ECgames 2017 – Case A

Team #13

April 6, 2017

1 Summary

In this report we answer the three questions outlined below. In this section, a brief preview of our procedures and main findings is provided.

- 1. What are the geographical variations in the local use of shutters? (Section 2)
 - First, we provide several plots displaying the prevalence rate of shutters throughout the Netherlands.
 - Using Moran's I, we test for spacial autocorrelation and conclude that spatial autocorrelation is present in the data.
 - Lastly, we run several regressions (OLS, Probit, Logit) at the individual level, investigating how geographical adoption patterns depend on longitude, latitude and their interaction. The results obtained confirm the existence of a West-South-West to East-North-East cut in shutter adoption. That is, below this line, shutter usage is higher than above it, corroborating the earlier found results at the municipality level.
- 2. What explains these geographical patterns? (Section 3)
 - Using binary choice models with latent processes which depend on household and municipality characteristics, a perceived risk and a sociability index, and interactions between those indices and the average adoption rate in each municipality we gauge the following two effects, that make shutters contagious: *i*) As more nearby homes adopt shutters, they become more socially accepted and therefore people are more likely to adopt shutters; and *ii*) The fact that more nearby homes adopt shutters makes any home without shutters look like a very easy target, so that individuals will adopt shutters as a response.

- Our findings show that the main mechanism driving the contagiousness of shutters is *i*). No statistically significant evidence was found favouring *ii*), though the fact that the estimated coefficient bears the expected sign suggests that it can also be at play.
- Further robustness checks confirm the findings. A placebo test provides additional confidence regarding our model specification.
- 3. By which year will each municipality in the Netherlands be saturated with shutters? (Section 4)
 - Building a tailor-made dynamic model to gauge the long-run dynamics of the data using both provided datasets, we conclude that it will take about 50 years until the vast majority of Dutch municipalities will be saturated in terms of windowshutters. The saturation points are positively related to the current share of shutters, but also depend on municipality characteristics.

$2 \quad (a)$

Geographical distribution of shutters

In this Section we analyze the geographical distribution of window shutters graphically. In order to do so, we collapsed the data at the municipality level, pooling data for 2005-2008. Figure 1 shows the prevalence of window shutters by municipality. In the South of the Netherlands the shutters are much more common than in the North and their prevalence declines gradually. Note that the 'iso-shutter lines' do not run exactly from West to East but rather from West-South-West to East-North-East.

This gradual change is very peculiar. As illustrated in Figure 2, we do not find a similar trend for other measures of burglary prevention. For comparison, we considered door locks, outdoor lights and burglary alarm. We consider these measures particularly similar to shutters, as they also include upgrades of the home that potentially prevent one from being victimized.

For the gradual change in shutters, many explanations are possible. This will be the topic of the following Section. However, we can use the graphical analysis in this Section to already rule out some simple alternative explanations for the gradual shift in shutter adoption. Figure 2 depicts the adoption of shutters conditional on various explanatory variables. First, Figure 2a shows that controlling for income does not explain much of the observed pattern - Figure 2a was prepared by regressing the prevalence of shutters on income and squared income. Controlling for the presence of other security measures does not change the picture either (Figure 2b). However, once we control for the share of catholics in a given municipality the picture changes significantly, see Figure 2c. Superficially, this appears to refute our theories of contagion and social copying mechanisms. However, there are three reasons to be careful: Firstly, it seems unlikely that catholics just have a taste for window shutters, although it might be that differing social cohesion among catholics leads to a more pronounced copying effect; Secondly, the prevalence of catholics is very highly correlated with latitude, so that it might make a latitude-based effect disappear by chance; Finally, there is still graphically evident and statistically significant clustering in the adoption of shutters even after controlling for religion. We conclude that none of the covariates that might be responsible for the effect explain it convincingly on their own, so that we need to continue to explore the social mechanisms behind it in the following sections.

Assessing Tobler's First Law of Geography

We now assess the extent to which the spacial distribution of shutters conforms with to Tobler's First Law of Geography, that is, whether there is evidence of spacial autocorrelation in the distribution of shutters prevalence in each municipality. In order to test for this, we compute Moran's I. In a nutshell, Moran's I tests the null hypothesis of absence of spacial autocorrelation in the data, against the alternative hypothesis of presence of autocorrelation. Two variations of the test are performed. The first one assesses spacial autocorrelation in the raw variable *shutters*, whereas the additional version assesses spacial autocorrelation in the residuals obtained from regressing *shutters* on some demographic and socio-economic variables.¹ The results of the spacial autocorrelation tests are presented in Table 1. Based on the associated p-values, we reject the null hypothesis of absence of spacial autocorrelation in both specifications of the test. So, we conclude that the spatial distribution of high values and/or low values in the data is more spatially clustered than would be expected if underlying spatial processes were random, thus conforming to Tobler's First Law of Geography.

How does the probability of having shutters vary with latitude and longitude?

We estimate how the probability of having shutters varies with latitude and longitude using a linear probability model, as well as logit and probit models. We start by simply regressing the binary variable capturing whether the individual has shutters on latitude, longitude and

¹Specifically, these variables are: household income, household income squared, the percentage of romancatholic population living in the municipality and the percentage of individuals in the municipality owning doorlocks, outdoorlight and burglary alarm, and the share of population in the municipality that experienced a burglary or an attempted burglary in the last 5 years.



Figure 1: Prevalence of home-security measures by municipality.

an interaction term between latitude and longitude. The results of this simple specification are shown in the first 3 columns of Table 2 (results reported for probit and logit models correspond to marginal effects evaluated at the mean). These convey the fact that the probability of having shutters decreases as latitude and longitude increase, but is positively associated with the interaction of the two, reflecting the patterns visualized in the maps above. The magnitude of the marginal effects slightly differs across models, but the direction of the effects is similar. The last 3 columns of Table 2 report the results from a including additional control variables.² Overall, including the additional controls does not change the effects previously found for latitude, longitude and their interaction.

²Controls in the regression include income, income squared, dummies for age categories, dummies for marital status, the percentage of Roman-Catholics in the municipality, a time trend and binary variables for whether the individual has doorlocks, an outdoor lock and a burglary alarm, respectively.

Variables	Ι	E(I)	sd(I)	Z	p-value*
Shutters	0.227	-0.003	0.004	51.413	0.000
Residuals	0.100	-0.003	0.004	23.037	0.000

Table 1: Moran's I tests for spacial autocorrelation - Results

*1-tail test

For performing the test we set the distance band for 0 to 5.

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	Probit	Logit	OLS	Probit	Logit
latitude	-0.318***	-0.385***	-0.414***	-0.353***	-0.402***	-0.411***
	(0.021)	(0.022)	(0.021)	(0.022)	(0.023)	(0.022)
longitude	-0.927^{***}	-1.683^{***}	-2.009^{***}	-1.821***	-2.254^{***}	-2.368***
	(0.186)	(0.193)	(0.191)	(0.207)	(0.210)	(0.203)
latitude \times longitude	0.0191^{***}	0.0334^{***}	0.0396^{***}	0.0358^{***}	0.0441^{***}	0.0463^{***}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Constant	16 /17***			F 700		
Constant	10.417			5.700		
	(1.078)			(3.716)		
Demographic Controls	NO	NO	NO	YES	YES	YES
Other Security Controls	NO	NO	NO	YES	YES	YES
Time trend	NO	NO	NO	YES	YES	YES
Socio-economic Controls	NO	NO	NO	YES	YES	YES
N	65033	65033	65033	58648	58648	58648

Table 2: Marginal effects on the probability of having shutters

Marginal effects; Standard errors in parentheses

(d) for discrete change of dummy variable from 0 to 1

* p < 0.05, ** p < 0.01, *** p < 0.001



Figure 2: Prevalence of window shutters conditional on different sets of explanatory variables.

3 Drivers of shutter adoption

We aim at gauging the following two mechanisms, which may drive shutter contagion:

- 1. A higher local adoption rate of shutters makes them more socially acceptable.
- 2. Shutters are a visible means of protection, households do not want to stay behind when their neighbors adopt them.

Assumptions

- Shutters are viewed as a nuisance in and of itself, on average. The two mechanisms driving the adoption choice are driven by this underlying assumption, otherwise people would adopt shutters for different purposes, e.g. sun protection.
- 2. More sociable households are more susceptible to social convention. Social acceptance of shutters is not directly observable, to gauge it we require variation in auxiliary variables. Under this assumption, sociability variables will properly measure the household's susceptibility to social convention.

Baseline model

In order to understand and distinguish between the two potential mechanisms driving shutter adoption behavior, we begin with specifying a model confined to the municipality level without accounting for spatial effects between municipalities. Note that to understand these effects, the spatial aspect is not necessary. Therefore, we will later include a robustness specification including a spatial dimension to control for potential spillovers.

Household *i*'s shutter adoption $y_i \in \{0, 1\}$ is given by

$$y_i = 1\{y_i^* \ge 0\},\$$

where y_i^* is a latent process modeled as

$$y_{it}^* = s_{mt}(\beta_1 + \text{SOC}_{it}\beta_2 + \text{RISK}_{it}\beta_3 + \text{SES}_{it}\beta_4) + \beta_0 + \text{SOC}_{it}\beta_5 + \text{RISK}_{it}\beta_6 + \text{SES}_{it}\beta_7 + \text{MUN}_m\beta_8 + \epsilon_{imt},$$
(1)

were s_{mt} is the average shutter adoption in municipality m, the municipality in which i resides in year t; SOC_{mt} is a sociability indicator for household i in year t; RISK_{imt} is a crime perception index of household i regarding municipality m including past victimization of household *i* and additional employed security measures; SES_{it} are household *i*'s socio-economic characteristics at time *t*; MUN_m are municipality dummies; finally, ϵ_{imt} is a household unobserved taste shock which are iid across households and follow a uniform distribution.

Endogeneity

Incorporation of municipality dummies will account for any correlation of the average shutter adoption rate with unobserved factors. With that, since the average adoption rate in each municipality is a direct function of household *i*'s adoption decision, this measure may be endogenous. In a sufficiently large sample, this endogeneity would dissapear. However, since our sample has a very limited number of households per municipality, we will conduct a robustness specification in which the average municipality adoption rate excludes the choice of household *i*, specifically $s_{m\setminus i}$.

One might worry about endogeneity of the risk indicator and additional security measures. The individual specific taste shock incorporates the factors which make a household more or less averse to shutters, ceteris paribus. To account for this possible effect, we will test a specification where we instrument for the risk indicator using the average risk perceptions of all other households in the municipality excluding household i.

SOC and RISK Factors

In order to keep the model simple and interpretable we reduce the SOC and RISK variables to a single index each, by means of principle component analysis, henceforth PCA (see below for a brief description). Another, and more important, reason for this is the fact that these variables are most likely to be endogenous and, thus, we would have to find at least one instrument for each variable. A daunting task. However, after collapsing them to 2 indices, only 2 instruments are needed.

Since variables within SOC and RISK are ordinal³ and, therefore, violate the vital multivariate normality assumption needed for standard PCA. Standard PCA is therefore not feasible. To circumvent this issue, we make use of polychoric PCA (cf. Kolenikov and Angeles, "The Use of Discrete Data in PCA: Theory, Simulations, and Applications to Socioeconomic Indices", 2004, Working Paper, for a summary), accounting for the ordinal nature of the data. The underlying idea is similar to that of estimating discrete choice models. Suppose you have \mathcal{K} ordinal variables $x_k, k \in \{1, 2, ..., \mathcal{K}\}$, with levels $1, 2, ..., K_k$. Moreover, suppose

 $^{^{3}}$ Note that some variables are of qualitative nature and we, therefore, recoded them into appropriate ordinal scales.

that you have a multivariate, latent process x^* such that

$$x^* = \Lambda \xi + \delta$$
, where $\operatorname{Var}(\delta) = \Theta$ and $\operatorname{Var}(\xi) = \Phi$

with Θ being a diagonal matrix and ξ the unobserved PCs. Of course, x^* is such that the remaining PCA assumptions are met. The relationship between x_k and x_k^* is then as follows

$$x_{k} = 1\{\alpha_{k,r-1} < x_{k}^{*} \le \alpha_{k,r}\}$$

with $\alpha_{k,0} = -\infty$ and $\alpha_{k,K_k} = \infty$ for all $k \in \{1, 2, ..., \mathcal{K}\}$. Based on this representation, the PCs for the ordinal x_k 's are then estimated sequentially where the thresholds $\alpha_{k,r}$, $k \in \{1, 2, ..., \mathcal{K}\}$ and $r \in \{1, 2, ..., K_k\}$, are estimated first and the PCs afterwards via conditional maximum likelihood. The composite indices are then chosen to be the PC with largest eigenvalue. Finally, in order to implement this idea the Stata-package polychoric is used.

Observed spatial effects

To incorporate spatial effects in the estimation, we will account for spatial effects in the observed characteristics, as specified below:

$$y_i^* = s_k (\beta_1 + \text{SOC}_i \beta_2 + \text{RISK}_i \beta_3 + \text{SES}_i \beta_4) + \beta_0 + \text{SOC}_i \beta_7 + \text{RISK}_i \beta_8 + \text{SES}_i \beta_9 + \text{MUN}_m \beta_{10} + \epsilon_i$$
(2)

where s_k pertains to the average shutter adoption in municipality $k, k \neq m$, where all the neighboring municipalities to municipality m are included in k.

Following the econometric specification and the specified assumptions, the interaction between sociability and the average adoption rate yields the average tendency to conform to social convention as shutters become more commonplace, i.e. the first effect. The interaction between risk perception and the average adoption rate will identify the effect of increased risk perception, while controlling for other protection mechanisms, on the adoption rate as shutters become more prevalent, i.e. the second effect.

Results

The results obtained for our model specifications are shown in Table 3. Only the results associated with the interaction variables will be discussed along the text, in order to keep it brief. Column (1) shows the estimates for a simple linear probability model without municipality fixed-effects. Column (2) adds these fixed-effects. Overall, results convey the

fact that shutter adoption is more driven by the tendency to conform to social convention as shutters become more common in the municipality where the individual lives. This is given by the positive coefficient on the interaction between individual sociability and rate of adoption in the municipality where the individual lives and it is significant at 0.1%. In contrast, we find no statistically significance evidence for the effect of increased risk perception, despite the fact that the estimated coefficients have the right (positive) sign. These effects are robust to the use of distinct measures of average rate of adoption in the municipality where the individual lives, as well as in the neighbour municipalities (columns (3) and (4), respectively). The use of instruments for potential endogenous variables also leaves these results unchanged (column (6)). Moreover, our placebo test for the adoption of doorlocks shows evidence of neither of these effects, providing extra confidence on our model specification (column (5)) - Indeed, as discussed above, the adoption of doorlocks is not easily visible as it is the case with window shutters.

Overall, we conclude that the main mechanism driving the contagiousness in the adoption of shutters is related to their increased social acceptability. That is, despite the fact that shutters are ugly and expensive, the more common they are in a given neighbourhood, the more socially acceptable they become and thus the more likely an individual is to get shutters as well.

SES includes household income and its square, binary variable for whether individual is married, binary variable for whether individual has higher education. MEASURES SOC and RISK are the factors obtained from PCA procedure. * p < 0.05, ** p < 0.01, *** p < 0.001

	(1)	(3)	(3)	(7)	(2)	(9)
	OLS	OLS_FE	OLS_FE_robust1	OLS_FE_spatial	OLS_FE_locks	IV_FE
Sm	0.146	0.106	-1.123***	-0.737	0.255^{**}	-0.0141
	(0.0795)	(0.0846)	(0.0881)	(4.150)	(0.0797)	(0.327)
$s_m \times RISK$	0.00803	0.00282	-0.00770	0.00194	0.0212	0.0453
	(0.0179)	(0.0185)	(0.0193)	(0.0205)	(0.0181)	(0.187)
$s_m imes SOC$	0.0195^{***}	0.0213^{***}	0.0233^{***}	0.0214^{***}	0.00469	0.0236^{**}
	(0.00483)	(0.00491)	(0.00512)	(0.00556)	(0.00478)	(0.00904)
$s_m \times MEASURES$	0.335^{***}	0.340^{***}	0.276^{***}	0.326^{***}		0.296
	(0.0297)	(0.0303)	(0.0303)	(0.0314)	(0.0340)	(0.337)
SOC	-0.000390	-0.000337	-0.000387	-0.000277	0.00822^{***}	-0.00575
	(0.00112)	(0.00114)	(0.00115)	(0.00126)	(0.00111)	(0.00363)
RISK	0.00563	0.00416	0.00614	0.00453	0.0195^{***}	0.00862
	(0.00412)	(0.00426)	(0.00432)	(0.00467)	(0.00416)	(0.0694)
MEASURES	0.0517^{***}	0.0551^{***}	0.0711^{***}	0.0588^{***}		0.464^{**}
	(0.00667)	(0.00685)	(0.00693)	(0.00756)		(0.142)
Constant	-2.469	-2.597	-21.31^{***}	-15.75***	-4.673	5.313
	(3.245)	(3.283)	(3.321)	(3.305)	(3.206)	(4.724)
$SES \times s_m$	YES	YES	YES	YES	YES	YES
SES	\mathbf{YES}	YES	YES	YES	YES	YES
Municipality FE	NO	YES	\mathbf{YES}	YES	YES	YES
Time trend	YES	YES	YES	YES	YES	YES
Ν	64024	64024	64024	64024	64024	64024
Standard errors in par	entheses					

Table 3: Explanations for geographical patterns - Results

4 Geographical adoption prediction

To predict infiltration rate of shutters throughout the Netherlands, we take into account two considerations: first, a municipality's adoption of shutters is affected by the contemporaneous propensity of adoption across other municipalities. Second, a municipality's adoption rate is affected by its previous adoption, which can be generally framed as a differential trend between municipalities at differing adoption rates.

To predict the long run time-effects, we need to build a dynamic model for which it is paramount to have a panel structure. Therefore, we aggregate at the municipality level. As it is very difficult to estimate long-run effects from short-run dynamics, we use the additional data from the nineties. By creating a specific model for the long-run dynamics, we will be able to estimate the saturation points and durations from just two time periods. This is important, as – from a long-run perspective – we only have two time points (the early and the late datasets) and merging them into one panel would be problematic as we then our time periods would differ in length. In the following, a time period thus corresponds to about 15 years, i.e., the time between the two datasets.

Theoretical setup

Following the aforementioned considerations, the growth in the shutter adoption in municipality m can be written as:

$$\Delta s_{m,t} = \alpha s_{m,t-1} + c(X_{m,t}, s_{-m,t}) \tag{3}$$

Where $s_{m,t}$ is the adoption rate in municipality m at time t; $\Delta s_{m,t} = s_{m,t} - s_{m,t-1}$; $s_{-m,t}$ is the adoption rate of all other municipalities except for municipality m at time t; finally $X_{m,t}$ is municipality m's demographics.

At the municipality saturation point $\Delta s_{m,t} = 0$ which implies that $s_{m,t} = s_m^* = -\frac{c(\cdot)}{\alpha}$ for all $t \ge t^*$ and t^* denoting the point in time at which saturation occurs first. Similarly, rewriting the estimation equation we get:

$$s_m^* = c(X_{m,t}, s_{-m,t}) + (1+\alpha)s_{m,t-1} = \dots = -\frac{c(\cdot)}{\alpha}[1 - (1+\alpha)^t] + (1+\alpha)^t s_m^0$$
(4)

Using the definition of the saturation adoption rate in the above equation yields:

$$t(s_m^0, s_m^*) = \frac{\log\left[\frac{s_m^* - s_0^*}{s_m^*}\right]}{\log(1 + \alpha)}$$
(5)

Where s_m^0 is municipality *m*'s starting saturation level and s_m^* is a function of the municipality's characteristics and the other municipalities' saturation level.

Econometric specification

The econometric specification used to estimate the above equation is given by:

$$\Delta s_{m,t} = \alpha s_{m,t-1} + X'_{m,t}\beta + \sum_{k \neq m} \lambda_k w_{m,k}^{-1} s_{k,t} + \epsilon_{m,t}$$
(6)

Where $s_{m,t}$ is the adoption rate of shutters in municipality m at time t; $w_{m,k}$ is pre-defined weight for the adoption rates between municipality m and k; $X_{m,t}$ is a linear combination of municipality m characteristics that predict the adoption behavior; $\epsilon_{m,t}$ is a municipality-time specific error term. The weighting for municipality adoption rates, $w_{m,k}^{-1}$, is defined as the inverse of the ℓ_2 distance between every two municipalities:

$$w_{m,k}^{-1} = \left[\sqrt{(\log_m - \log_k)^2 + (\operatorname{lat}_m - \operatorname{lat}_k)^2}\right]^{-1}$$

Results

To estimate the model we averaged over all households within the municipality to construct municipality level datasets. We used the difference between the two datasets, the one from the 1990's and the one from the late 2000's, to identify the effect on shutter adoption growth within this time frame. As such, when conducting counterfactual predictions regarding the required time to municipality saturation, we normalized according to the time difference between the two datasets.

Using our steady state equation, we obtained a theoretical saturation point for each municipality, depending on the characteristics of the municipality as well as the current state of shutterization. As expected, Figure 4 shows a positive but not perfect correlation between the current share of window shutters and our theoretical maximum.

4 shows the results from the estimation of equation (6). The spatial effects, namely the sum of the weighted adoption shares of the other municipalities, although not significant, we decided to still include in the specification. The final variable summarizes all the additional information on municipality m as the predicted probability of adoption derived from the preferred specification in the previous part.

Using these estimates we are able to derive the saturation rate of each municipality based on it's characteristics and the spatial effects. The graph below depicts the distribution of the saturation levels across all municipalities in the Netherlands as a function of their level of



Figure 3: Kernel density estimates of the expected time until saturation.

	(1)
	ds
$s_{m,t-1}$	-0.661***
	(0.0325)
$X_{m,t}$	1.815***
,	(0.0667)
$w_{mk}^{-1} s_{k,t}$	0.000187
110,10	(0.000242)
β_0	0.0121
	(0.0112)
N	341

Table 4: Municipality saturation point - Results

Standard errors in parentheses

* p < 0.05, ** p < 0.01, *** p < 0.001



Figure 4: The saturation point of a given municipality vs. it's current share of window shutters.

adoption in .

Table 3 shows the smoothed distribution of the time required for each municipality to reach its saturation level. The graph shows most municipalities will converge within 40 years and that the average time required to saturation is roughly 25 years.

5 Discussion

Before concluding we briefly discuss shortcomings of our methodology and how these can be addressed in future research.

- Available data. There are, in our opinion, several issues with the available data set:
 - First, and foremost, we think that having a proper panel data set with more time observations would greatly facilitate the analysis and highlight the asked effects.
 For one, with a proper panel we could gauge unobserved heterogeneity and thereby eliminate several potential sources of endogeneity.
 - Moreover, we wished to have a more detailed data set, for example, more information on an individual level. This would allow to incorporate building/room level

height relative to the ground. A useful feature in predicting the shutter usage since it can safely be assumed that you do not invest in shutters when living on the 20th floor.

- Information on whether an individual owns or rents his current residence since people who rent usually cannot adopt that easily (they need written permission from the owner). This should have a clear effect for cities in terms of adaption speed and saturation level.
- A more complete data set would us also have allowed to construct more/better IVs in Section 3, respectively use less strong aggregates for the risk and sociability indices.
- In Section 4 we use the ℓ_2 distance to construct spatial weights. Technically, they can only be seen as a crude proxy since we are dealing with distances on a sphere rather than a plane. However, we also believe that the effect of this approximation should not be too big given the relatively negligible size of the Netherlands on planet Earth.
- Even though this point relates to the availability of data, it is worth mentioning it separately due to its implications. If individuals would/could be tracked, at least on the level whether they moved (if they did), this would render a, in our view, nice experimental design. To see why, consider the case of people moving within the municipalities below/above the big rivers. In addition, consider individuals who move across this natural border. By having access to such data one can construct a straightforward experiment to test the hypothesis whether there are cultural differences between north/south in adopting shutters since, if there are, it can be expected that movers from south to north (north to south) would choose a place with (without) shutters due to their cultural background, ceteris paribus.

6 Conclusion

In this report, we investigated the drivers of shutter adoption in the Netherlands first and then predicted when the Netherlands can be expected to be saturated in terms of shutter usage. We are able to show that the main driver of individual shutter adaptation can be attributed to sociability. In addition we are able to show the perceived risk does not contribute to an individual's decision to adopt shutters.

As for the predictions when the Netherlands will finally be saturated we found that the majority of municipalities should have saturated after roughly 25 years from 2008 on. That is, we expect that the majority adopted shutters by 2033.