MODELLING OVERWEIGHT DEVELOPMENT IN UTRECHT: AN AGENT-BASED SIMULATION STUDY

WEIPER, M. L. V.
LEAL CASTILLO, J. E.
VAZQUEZ SANCHEZ, I. A.

Modelling Overweight Development in Utrecht: An Agent Based Simulation Study

Introduction

Overweight and obesity is one of the main health threats to millions of people worldwide (World Health Organization, 2021). People with a higher body-mass index (BMI) are more prone to develop cardiovascular diseases, diabetes and even certain forms of cancer (World Health Organization, 2021). In 2021, over half of the adult population in the Netherlands was classified as overweight, with a BMI of 25 or greater, while 14% were considered obese, with a BMI of 30 or higher (CBS, 2022). As a result of this, the Dutch government has issued a national prevention strategy (Ministerie van Volksgezondheid, Welzijn en Sport, 2018) with its goal to reduce overweight to 38% or less by 2040. The government has proposed several strategies to fulfil the goal of reducing overweight, such as cutting taxes on healthy fruits and vegetables (Ministerie van Volksgezondheid, Welzijn en Sport, 2018). However, a recent analysis shows that overweight has barely reduced since the start of this strategy in 2018 (Centraal Bureau voor de Statistiek, 2022, p. 20). Therefore, governmental strategies to reduce overweight must be improved to be successful in the near future.

One approach to model overweight and the potential consequences of governmental interventions is agent-based modelling (ABM; Auchincloss & Diez Roux, 2008). ABM is a simulation-based modelling approach that can model interactions of spatial, temporal and social aspects of a society with regard to a phenomenon such as overweight and obesity. Over time, agents (e.g., households in a city) can develop certain characteristics and behaviours that shape their surroundings, which in turn shape the agent's characteristics and behaviour in the future. This way, ABM is an effective way to model the complex multi-level interactions and feedback loops which shape phenomena such as obesity in society.

In the past years, ABM has been increasingly used for modelling obesity and overweight (Bruzzone et al., 2012; Giabbanelli et al., 2021; Li et al., 2018; Orr et al., 2016). For example, Orr et al. (2016) modelled racial disparities in obesity over multiple generations of agents and found that school quality has a strong influence on the inequality between communities in obesity over a larger period of time, while the influence of food store quality or sport facilities barely changed over time. This shows that, in contrast to other methods like linear regression, ABM is an important method to model the complex interactions over a large period of time (Auchincloss & Diez Roux, 2008). Put differently, ABM is a novel, yet validated approach to investigate the development of overweight within a population.

Therefore, in this case study, an ABM will be used to model the development of BMI in the Dutch city of Utrecht. As one of the biggest cities of the country, Utrecht's inhabitants are people with many different backgrounds and characteristics. Thereby, the city represents an interesting case for modelling overweight in The Netherlands. By combining real-world data on population density, BMI, and income, this study addresses an important limitation of earlier studies, which often did not use real data in the calibrations of their model (Giabbanelli et al., 2021). Furthermore, this research contributes to the current state of the art literature in two main ways. First, it extends upon the growing body of scientific literature on the use of ABM to model overweight. And second, it contributes to the understanding of the development of obesity in the Dutch urban context with and without interference from governmental authorities.

Specifically, the simulation will model the development of overweight in Utrecht over 28 years (2012 to 2040), with a calibration period of the first 8 years. Additionally, it will examine the potential impact of a tax increase on fast food stores in Utrecht. The hypothesis is that increasing taxes on fast food will reduce the probability of consumption of unhealthy, high-calorie foods, leading to a decrease in overall calorie intake and potentially improved BMI for the population in the simulated environment from 2020 onwards. This is supported by the evidence that taxing unhealthy food and drinks can result in reduced consumption and improved population health outcomes. For instance, after Mexico introduced a 10% tax on fast food in 2014, there was a 6% reduction in the purchase of taxed food and a decrease in average daily calorie intake, leading to improved population health (Bonilla-Chacin et al., 2016).

The outcomes of this simulation will be compared to the outcomes of a scenario without governmental interference. It is expected that in the scenario without government intervention, the prevalence of overweight will be greater and more unevenly distributed compared to the scenario with the impact of the tax increase.

Methodology

Data collection

The initial step involved gathering district-wise information on the percentage of overweight population, total population, and average annual income in Utrecht from 2012 to 2020. Household and fast-food store locations were obtained using QuickOSM and a map was used to depict the division of Utrecht into districts. Table 1 and Table 2 provide a comprehensive overview of the collected datasets.

Name	Scale Type	Temporal Range	Temporal Resolution
Total adult population per district	ratio	2010-2021	1 year
Average income per district	ratio	2012-2020	1 year
Percentage of people with overweight per district	ratio	2012-2020	2 years

Table 1. Description of non-spatial datasets, sources available in *Appendix A*

Name	Data geometry	CRS	Number of features	Extent
Districts map	Region	EPSG: 3857 (transformed to 28992)	10	(126434.76, 141834.39, 448710.01, 461609.64)
Household	Point	EPSG: 28992	163,621	(129835.87, 141614.53, 450297.90, 460842.46)
Fast-Food Stores	Point	EPSG: 28992	212	(130566.50, 140271.89, 452331.0, 459573.79)

Table 2. Description of spatial datasets, sources available in *Appendix B*

Data preparation

An intersection operation was performed in the Graphical Modeller tool of QGIS to determine the district of each household by finding the overlap between the household and district data. The algorithm clipped the household vector layer using the district polygon boundaries, generating a layer containing only the parts of the household features within each district. As a result, the attributes table of the resulting layer was enriched with attributes from both the input and clipping layers, with the district attributes added as new columns, providing the needed information about the district location of each household.

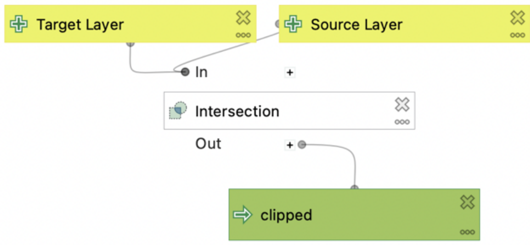


Figure 1. Intersection operation implemented in QGIS

Then, the ratio of households (agents) to be assigned to each district was chosen while maintaining the original population proportions of each district. The annual income and BMI of each household was established based on a normal distribution that followed the original district proportions. Additionally, the shortest distance of each household to a fast-food restaurant was calculated using the Euclidean distance method.

Model Proposal

A system of differential equations was proposed to model the evolution of BMI over time, considering factors such as annual income, proximity to the nearest fast-food store, and taxes on fast-food consumption. Furthermore, it was assumed that the annual income would increment each year.

$$\frac{d(BMI)}{dt} = \left[k_1 \times \frac{1}{d} \times \left(f(d,I,p)\right)\right] - [k_2 \times BMI]$$

$$\frac{d(I)}{dt} = k_5 \times I$$

Where:

$$f(d,I,p) = \frac{1}{1 + e^{(-k_1 \times d + k_3 \times I - (1 + p \times k_4))}}$$

The left side of the equation, $\frac{d(BMI)}{dt}$, represents the rate of change in BMI over time and is determined by the product of the distance to the nearest fast-food outlet (d) and the probability of frequenting fast-food stores ($f(d,I,p)$). This probability is based on the relationship between distance (d), annual income (I), taxes on fast-food (p).

$f(d,I,p)$ was modeled using a logistic function, which includes the relationship between distance, income, and taxes on fast-food. The function represents the probability of frequenting fast-food outlets and ranges from 0 to 1, with 0 indicating a low likelihood and 1 indicating a high likelihood. The logistic function considers the distance as a negative influence and the income as a positive influence on the rate of change of BMI. Additionally, the equations have weight constants (k_1, k_2, k_3, k_4) that represent the effect of distance to fast-food stores, previous BMI, income, and taxes on the current BMI, respectively.

The rate of change in annual income is represented by the second equation, $\frac{d(I)}{dt}$, the equation takes into account the current income and the increase percentage k_5, which was calculated using the average annual income data per district from 2012 to 2020.

Calibration

To make sure the model accurately represents the phenomenon being studied, the model's parameters k_1, k_2 and k_3 were calibrated using the brute-force method to find the most accurate values for these parameters per district. The mean absolute error (MAE) was used as the objective function and the values with the lowest MAE, as listed in Table 3, were selected for analysis.

District	k1	k2	k3	MAE
Binnenstad	20	0.01	0.01	3.50
Leidsche Rijn	10	5	5	1.50
Noordoost	15	0.01	5	0.81
Noordwest	5	5	5	3.50
Oost	15	5	0.02	0.04
Overvecht	20	5	5	3.08
West	15	0.01	0.01	2.06
Zuid	15	5	15	1.39
Zuidwest	15	0.01	5	1.56

Table 3. Calculated values for the model's parameters per district, MAE definition available in *Appendix C*

Implementation

Finally, with the values obtained in the calibration step, three scenarios were implemented: the first with no tax increase on fast food, a second with a 10% tax increase, and a third with a 20% tax increase. The obtained results will be reviewed in the following section.

Results

The results of our simulation study show two main patterns of BMI development. First, the results depicted in Figure 2 and Map 1 demonstrate that in most districts, the percentage of people with overweight naturally decreases in this simulation. This is true for both the calibrated, and the simulated timesteps. The districts with the strongest decrease without governmental interventions are West and Zuidwest. In the latter timesteps of this simulation, the BMIs even decrease to an unhealthy extent with 20% of households having a BMI of lower than 18.5 and thereby being classified as underweight (in the no taxes simulation).

Second, the results show that an increase in taxes on fast-food generally has little to no impact on the proportion of people with overweight in the different districts of Utrecht. There are, however, two noticeable exceptions to this rule as the districts Binnenstad and Oost do show an effect of a tax increase. Taking a closer look at these two districts, one can see that both districts are characterized for having below-average annual incomes compared to other areas in Utrecht, as well as a high density of fast-food restaurants (see Map 2). Furthermore, these two districts are the only ones in which the proportion of people with overweight naturally (i.e. without governmental intervention) increases instead of decreases, which leads to interesting conclusions regarding the effects of taxes on fast-food.

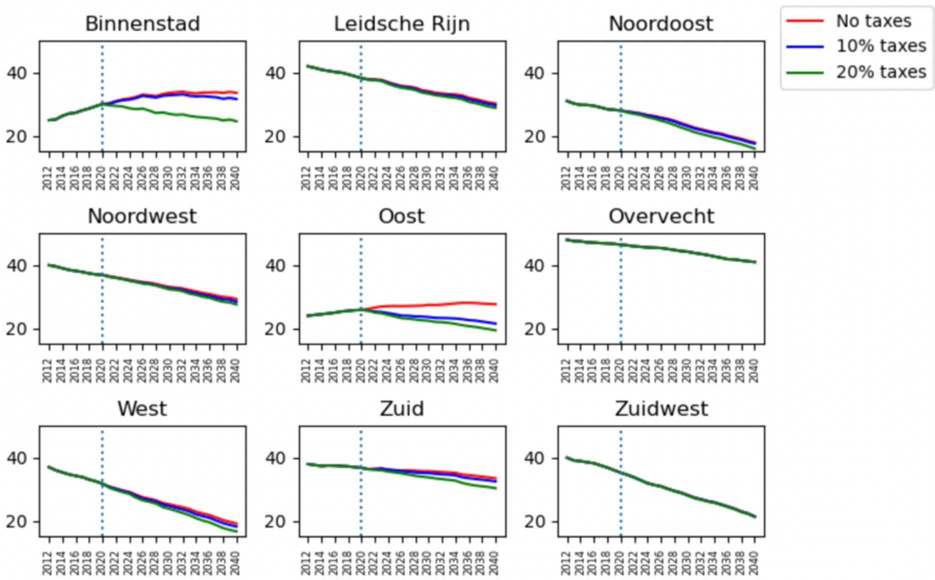
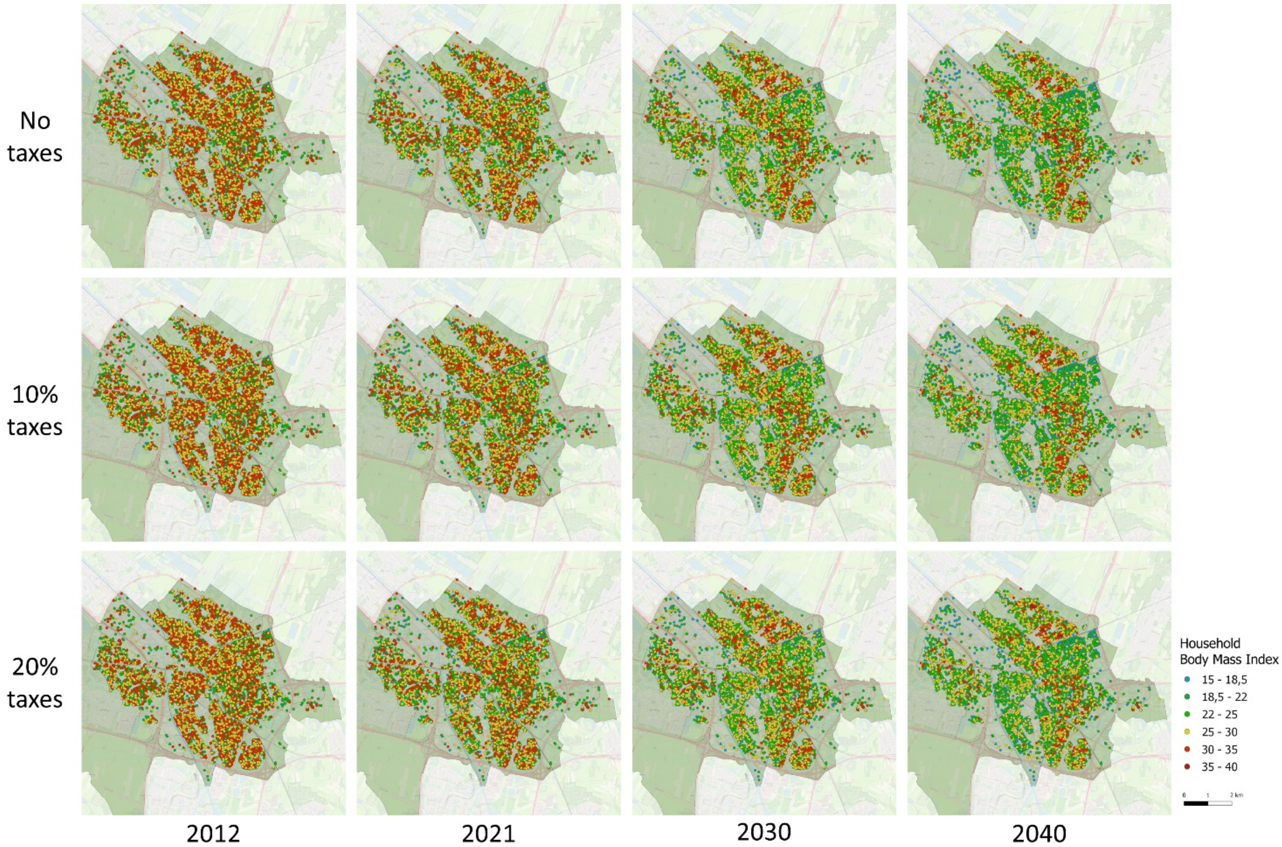
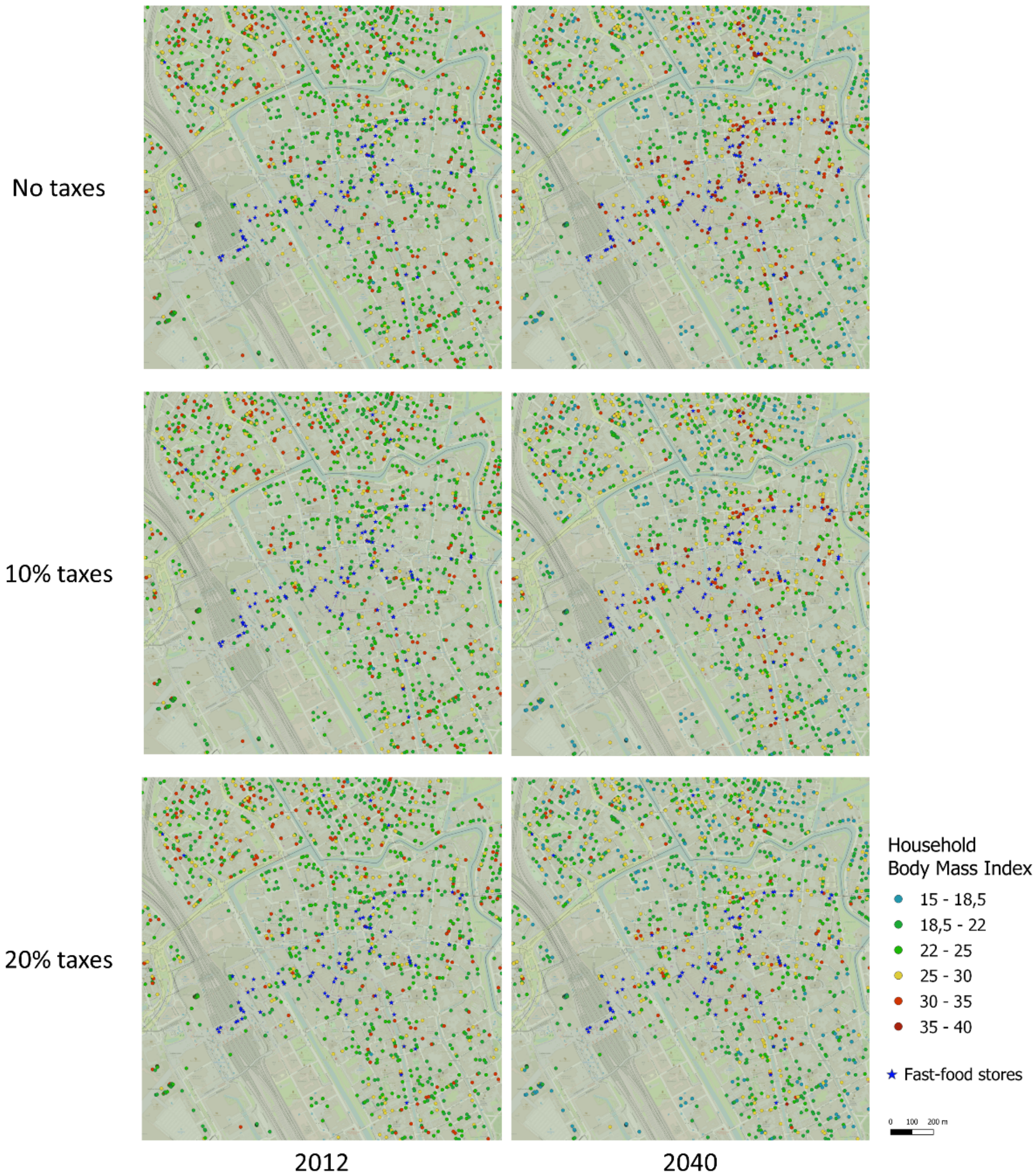


Figure 2. Percentage of overweight people (BMI > 25) per district from 2012 to 2040



Map 1: Distribution of household BMI in the study area over time and simulated conditions.



Map 2: Distribution of household BMI in the district Binnenstad near fast-food restaurants in the simulations of 2012 and 2040. Note the distribution of BMI in 2040 with regard to the location of the fast food restaurants.

Discussion

The results showed a structural decrease of overweight in most districts, regardless of the fast-food taxes imposed. This could be due to factors such as relatively high annual incomes, which provide access to healthier food options, and greater distances from fast-food restaurants, reducing the opportunity for unhealthy eating habits. In contrast, the proportion of people with overweight in the districts Binnenstad and Oost naturally increased. These districts are characterized by below-average annual incomes and a high density of fast-food restaurants within the district, potentially making the inhabitants of these districts more susceptible to unhealthy eating habits and leading to higher BMI when there is no governmental intervention (Orr et al., 2016). These results highlight the crucial role of income and access to fast-food in shaping eating habits and BMI.

The imposition of taxes did result in a noticeable decrease in the percentage of overweight people in the two districts that otherwise had an increasing proportion of people with overweight. This is likely due to the lower annual incomes in these districts limiting the ability of residents to afford fast-food, despite the close proximity to fast-food restaurants. This effect of taxes was not observed in districts with higher incomes and a lower density of fast food stores. A potential explanation is that households with higher annual incomes and greater distances to fast-food restaurants have better opportunities to consume healthier food, both because of their financial access to such, and because of the increased distance to unhealthy food options.

Our results highlight the importance of considering the interconnectedness of multiple factors, including income, access to fast-food, and public health initiatives in addressing the overweight and obesity epidemic. The findings suggest that the imposition of taxes on fast food can play a critical role in reducing unhealthy food consumption and improving the overall health of the population. However, it is important to consider the potential for unintended consequences, such as a decrease in revenue for fast-food restaurants, and to weigh the benefits against the costs when considering the implementation of such policies.

In comparison to more simplistic approaches, the methodological approach of ABM allowed us to simulate such complex phenomena as (un-)healthy eating habits. This way, the current study contributes to both, the scientific literature on ABM and health behaviors, but also to the broad base of knowledge of how governmental interventions can influence overweight and obesity in The Netherlands.

Conclusion and Further Research

Nevertheless, this study also has important limitations. First, the system of differential equations proposed, and the ABM used in this study are oversimplified representations of the complex relationship between taxes, fast-food consumption, and BMI. Even though we did calibrate the model for each district separately, the model still assumes certain generalized principles that affect the development of the BMI and thereby ignores individual and group differences to a certain extent. Furthermore, our set of differential equations did not include a regression to the mean BMI, but instead the BMI solely decreased or increased with our parameters. This way, at the end of the simulation, many households had an underweight BMI instead of a healthy one. Since this simulation was used to model the development of overweight and obesity, this effect is of little importance to the conclusions that we can draw from this model. However, this development should be taken into account in future research in order to model healthy eating habits and BMI in general.

Second, it is essential to note that the study only focused on the effect of fast-food taxes on BMI and did not consider other factors that may also influence BMI, such as genetics, physical activity levels, and environmental and socioeconomic factors. Future research should include such factors for an improved modelling of overweight and obesity. Third, next to the limitations of the system of differential equations, ABM has its own limitations, such as not being able to accurately capture the real-world behavior and interactions between individuals. While the model offers valuable insights, it may not fully reflect the real-world complexities and more comprehensive models may be necessary to better understand the connection between fast-food taxes and public health.

Lastly, although the hypothesis was supported by the results in two districts, it is not possible to apply the findings to other populations or regions without further research. Therefore, more studies are required to assess the generalizability of the findings and evaluate the impact of fast-food taxes in different settings. Future research could also explore other policy options aimed at improving public health, such as subsidies for healthy foods or marketing campaigns promoting healthy eating habits. Finally, future studies should aim to include a more robust model that considers factors such as the appearance of new fast-food stores in the city, the quality of food, and the inclusion of restaurants and supermarkets.

In conclusion, while this study provides evidence of the potential impact of fast-food taxes on BMI in certain populations, it is crucial to consider the limitations of the study and the need for further research to determine its generalizability and the interplay of various factors that influence BMI.

References

Auchincloss, A. H., & Diez Roux, A. V. (2008). A New Tool for Epidemiology: The Usefulness of Dynamic-Agent Models in Understanding Place Effects on Health. *American Journal of Epidemiology*, *168*(1), 1–8. https://doi.org/10.1093/aje/kwn118

Bonilla-Chacin, M. E., Iglesias, R., Suaya, A., Trezza, C., & Macías, C. (2016). Learning from the Mexican Experience with Taxes on Sugar-Sweetened Beverages and Energy-Dense Foods of Low Nutritional Value: Poverty and Social Impact Analysis. *Health, Nutrition and Population Discussion Paper*. https://openknowledge.worldbank.org/handle/10986/24701

Bruzzone, A., Novak, V., & Madeo, F. (2012). *Agent based simulation models for obesity epidemic analysis*.

CBS. (2022). *Hoeveel volwassenen hebben overgewicht? - Nederland in cijfers 2022 | CBS* [Webpagina]. Hoeveel volwassenen hebben overgewicht? - Nederland in cijfers 2022 | CBS. https://longreads.cbs.nl/nederland-in-cijfers-2022/hoeveel-volwassenen-hebben-overgewicht

Centraal Bureau voor de Statistiek. (2022, March 10). *Overgewicht, roken en alcoholgebruik nauwelijks gedaald sinds 2018* [Webpagina]. Centraal Bureau voor de Statistiek. https://doi.org/10/overgewicht-roken-en-alcoholgebruik-nauwelijks-gedaald-sinds-2018

Centraal Bureau voor de Statistiek. (2022, October 3). *Meer obesitas en diabetes bij volwassenen met armoederisico* [Webpagina]. Centraal Bureau voor de Statistiek. https://www.cbs.nl/nl-nl/nieuws/2022/40/meer-obesitas-en-diabetes-bij-volwassenen-met-armoederisico

Giabbanelli, P. J., Tison, B., & Keith, J. (2021). The application of modeling and simulation to public health: Assessing the quality of agent-based models for obesity. *Simulation Modelling Practice and Theory*, *108*, 102268.

Li, Y., Du, T., & Peng, J. (2018). Understanding Out-of-Home Food Environment, Family Restaurant Choices, and Childhood Obesity with an Agent-Based Huff Model. *Sustainability*, *10*(5), Article 5. https://doi.org/10.3390/su10051575

Ministerie van Volksgezondheid, Welzijn en Sport. (2018a, April 19). *Maatregelen Nationaal Preventieakkoord* [Onderwerp]. Ministerie van Algemene Zaken. https://www.rijksoverheid.nl/onderwerpen/gezondheid-en-preventie/nationaal-preventieakkoord

Ministerie van Volksgezondheid, Welzijn en Sport. (2018b). *Nationaal Preventieakkord—Naar een gezonder Nederland*.

Orr, M. G., Kaplan, G. A., & Galea, S. (2016). Neighbourhood food, physical activity, and educational environments and black/white disparities in obesity: A complex systems simulation analysis. *Journal of Epidemiology and Community Health*, *70*(9), 862–867. https://doi.org/10.1136/jech-2015-205621

World Health Organization. (2021, June 9). *Obesity and overweight*. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

Appendices

Appendix A)

Non-spatial datasets sources:

- **Total adult population per district:** Gemeente Utrecht Data:
  - Link: https://utrecht.incijfers.nl/jive/
- **Average income per district:** Gemeente Utrecht Data:
  - Link: https://utrecht.incijfers.nl/jive/
- **Percentage of people with overweight per district:** Gemeente Utrecht Data:
  - Link: https://utrecht.incijfers.nl/jive/

Appendix B)

Spatial datasets sources:

- **Districts map:** ARGIS:
  - Link: https://www.arcgis.com/home/item.html?id=3640a7b4e914417e815331cec2ee00db&sublayer=0
- **Household:** QuickOSM query:
  - Key: “addr:housenumber” in Utrecht, filtering: website and amenity = NULL
- **Fast-food stores:** QuickOSM query:
  - Key: “amenity”=”fast_food” in Utrecht

Appendix C)

Mean Absolute Error (MAE) formula:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \check{y}_i|$$

Where:

- n: represents the number of observations
- y_i: represents the true value
- $\hat{y}_i$: represents the predicted value