

# Applying disclosure control on average values

Niek van Leeuwen  
Statistics Netherlands

[nlwe@cbs.nl](mailto:nlwe@cbs.nl)



# Disclosure control average value (1)

Applied for average value of dwellings

## Publication

- Average value
- Total number dwellings  
published number  
aerial photograph



## Disclosure control average value (2)

- **Primary** non disclosure, **variation** of values
- **Secondary** non disclosure,  
in **combination** with other small areas

# When is disclosure control needed

Values are almost identical

One divergent value in series, high or low

Criterion

When one value is known, others may be estimated with a minimum uncertainty.

Otherwise non disclosure

# Calculation

Estimation

$$\hat{w}_1 = W - (n - 1).w_2$$

$\hat{w}_1$  = estimation

$w_2$  = known value

$n$  = number of dwellings

$W = \sum_{i=1}^n w_i$  = sum of values

$p$  = minimum deviation

Uncertainty estimated value

$$abs(w_1 - \hat{w}_1) > \frac{p}{100} . w_1$$

Best estimation of highest value is by knowing the highest but one.

# Results primary non disclosed

Dwellings 2008, Netherlands, total 6.913.293 with x, y-coordinate,  $p > 15\%$

Dwellings per grid	100m grid		500m grid	
	Dwellings involved	Dwellings disclosed (%)	Dwellings involved	Dwellings disclosed (%)
<b>1-2</b>	290.307	<b>0.0</b>	37.155	<b>0.0</b>
<b>3</b>	84.252	<b>37.8</b>	22.153	<b>51.0</b>
<b>4</b>	76.672	<b>43.7</b>	22.532	<b>62.8</b>
<b>5</b>	73.260	<b>62.7</b>	20.665	<b>79.6</b>
<b>6</b>	76.374	<b>78.1</b>	20.214	<b>93.1</b>
<b>7</b>	77.952	<b>87.6</b>	18.130	<b>97.3</b>
<b>8</b>	83.648	<b>91.4</b>	16.888	<b>98.5</b>
<b>9</b>	85.824	<b>93.6</b>	15.183	<b>99.5</b>
<b>10</b>	89.590	<b>95.5</b>	13.750	<b>99.7</b>
<b>11-plus</b>	5.975.414	<b>98.6</b>	6.726.624	<b>100.0</b>

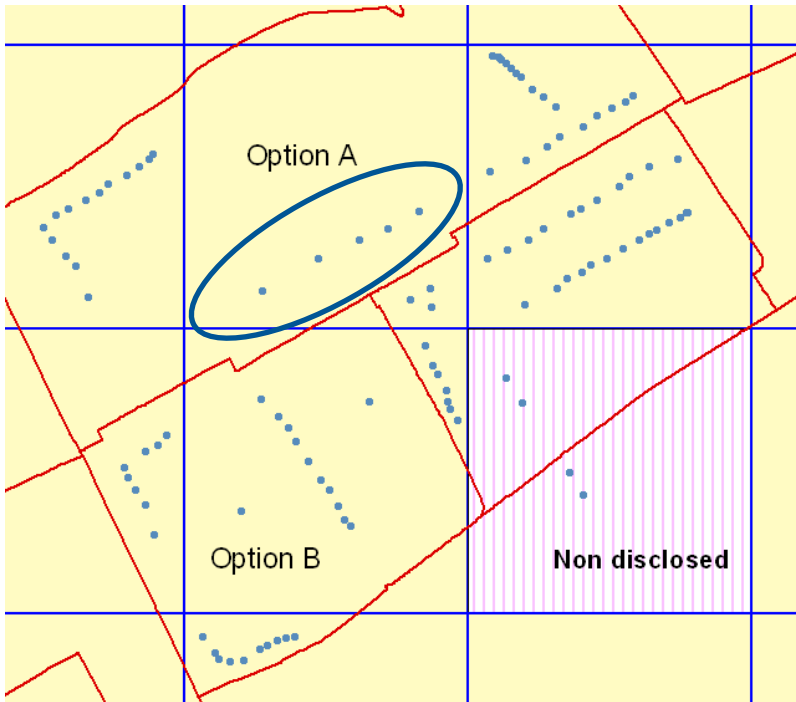
## Secondary non disclosed (1)

Combinations 100m grids with other small area publications

- Neighbourhoods and districts
- Hierarchical grid (500m)

## Secondary non disclosed (2)

Combining 100m grid and small neighbourhoods



Avoid inheritance of disclosure.  
Not enough variation in distinct clusters (option A)

Hierarchical grid is special case  
in combining different areas



# Contribution to non disclosure average dwelling value

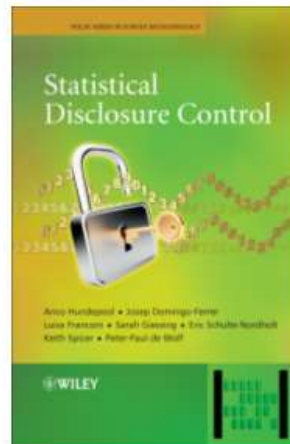
	100m grid	500m grid
<b>Primary non disclosure</b>		
Small number (1 – 2)	4.2 %	0.5 %
No variation ( $p < 15\%$ )	3.6 %	0.4 %
<b>Secondary non disclosure</b>		
Combination with neighbourhood areas	0.1 %	0,6 %
<b>Total</b>	<b>7.9 %</b>	<b>1.6 %</b>

# AVAILABLE NOW!

## Statistical Disclosure Control

Anco Hundepool, Josep Domingo-Ferrer, Luisa Franconi, Sarah Giessing,  
Eric Schulte Nordholt, Keith Spicer, Peter-Paul de Wolf

Order your copy  
online at  
[www.wiley.com](http://www.wiley.com)



ISBN: 978-1-1199-7815-2  
Hardcover  
302 pages  
September 2012  
Wiley  
US \$99.95

### Statistical Disclosure Control:

- Presents a combination of both theoretical and practical solutions
- Introduces all the key concepts and definitions involved with statistical disclosure control.
- Provides a high level overview of how to approach problems associated with confidentiality.
- Provides a broad-ranging review of the methods available to control disclosure.
- Explains the subtleties of group disclosure control.
- Features examples throughout the book along with case studies demonstrating how particular methods are used.

WILEY

# Questions?