974***	2.352***	2.347***
	(0.672)	(0.705)	(0.708)
Negative C^5_{it-2}	0.566		
	(0.738)		
Positive $\times$ C^5_{it-2}	-0.967		
	(0.809)		
C^5_{it-2}	1.119***		
	(0.0320)		
$\mathbb{I}[C^5_{it-2} = 1]$		1.362***	
		(0.0504)	
$\mathbb{I}[C^5_{it-2} = 2]$		2.058***	
		(0.0848)	
$\mathbb{I}[C^5_{it-2} = 3]$		3.137***	
		(0.188)	
$\mathbb{I}[C^5_{it-2} = 4]$		4.697***	
		(0.543)	
$\mathbb{I}[C^5_{it-2} = 5]$		6.418***	
		(2.075)	
Marginal effects of positive treatment by initial trampolines			
$C^5_{it-2} = 0$		2.352***	
		(0.705)	
$C^5_{it-2} = 1$		-0.439	
		(0.883)	
$C^5_{it-2} = 2$		3.166	
		(2.190)	
$C^5_{it-2} = 3$		-2.571	
		(4.805)	
$C^5_{it-2} = 4$		-10.52***	
		(1.596)	
Observations	1,533,450	1,533,450	919,588
Households	800,499	800,499	569,666
R-squared	0.019	0.019	0.022
Outcome mean	7.06%	7.06%	6.03%

The table presents the estimation of equation (2) on a sample of initial non-owners ($y_{it-1} = 0$). Coefficients expressed as percentage points. The first column treats C^5_{it-2} as continuous. The second column includes a full set of interactions between level indicators for C^5_{it-2} and treatment. The marginal effect of positive treatment over C^5_{it-2} is reported in the bottom panel. The third column only includes neighborhoods with $C^5_{it-2} = 0$ All regressions include a full set of district by year fixed effects. Standard errors clustered on households. *** p<0.01, ** p<0.05, * p<0.1

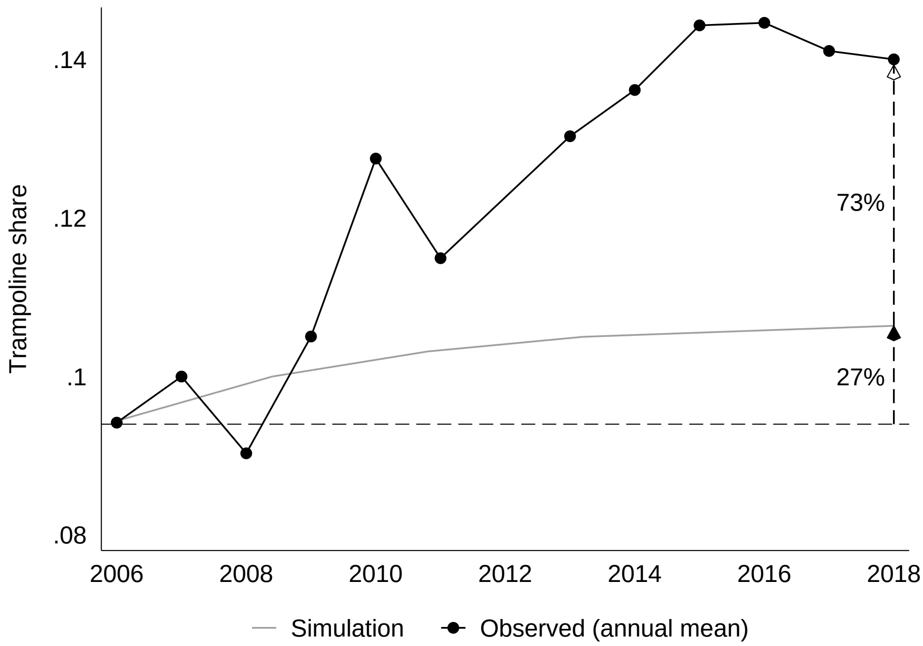

Figure 8: Simulated aggregate impact of the neighborhood effect

Step 3: Iterate step 1 and 2 for the desired number of periods.

For the simulation, I set β to 0.0142 (corresponding to the estimate in column 2 of table 5). The average time between photos is 2.5 years. I run the simulation for 5 periods, which given a period length of 2.5 years is about equivalent to the 12 years between 2006 and 2018. As figure 8 shows, the neighborhood effect alone can explain about 25 percent of the aggregated rise in trampoline ownership between 2006 and 2018. This simple exercise suggests that the neighborhood effect is a quantitatively important determinant of demand for this particular good.

5 Conclusion

This study provides evidence of a neighborhood effect in the consumption of a popular household durable good. I address the major concerns related to the confounders of the true outcome-on-outcome neighborhood effect using two different approaches. First, I estimate a diff-in-diff specification similar to Grinblatt et al. (2008). Second, I use plausibly exogenous variation stemming from new households moving into the neighborhood which affects the exposure of incumbent neighbors. My preferred specification indicates that the neighborhood effect can have a substantial impact of demand for this particular product. The fact that it is only the first neighboring trampoline that seem to matter suggests that information transmission between neighbors is an important mechanism.

However, it is not clear what information is being conveyed. As this is a relatively simple, homogeneous and inexpensive good, it could be that information about utility is more important than technical characteristics, local pricing etc. Additionally, since this a product where children likely have a lot of influence regarding the purchase decision, information about utility rather than price could be the most important message transmitted between neighbors.

An additional contribution of this study is to provide an example of how deep learning algorithms can be used by economists to collect data from image-based sources with a level of accuracy comparable to manual methods but at a cost that is orders of magnitude lower. I believe that these recent advances in computer vision will open up new mountains of data to researchers in social science.

References

Agarwal, S., Jensen, J. B., and Monte, F. (2017a). The Geography of Consumption. SSRN Scholarly Paper ID 3007488, Social Science Research Network, Rochester, NY.

Agarwal, S., Qian, W., and Zou, X. (2017b). Thy Neighbor's Misfortune: Peer Effect on Consumption. SSRN Scholarly Paper ID 2780764, Social Science Research Network, Rochester, NY.

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, 30:98–108.

Anselin, L. (2009). Spatial Hedonic Models. In Mills, T. C. and Patterson, K., editors, *Palgrave Handbook of Econometrics*, pages 1243–1250. Palgrave Macmillan UK, London.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, 58(2):277.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, page 29.

Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150(1):41–55.

Buda, M., Maki, A., and Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, 106:249 – 259.

Dahl, G. B., Løken, K. V., and Mogstad, M. (2014). Peer Effects in Program Participation. *American Economic Review*, 104(7):2049–2074.

De Giorgi, G., Frederiksen, A., and Pistaferri, L. (2016). Consumption Network Effects. Working Paper 22357, National Bureau of Economic Research.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*.

Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L., and Fei-Fei, L. (2017). Using Deep Learning and Google Street View to Estimate the Demographic Makeup of the US. *arXiv:1702.06683 [cs]*. arXiv: 1702.06683.

Gilchrist, D. S. and Sands, E. G. (2016). Something to Talk About: Social Spillovers in Movie Consumption. *Journal of Political Economy*, 124(5):1339–1382.

Grinblatt, M., Keloharju, M., and Ikäheimo, S. (2008). Social Influence and Consumption: Evidence from the Automobile Purchases of Neighbors. *The Review of Economics and Statistics*, 90(4):735–753.

Haenssle, H. A., Fink, C., Schneiderbauer, R., and [et al.] (2018). Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Annals of Oncology*, 29(8):1836–1842.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring Economic Growth from Outer Space. *American Economic Review*, 102(2):994–1028.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Pereira, F., Burges, C. J. C., Bottou, L., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc.

Kuhn, P., Kooreman, P., Soetevent, A., and Kapteyn, A. (2011). The Effects of Lottery Prizes on Winners and Their Neighbors: Evidence from the Dutch Postcode Lottery. *American Economic Review*, 101(5):2226–2247.

Lavy, V., Silva, O., and Weinhardt, F. (2012). The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools. *Journal of Labor Economics*, 30(2):367–414.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60(3):531–542.

Moretti, E. (2011). Social Learning and Peer Effects in Consumption: Evidence from Movie Sales. *Review of Economic Studies*, 78(1):356–393.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49(6):1417–26.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, 9(1):86–136.

Szegedy, C., Ioffe, S., and Vanhoucke, V. (2016). Inception-v4, inception-resnet and the impact of residual connections on learning. *CoRR*, abs/1602.07261.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288.

Vasconcelos, C. N. and Vasconcelos, B. N. (2017). Increasing deep learning melanoma classification by classical and expert knowledge based image transforms. *CoRR*, abs/1702.07025.

A Appendix

A.1 Image prepossessing

In this section, I detail the data collection process. To obtain the trampoline ownership status of the homes in my sample using a neural network, a number of preprocessing steps is necessary. The aerial photographs used are produced by the Swedish mapping, cadastral and land registration authority (Lantmäteriet). The set of images I use covers the major population centers and the surrounding areas (tables A.2a and A.2b present the distribution of land lot-year observations at the municipal level). The photographs are taken during 2006 until 2017. The observation weighted mean time between consecutive photos of a land lot is 2.5 years. A land lot in my sample is on average photographed 4.2 times. The images are orthorectified, meaning that the photo has been projected on an elevation model of the earth's surface and a standard mapping coordinate system is embedded in the image. This allows maps and other spatial data to be accurately projected on the image. The images are typically taken during spring, just after trees and bushes start to leaf out. The median photo date for 2010 until 2017 is May 19$^{\text{th}}$.

To detect the presence of a trampoline on a specific lot, the first prepossessing step is to cut the large [26] aerial photos down to small image chips, each containing a single land lot. As an input, the neural net accepts a fixed three-dimensional matrix. In this study, I set the shape of the input as (300,300,3), which corresponds to a 300 by 300 pixel using the standard 24-bit three-channel RGB color space[27]. The aerial photos have a spatial resolution of 0.25 meters per image pixel, meaning that the largest eligible dimension for a land lot is 75 by 75 meters $(300 * 0.25)$. This is sufficiently large to fit the majority of single-family home lots and only about 2 percent of eligible lots have to be discarded for being too large. Increasing the input size increases the computational burden exponentially and scaling down images of larger lots to fit the input size would decrease the pixel size of trampolines, making them more difficult to detect.

It is crucial that each image only shows a single land lot, or we risk contaminating the input with e.g. neighboring trampolines. To achieve this, I use a vector based map of land lot borders produced and maintained by Lantmäteriet. Projecting this map on

[26] The original images are 10,000 $\times$ 10,000 pixels.

[27] The third dimension corresponds to the red, green and blue color space. Each 300 by 300 channel takes an integer value between 0 and 255 (8 bits).

the aerial photo allows me to cut along the borders to isolate a single property. As I am only interested in properties occupied by single-family homes, I overlay a map of building footprints and their legal building code and only include land lots that intersect a build registered as a single-family home. While land lots with multiple family homes are rare, I have to exclude these since I would not be able to attribute a trampoline to a single household. Figure A.1 illustrates the property selection procedure. Since the image is always cut along the x-y axis of the original photo (roughly corresponding to the east-west and north-south axis of the map), the cardinal alignment of the land lot matters for the dimensions of the cutout. To avoid selecting my sample on this rather arbitrary factor, I implement a simple rotation algorithm. Before discarding a land lot as being too large, the algorithm attempts to rotate lots that exceed the 300 pixel restriction in either dimension up to 90 degrees. Figure A.3 illustrates the rotation procedure.

Because the map detailing property borders is continuously updated and only the latest version is made available, there are discrepancies between the ground truth at the time the photo was taken and the border map. For example, a property observed in 2017 may not have been formed and built upon back in 2006. Consequently, the lot cutout from the 2006 photograph would just depict an empty plot of land. Since a lot has be occupied in order to meaningfully be classified into having or not having a trampoline, this requires me to first classify lots as occupied or empty. I address this ex ante classification problem using a secondary neural net, trained on images containing occupied and empty lots. I manually label a set of 9,618 images (3,084 empty and 6,534 occupied lots). The set of labeled images is balanced and split 85/15 into a training and validation data set. After training, the neural network achieves 97.6 percent accuracy in the validation data set. The trained network is then used to predict the full set of images. If the predicted probability of being occupied is 0.5 or above, I consider the land lot to be occupied in the current and all future years. Because new housing is marginal compared to the existing stock, the vast majority of lot-year observations (98 percent) are classified as occupied. As changes to this secondary neural network only concerns sample selection along a very small margin, I do not detail the performance of this network further.

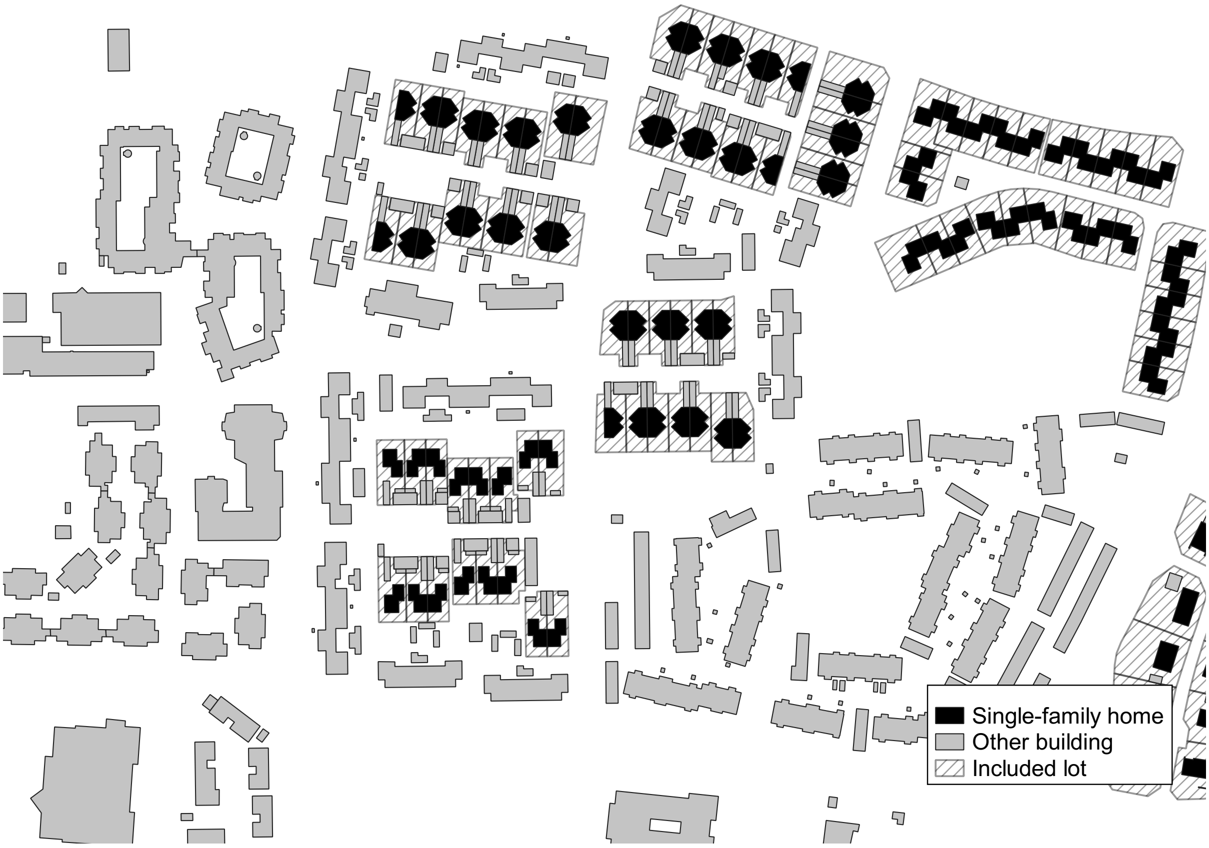

Figure A.1: Black shapes represent the footprint of single family homes. Their corresponding land lot is included in my sample (striped areas).

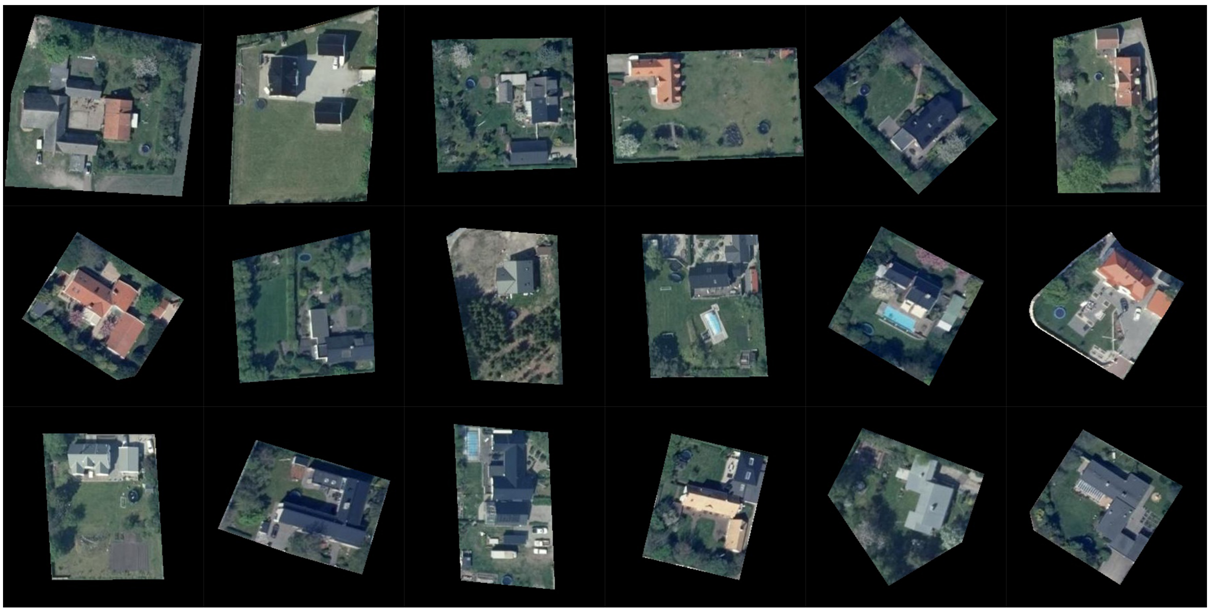

Figure A.2: Examples of images fed into the neural network. Each property is a separate 300 by 300 pixel image, corresponding to 75*75 meters.

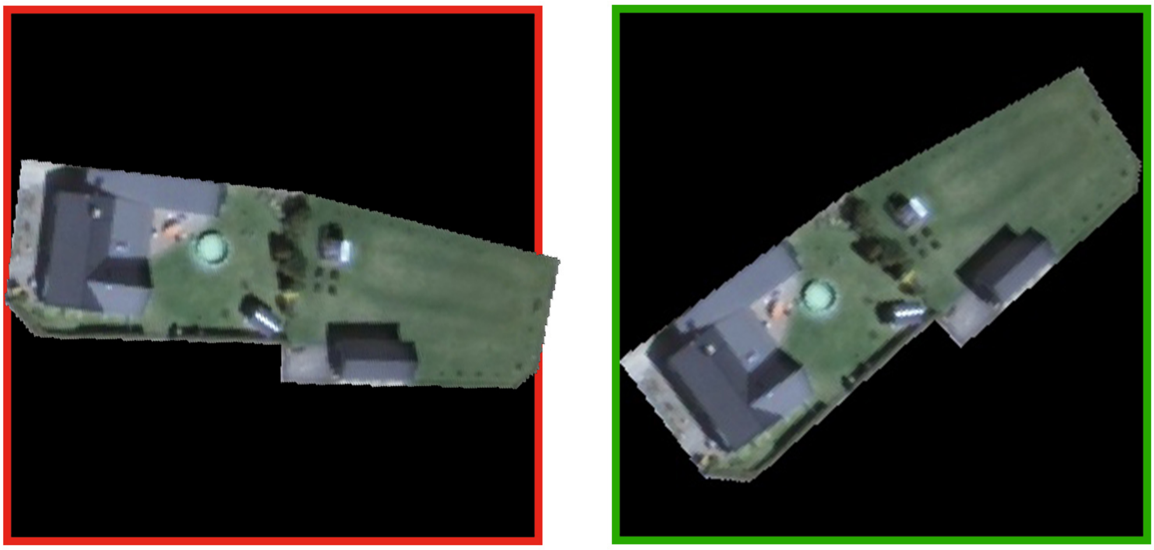

Figure A.3: An example of an image violating the 300 pixel restriction due to the x-y alignment of the land lot (left). After a 43 degree counter-clockwise rotation, the lot satisfies the restriction (right).

A.2 Training the neural network

The neural network used in this study is called Inception-Resnet (Szegedy et al., 2016).[28] This publicly available neural network has demonstrated state-of-the-art performance on the benchmark ImageNet (Deng et al., 2009) dataset. In addition to excellent performance, it is less memory intensive than competing networks, making it feasible for consumer grade hardware.[29]

After preprocessing the aerial photos (see section A.1), I am left with a set of about 4,484,665 images, each containing a single property lot with a single-family home. To train the network to recognize a trampoline, I create a training data set by manually labeling a subsample of 22,435 images as being a member of either of the two classes, "Trampoline" (2,375 images) or "No trampoline" (20,060 images). Images of trampolines are underrepresented due to the simple fact that trampolines are fairly rare. The frequency imbalance between the two classes poses a problem for training because the iterative updating of the network weights can easily take the loss function to a local minimum where every image is predicted to be of the majority class without actually learning

[28] As a side note, I initially tried to train an older network called VGG16, the 2014 winner of ILSVRC, with poor results. I was only able to achieve high accuracy by overfitting the training data, suggesting that the network was not "deep" enough to capture robust representations of the images.

[29] I perform all training and classification on a standard desktop computer equipped with an Nvidia Geforce GTX 1080ti GPU using the open source Keras Python library with a Tensorflow backend.

anything about the features of the images. To induce learning, I manually oversample the "Trampoline" class in the training and validation data to achieve a 50-50 balance of images containing trampolines and images without (Buda et al., 2018).

In the initial stages of training, I manually inspected all erroneous predictions made by the model in order to identify potential weaknesses. I initially found that for type 1 errors. i.e. finding a mat in an image labeled "No trampoline", there was often, upon review, actually a trampoline present in the image. I had simply missed it during labeling due to e.g. partial occlusion, low contrast, placement or, most importantly, human error. While this study is not a formal comparison between manual and automated image classification, this provides some anecdotal evidence that the accuracy of manual classification in this seemingly trivial task is being challenged by neural networks.

While it is reassuring that the model can identify even elusive trampolines, it is precarious to let the model influence the labeling of training data. The risk is that any flaws in the model is reinforced during training and that the "ground truth" that should be represented by the label is compromised. This makes it impossible to assess true model performance. Rather than simply trusting the model when identifying an incorrectly labeled image, I set up a script that fetches an image of the same property using the Google Maps Static API.[30] This allows me to get two independently sourced images of the same property, which is helpful in borderline-cases. Additionally, the Google Maps image is often of higher resolution and taken from a different angel. In ambiguous cases where the model and I disagree on the correct label, I examine the corresponding Google Maps image and use it as a tie-breaker. Figure A.4 present a few examples of the visual differences between the two image sources.

With images labeled as correctly as possible, I then randomize the data into an 80-10-10 training-validation-testing split. The validation set is used for model selection while the testing set is used to evaluate final model performance. During training, I use the Adam optimization algorithm and an initial learning rate of 0.001. When validation accuracy stops improving, the learning rate is reduced by half and so on until the accuracy converges. I set the maximum number of epochs to 30 with each epoch consisting of a complete pass of all training images, but the model typically converges faster. To get the most mileage out of my relatively scarce trampoline images, I augment the training

[30] Using only Google Maps data for this study was not possible at the time of writing for several reasons.

Figure A.4: Examples of properties where the trampoline is difficult to make out in the original image (columns 1 and 3). In the high resolution Google-sourced image of the same property (to the right of the original), it is easy to verify the presence of a trampoline.

and validation data by applying random shifts and flips around each axis and random 90-degree rotations. This has been shown to increase model robustness (Vasconcelos and Vasconcelos, 2017).

A.3 Additional tables and graphs

Table A.1: Results, new neighbors (120 days, 5 closest neighbors)

Negative ($D_{it-1}\mathbf{\Delta y}_{t-1} < 0$)	-0.753*	-0.622*	-0.346
	(0.440)	(0.353)	(0.352)
Positive ($D_{it-1}\mathbf{\Delta y}_{t-1} > 0$)	-0.734*	0.920***	0.948***
	(0.432)	(0.345)	(0.343)
Negative × $\mathbb{I}[y_{t-1} = 1]$		-50.48***	
		(1.336)	
Control × $\mathbb{I}[y_{t-1} = 1]$		-52.30***	
		(0.111)	
Positive × $\mathbb{I}[y_{t-1} = 1]$		-52.91***	
		(1.372)	
Marginal effects for $y_{t-1} = 1$			
Negative vs Control		1.204	
		(1.295)	
Positive vs Control		0.311	
		(1.334)	
Observations	1,766,022	1,766,022	1,533,450
Households	872,152	872,152	800,499
R-squared	0.008	0.262	0.019
Outcome mean	13.49%	13.49%	7.06%
Fixed effect	District × Year	District × Year	District × Year

The table presents estimates of equation (2). Coefficients expressed as percentage points. The outcome is the change in trampoline ownership status, Δy_t. Treatment, $D_{it-1}\mathbf{\Delta y}_{t-1}$, is categorized as positive or negative. The second column includes interactions between initial ownership and treatment. The third column only includes initial non-owners. All regressions include a full set of indicators for the number of neighboring trampolines in $t-2$. Standard errors clustered on households. *** p<0.01, ** p<0.05, * p<0.1

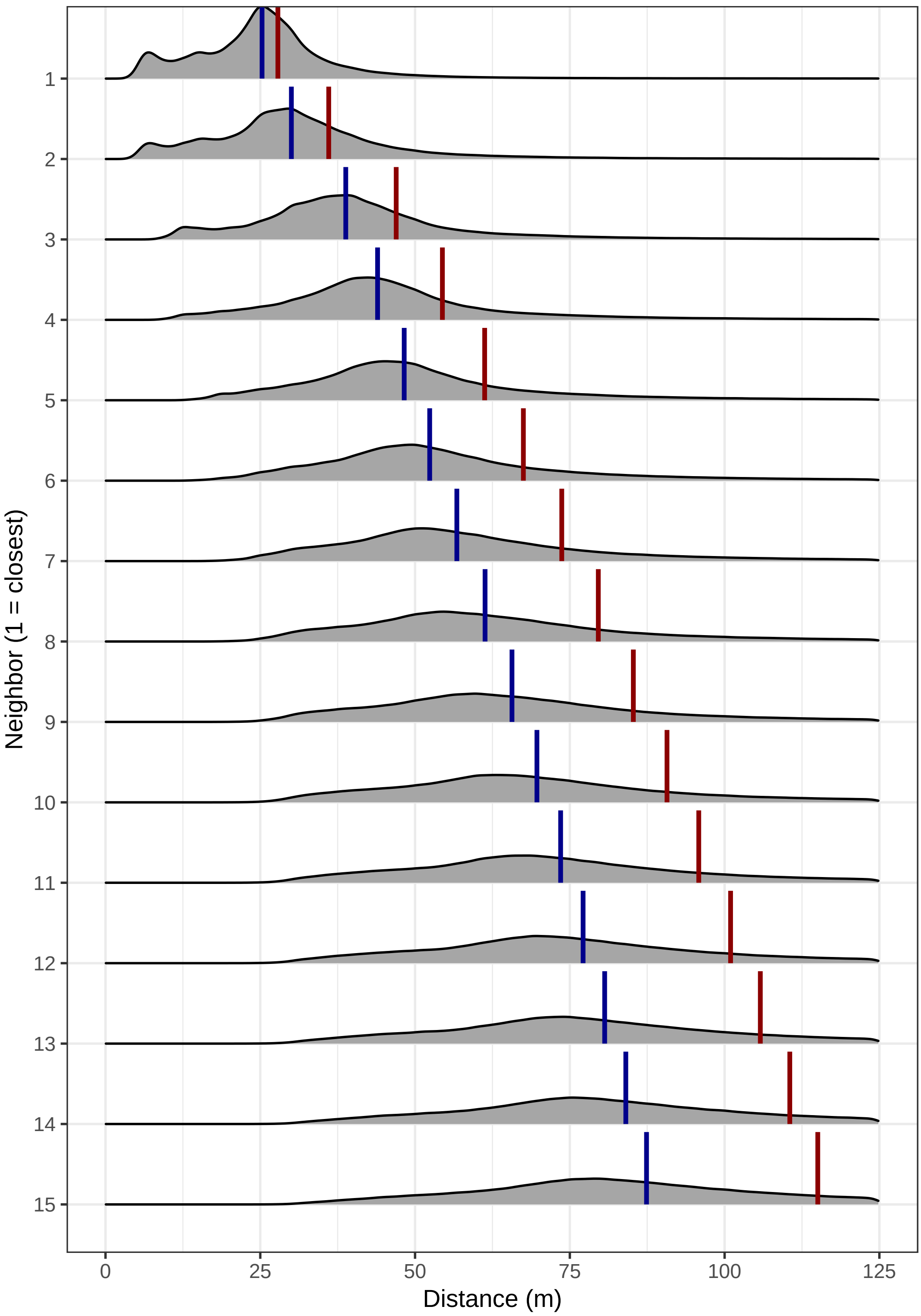

Figure A.5: Density plots of the sample centroid-to-centroid distance to the 15 closest neighbors. The red lines indicate means, blue lines are medians.

Table A.2a: Observations by municipality

Municipality	No.	%	Municipality	No.	%	Municipality	No.	%
Ale	17,871	0.40%	Gnosjö	1,817	0.04%	Kristianstad	66,615	1.49%
Alingsås	16,051	0.36%	Grums	3,795	0.08%	Kristinehamn	14,324	0.32%
Alvesta	11,990	0.27%	Grästorp	1,661	0.04%	Kumla	12,611	0.28%
Aneby	1,505	0.03%	Gullspång	2,466	0.05%	Kungsbacka	72,823	1.62%
Arboga	9,186	0.20%	Gävle	42,352	0.94%	Kungsör	3,176	0.07%
Arvidsjaur	1,153	0.03%	Göteborg	190,442	4.25%	Kungälv	32,201	0.72%
Arvika	4,565	0.10%	Götene	5,987	0.13%	Kävlinge	19,774	0.44%
Askersund	2,998	0.07%	Habo	3,942	0.09%	Köping	10,772	0.24%
Avesta	4,708	0.10%	Hallsberg	4,447	0.10%	Laholm	19,565	0.44%
Bengtsfors	5,870	0.13%	Hallstahammar	8,481	0.19%	Landskrona	22,212	0.50%
Bjurholm	798	0.02%	Halmstad	74,309	1.66%	Laxå	2,148	0.05%
Bjuv	10,791	0.24%	Hammarö	15,844	0.35%	Lekeberg	1,037	0.02%
Boden	12,612	0.28%	Haninge	47,697	1.06%	Leksand	4,828	0.11%
Bollebygd	2,943	0.07%	Heby	2,298	0.05%	Lerum	34,980	0.78%
Bollnäs	10,242	0.23%	Hedemora	2	0.00%	Lessebo	5,581	0.12%
Borgholm	5,997	0.13%	Helsingborg	76,243	1.70%	Lidingö	27,875	0.62%
Borlänge	21,137	0.47%	Herrljunga	3,270	0.07%	Lidköping	24,483	0.55%
Borås	44,027	0.98%	Hjo	4,594	0.10%	Lilla Edet	6,280	0.14%
Botkyrka	40,778	0.91%	Hofors	4,060	0.09%	Lindesberg	9,145	0.20%
Boxholm	1,026	0.02%	Huddinge	68,114	1.52%	Linköping	58,787	1.31%
Bromölla	11,386	0.25%	Hudiksvall	13,232	0.30%	Ljungby	15,271	0.34%
Burlöv	16,067	0.36%	Hultsfred	5,170	0.12%	Ljusnarsberg	766	0.02%
Båstad	23,302	0.52%	Hylte	3,384	0.08%	Lomma	23,449	0.52%
Dals-Ed	1,437	0.03%	Härryda	28,571	0.64%	Ludvika	2,814	0.06%
Danderyd	26,004	0.58%	Hässleholm	40,698	0.91%	Luleå	36,127	0.81%
Degerfors	6,237	0.14%	Håbo	17,045	0.38%	Lund	46,062	1.03%
Ekerö	20,288	0.45%	Höganäs	38,789	0.86%	Lysekil	12,875	0.29%
Eksjö	7,678	0.17%	Högsby	1,806	0.04%	Malmö	135,966	3.03%
Emmaboda	3,119	0.07%	Hörby	8,399	0.19%	Mariestad	14,998	0.33%
Enköping	19,957	0.45%	Höör	13,646	0.30%	Mark	14,429	0.32%
Eskilstuna	43,642	0.97%	Järfälla	38,219	0.85%	Markaryd	6,996	0.16%
Eslöv	20,208	0.45%	Jönköping	72,893	1.63%	Mellerud	4,514	0.10%
Essunga	1,302	0.03%	Kalix	510	0.01%	Mjölby	10,212	0.23%
Fagersta	10,378	0.23%	Kalmar	39,219	0.87%	Mora	8,286	0.18%
Falkenberg	36,947	0.82%	Karlsborg	3,011	0.07%	Motala	24,447	0.55%
Falköping	11,518	0.26%	Karlshamn	28,541	0.64%	Mullsjö	3,027	0.07%
Falun	35,630	0.79%	Karlskoga	21,637	0.48%	Munkedal	2,825	0.06%
Finspång	7,373	0.16%	Karlskrona	37,678	0.84%	Mölndal	39,820	0.89%
Flen	6,004	0.13%	Karlstad	46,293	1.03%	Mönsterås	8,937	0.20%
Forshaga	1,724	0.04%	Katrineholm	13,340	0.30%	Mörbylånga	7,690	0.17%
Färgelanda	2,077	0.05%	Kil	3,787	0.08%	Nacka	56,078	1.25%
Gagnef	4,512	0.10%	Kinda	2,615	0.06%	Nora	3,852	0.09%
Gislaved	15,924	0.36%	Klippan	11,452	0.26%	Norberg	2,560	0.06%
Gnesta	2,834	0.06%	Knivsta	5,934	0.13%	Nordanstig	693	0.02%

Table A.2b: Observations by municipality, continued

Municipality	No.	%	Municipality	No.	%	Municipality	No.	%
Nordmaling	2,271	0.05%	Sunne	2,178	0.05%	Vellinge	47,563	1.06%
Norrköping	50,315	1.12%	Surahammar	5,227	0.12%	Vetlanda	13,425	0.30%
Norrtälje	34,726	0.77%	Svalöv	5,191	0.12%	Vimmerby	3,407	0.08%
Nybro	13,228	0.29%	Svedala	12,292	0.27%	Vingåker	2,105	0.05%
Nykvarn	4,949	0.11%	Svenljunga	1,942	0.04%	Vänersborg	23,743	0.53%
Nyköping	17,596	0.39%	Säffle	7,622	0.17%	Vännäs	2,847	0.06%
Nynäshamn	8,290	0.18%	Säter	31	0.00%	Värmdö	30,109	0.67%
Nässjö	15,077	0.34%	Sävsjö	3,870	0.09%	Värnamo	19,046	0.42%
Ockelbo	1,662	0.04%	Söderhamn	9,801	0.22%	Västervik	21,204	0.47%
Olofström	9,087	0.20%	Söderköping	3,453	0.08%	Västerås	68,631	1.53%
Orsa	1,131	0.03%	Södertälje	31,409	0.70%	Växjö	50,368	1.12%
Orust	10,593	0.24%	Sölvesborg	17,532	0.39%	Vårgårda	2,813	0.06%
Osby	8,195	0.18%	Tanum	13,331	0.30%	Ydre	396	0.01%
Oskarshamn	16,015	0.36%	Tibro	6,387	0.14%	Ystad	23,677	0.53%
Oxelösund	8,752	0.20%	Tidaholm	3,623	0.08%	Älmhult	9,273	0.21%
Partille	26,052	0.58%	Tierp	826	0.02%	Älvkarleby	4,150	0.09%
Perstorp	3,361	0.07%	Timrå	9,298	0.21%	Älvsbyn	4,865	0.11%
Piteå	28,116	0.63%	Tingsryd	4,581	0.10%	Ängelholm	31,605	0.70%
Robertsfors	1,661	0.04%	Tjörn	19,387	0.43%	Åmål	8,221	0.18%
Ronneby	27,366	0.61%	Tomelilla	6,705	0.15%	Åstorp	12,298	0.27%
Rättvik	2,545	0.06%	Torsås	3,918	0.09%	Åtvidaberg	3,477	0.08%
Sala	8,672	0.19%	Tranemo	3,584	0.08%	Öckerö	19,125	0.43%
Salem	11,706	0.26%	Tranås	9,803	0.22%	Ödeshög	1,433	0.03%
Sandviken	24,869	0.55%	Trelleborg	36,211	0.81%	Örebro	45,901	1.02%
Sigtuna	17,310	0.39%	Trollhättan	36,391	0.81%	Örkelljunga	6,272	0.14%
Simrishamn	19,685	0.44%	Trosa	5,104	0.11%	Österåker	35,929	0.80%
Sjöbo	15,997	0.36%	Tyresö	32,544	0.73%	Östhammar	911	0.02%
Skara	7,234	0.16%	Täby	54,651	1.22%	Östra Göinge	8,921	0.20%
Skellefteå	40,373	0.90%	Töreboda	2,265	0.05%	Total	4,484,665	100.00%
Skinnskatteberg	1,832	0.04%	Uddevalla	30,441	0.68%			
Skurup	8,835	0.20%	Ulricehamn	6,194	0.14%			
Skövde	25,939	0.58%	Umeå	49,915	1.11%			
Smedjebacken	1,007	0.02%	Upplands Väsby	20,998	0.47%			
Sollentuna	44,240	0.99%	Upplands-Bro	13,730	0.31%			
Solna	1,460	0.03%	Uppsala	65,920	1.47%			
Sotenäs	16,018	0.36%	Uppvidinge	2,604	0.06%			
Staffanstorp	21,122	0.47%	Vadstena	4,015	0.09%			
Stenungsund	16,105	0.36%	Vaggeryd	3,891	0.09%			
Stockholm	163,839	3.65%	Valdemarsvik	3,355	0.07%			
Storfors	1,794	0.04%	Vallentuna	20,404	0.45%			
Strängnäs	17,766	0.40%	Vansbro	2	0.00%			
Strömstad	7,204	0.16%	Vara	5,006	0.11%			
Sundbyberg	5,095	0.11%	Varberg	39,042	0.87%			
Sundsvall	38,351	0.86%	Vaxholm	9,438	0.21%			