**README. Project ‘Equity in the Dutch long-term care system’.**

**Overview of code and input**

Marianne Tenand[[1]](#footnote-1) Pieter Bakx[[2]](#footnote-2) Eddy van Doorslaer[[3]](#footnote-3)

April 2021

1. **Purpose of this document**

The purpose of this document is to provide an overview of the code that was written as part of the project *Equity in the Dutch long-term care system*.

By making our code open access, we hope it can be used for replication purposes, educational purposes or by any researcher interested in conducting analyses similar to the ones we implemented.

The objective to make our code open access came only *after* the project was already to a large extent completed. While this implies that the code may not be optimally designed for this purpose, we have made our best to ensure it can be read and understood by external parties.

If you have any question on the code, any remark or suspicion that there might be an error, please contact us (marianne.tenand.pro[at]gmail.com).

1. **About the project**

This project was led by Marianne Tenand, Pieter Bakx and Eddy van Doorslaer between 2017 and 2020. Its general aim was to assess horizontal inequity in eligibility for and use of long-term care services in the Netherlands, when a unified system of social insurance for long-term care was still in place (reference year: 2012).

This is a project in applied economics, relying on econometric methods.

The results from the empirical analyses were published in two scientific articles:

* Tenand, M., Bakx, P. & van Doorslaer, E. (2020). “**Equal long-term care for equal needs with universal and comprehensive coverage? An assessment using Dutch administrative data**”, ***Health Economics***, 29(4):435–451. <https://doi.org/10.1002/hec.3994> [[Article Open Access](https://www.google.com/url?q=https%3A%2F%2Fonlinelibrary.wiley.com%2Fdoi%2Fepdf%2F10.1002%2Fhec.3994&sa=D&sntz=1&usg=AFQjCNEb5Z3WmB7p7TM0cWdp9JeFKlQTvw)] [[Supplementary Material](https://www.google.com/url?q=https%3A%2F%2Fonlinelibrary.wiley.com%2Faction%2FdownloadSupplement%3Fdoi%3D10.1002%252Fhec.3994%26file%3Dhec3994-sup-0001-Supplementary%2Bmaterial.pdf&sa=D&sntz=1&usg=AFQjCNFhq7p96H5hjq3uH09vEGsrOiAoaQ)]

In this document, we refer to this article as **[Equity Paper 1].**

* Tenand, M., Bakx, P. & van Doorslaer, E. (2020). "**Eligibility or use? Disentangling the sources of horizontal inequity in home care receipt in the Netherlands**", [***Health Economics***](https://www.google.com/url?q=https%3A%2F%2Fonlinelibrary.wiley.com%2Fjournal%2F10991050&sa=D&sntz=1&usg=AFQjCNFJQLLJu2LDg2RjmWg5dGk3PKmVNQ), 29(10):1161-1179. <https://doi.org/10.1002/hec.4126>[[Article Open Access](https://www.google.com/url?q=https%3A%2F%2Fonlinelibrary.wiley.com%2Fdoi%2Ffull%2F10.1002%2Fhec.4126&sa=D&sntz=1&usg=AFQjCNEFpebInsW93jppBSyDtG7GAluUrw)] [[Supplementary Material](https://www.google.com/url?q=https%3A%2F%2Fonlinelibrary.wiley.com%2Faction%2FdownloadSupplement%3Fdoi%3D10.1002%252Fhec.4126%26file%3DManuscript%2BHEC-19-0626_R1_Supplementary%2BMaterial_Proof_v2.pdf&sa=D&sntz=1&usg=AFQjCNGCeWuoxNou_mYDB-ZnSHdqVu5STQ)].

In this document, we refer to this article as **[Equity Paper 2].**

Description of the analyses can be found in these publications and their Supplementary Materials.

1. **Licensing**

The code is licensed under a Creative Commons Non Commercial license (CC-BY-NC): it can be freely reused, in particular for scientific publications, provided the two following conditions are fulfilled:

* The document is appropriately referred to as:

Tenand, M., Bakx, P. & van Doorsaler, E. (2021). **“README. Project `Equity in the Dutch long-term care system’: overview of code and input**”, DOI: 10.25397/eur.14394068

A picture containing text, clipart

Description automatically generated

* The relevant scientific publication is referred to (either [Equity Paper 1] or [Equity Paper 2], as described above).

1. **Software used**

We used software Stata. The code is written in *dofiles* (files .do). Data is created in format .dta.

1. **Dofiles allowing the creation of the datasets used for [Equity paper 1]**

The dofile can be found using the following DOI: 10.25397/eur.14394158

[Equity Paper 1] leverages a rich set of administrative data.

A.0.0.Master\_phase1.dois the master dofile allowing the creation of the datasets that are used to produce the results of [Equity paper 1].

It calls a series of sub-programs that link together different administrative sources: A.1.Indications.do / A.2.Individual\_info.do / A.3.Death.do / A.4.HC\_use.do / A.5.IC\_use.do / A.6.PB\_use.do / A.7.Household\_info.do / A.8.Partners.do / A.9.Income.do / A.10.Wealth.do / A.11.Address.do

Finally, it calls two programs allowing the creation of additional variables, relevant sample selection and the creation of sub-datasets (to be used in different parts of the analysis): A.12.Variables\_creation.do / A.13.Selection.do / A.14.Subdatasets.do

1. **Dofiles outputting the results of** **[Equity paper 1]**

3.3.Code\_Article Equity 1\_R1.dois the dofile containing the code producing all the descriptive statistics and empirical analyses of [Equity paper 1].

1. **Dofiles allowing the creation of the datasets used for [Equity paper 2]**

[Equity Paper 2] leverages a rich set of administrative data, similar to the ones used in [Equity Paper 1].It additionally linked a large survey conducted in the Netherlands in 2012.

C.Health\_monitor.do uploads the Health Monitor survey and generate a dataset keeping only the individuals aged 65+.

B.Master\_phase2.do links the Health Monitory survey information with a range of administrative data sources. This code makes use of datasets created for [Equity Paper 1].

0.Labels.do creates labels for the variables we created.

1. **Dofiles outputting the results of [Equity paper 2]**

The dofile can be found using the following DOI: 10.25397/eur.14394158

3.5.Code\_Article Equity 2\_R1.do contains the code producing the descriptive statistics and analyses as presented in the published version of [Equity Paper 2]. We also provide 3.2.Code\_Article Equity 2.do, which contains the code before the revision of the paper requested by the editorial board of *Health Economics*.

3.3.Code\_Article Equity 2\_Robustness checks\_R1.do includes robustness checks, as presented in the published version of [Equity Paper 2].

3.5.Code\_Article Equity 2\_Revision.do provides additional checks that were performed during the revision process, but have not been included in the publication for the sake of conciseness.

3.2.Code\_Article Equity 2\_Variable definition.do is a sub-program that needs to be run along the four previous dofiles and allows the creation of additional variables.

4.Figures\_Article Equity 2.do leverages the analysis results to produce the figures presented in the article.

1. **Ado file used in [Equity paper 2]**

The empirical analyses in [Equity paper 2] include a Lasso used for variable selection. For this, the code calls an ado-file, 0.LassoShooting.ado.

This .ado was adapted from a code released together with the following publication:

Belloni, A., & Chernozhukov, V. (2013). **Least squares after model selection in high-dimensional sparse models.** *Bernoulli*, *19*(2), 521-547. DOI: 10.3150/11-BEJ410 [[link to paper](https://www.jstor.org/stable/23525734?seq=1)] [[link to code](https://faculty.fuqua.duke.edu/~abn5/BernoulliFiles.zip)]

1. **Data sources: Statistics Netherlands**

The empirical analyses leverage individual- or household-level data (called microdata). In addition, it combines these microdata with a number of aggregate level data, or listings.

* 1. ***Data access***

The results presented in this article are based on calculations by the authors using non‐public microdata from Statistics Netherlands (CBS). The datasets used include the *Gezondheidsmonitor Volwassenen en Ouderen* (Health Monitor), wave 2012, provided by the GGDs (Municipal Health Services), CBS and RIVM (Dutch National Institute of Public Health).

Under certain conditions and a confidentiality agreement, these microdata are accessible for statistical and scientific research. For further information: [microdata@cbs.nl](mailto:microdata@cbs.nl). If access is granted, data can be accessed via a remote access environment.

All micro-data include a pseudomyzed personal identifier that allows to link (*merge*) the various datasets. Exact linkage of the Health Monitor survey respondents with the administrative data is also possible thanks to the availability of the pseudomized identifier in the survey.

Exploitation of the data and publication of the results are made in compliance with the European privacy legislation (GDPR, May 25th, 2018).

* 1. ***List of microdata used***

Table 1 presents the micro data that were used. The first column provides the original names of the datasets, as recorded in CBS remote access environment.

Because some datasets were only available in a spss (.sav) format when we started the project, as a preliminary step we converted a number of datasets into .dta data, saving them under different names (which are then indicated in Column 2).

Table 1: Name of microdata in the CBS remote access environment and in our code (if different)

|  |  |
| --- | --- |
| ***Name of dataset in the CBS remote access environment*** | ***Name in our code* (if different)** |
| **(1)** | **(2)** |
| GEMON 2012V3 | data\_healthmon |
| GBAPERSOON2017TABV1 |  |
| GBAOVERLIJDENTAB2017V1 |  |
| GBAHUISHOUDENS2017BUSV1 |  |
| GBAVERBINTENISPARTNER2014BUSV1 |  |
| GBAADRESOBJECT2017BUSV1 |  |
| VSLGWBTAB2017091V1 |  |
| NIETVSLGWB2018TABV4 |  |
| 130704 KINDOUDERTAB 2012v1 |  |
| 130614 MEDICIJNTAB 2011V1 |  |
| ZVWZORGKOSTEN2011TABV1 |  |
| ZORGMVTAB 2012V1 |  |
| GBAPERSOONTAB 2014V2 | individualinfo\_2014 |
| 131001 GBAOVERLIJDENTAB 2012V2 | dod\_2012 |
| 140606 GBAOVERLIJDENTAB 2013V1 | dod\_2013 |
| INDICAWBZTAV2012V1 | *indication\_2012.dta* |
| GEBZZVTAB2012V2 | zzv\_new\_2012 |
| ZORGMVTAB 2012V1 | zmv\_2012 |
| PGBAWBZ2012TABV1 | pgb\_2012 |
| GBAHUISHOUDENS2017BUSV1 | hhinfo\_2012 |
| GBAVERBINTENISPARTNERBUS 2012V1 | partners\_2012 |
| Integraal Huishoudens Inkomen 2012V2 | hhincome\_2011 |
| ipi 2012V2 | rinpersoonkern\_2011 |
| Integraal-Vermogen 2012V2 | hhwealth\_2012 |
| GBAADRESOBJECTBUS 2014V3 | address\_2012 |
| 141209 VSLGWTAB 2014V2 | houses |

* 1. ***Code listings***

We also combine aggregate level datasets that are directly provided by CBS. In the remote access environment they are called Code Listings, and can be found under: *K:\Utilities\Code\_Listings.* The code listings used in our project are listed in Table 2.

Table 2: Name of code listings in the CBS remote access environment and in our code (if different)

|  |  |
| --- | --- |
| ***Name of dataset in the CBS remote access environment*** | ***Name in our code* (if different)** |
| **(1)** | **(2)** |
| Gebieden in Nederland\geconverteerde bestanden\131204 GIN 2013V1 | data\_ CIZ\_regions\_clean |
| Gebieden in Nederland\geconverteerde bestanden\131204 GIN 2013V1 | CAK\_regions\_2013 |
| Landen- en nationaliteitscodes\ 120123 omrekentabel land naar herkomstgroepering | Origin\_codes |

* 1. ***External information used***

Finally, we combine the microdata with external information relating to care benefits and their tariffs, in order to compute the monetary value of care entitlements and use.

1. We import information relating to rules for the conversion of entitlements to in-kind care into the receipt of personal budgets (long-term care vouchers). This information was taken from: The figures and information come from:

College voor Zorgverzekeringe.n "[Zelf uw zorg inkopen in 8 stappen 2012. Persoonsgebonden budget AWBZ  
Vergoedingsregeling persoonlijke zorg](https://www.hetvenster.com/Content/DownloadDocs/2012/pgb%20brochure,%20zelf%20zorg%20inkopen,%202012.pdf)”

We coded the information to the following file, which we make available together with the dofiles of the project (using the following DOI: 10.25397/eur.14394158)

Conversion grids\_personal budget.xlsx

1. We import information on the daily tariff paid for institutional care. We use information publicly available on CBS website (StaLine, Monitor Langdurige Zorg). For each care package (ZZP), we compute the average tariff per day spent in institutional care in 2012 as the ratio of total spending (*uitgaven*) to number of days of care (*volume*), based on the a posteriori accounting approach (*nacalculatie*). The data can be found at:

<https://mlzopendata.cbs.nl/#/MLZ/nl/dataset/40026NED/table?dl=9BE4> (consulted on April 1st 2021).

In the code, the tariffs for institutional care are called from a dataset named zmv tarieven nacalculatie.dta.

1. Similarly, we import information on the daily tariff paid for the main types of home care (PV, VP, BG and BH), using the same source.

In the code, the tariffs for home care are called from a dataset named zzv tarieven nacalculatie.dta.

1. Erasmus School of Health Policy & Management, Erasmus Center for Health Economics Rotterdam, Erasmus University Rotterdam. Corresponding author: marianne.tenand.pro[at]gmail.com [↑](#footnote-ref-1)
2. Erasmus School of Health Policy & Management, Erasmus Center for Health Economics Rotterdam, Erasmus University Rotterdam. [↑](#footnote-ref-2)
3. Erasmus School of Health Policy & Management & Erasmus School of Economics Erasmus Center for Health Economics Rotterdam, Erasmus University Rotterdam & Tinbergen Institute. [↑](#footnote-ref-3)