APPENDICES

A Detailed Analysis of the Base Ensemble

Since ESDMA requires a systematic and structured research approach, these techniques and tools were applied –and are reported on in appendix A– in the following order:

1. Preliminary (visual) inspection and if interesting/necessary clustering of the data set:
   (a) Visualization with Envelope Graphs to identify Key Perfomance Indicators (KPIs) that require further analysis
   (b) Visualization with Lines Graphs to inspect the details of the ensemble and characteristics of runs
   (c) If necessary further inspection with Interactive Plots (not shown in this paper)
   (d) If desirable visualization of the 3D KDE plots (showing the kernel density estimate in 3D over time),
   (e) If different modes of behaviors exist, visualizations with dendrogram and cluster plots of interesting clusters obtained by means of time-series clustering.
2. Further analysis if and only the particular key performance indicator and the preliminary inspection of its dynamics require it:
   (a) Identification of undesirable outcomes based on the preliminary (visual) inspection and clustering, or specific directed searches
   (b) Classification of the data set in terms of the end states or the minimum/maximum over each run, or in terms of other characteristics
   (c) Identification of the uncertainties that contribute most to generating this classification by means of the Feature Selection and Random Forest algorithms
   (d) Identification of sets of uncertainties and their values that generate (un)desirable behaviors by means of PRIM and by means of manual testing and visualization of the influence of discrete values of important switch uncertainties (identified in the previous step by means of the Feature Selection and/or Random Forest algorithms).

A.1 Overview: Envelopes & KDEs of the Base Ensembles

Figure 15 shows a multiplot with visualization of the density of 10000 runs for all outcomes of interest. Suppose we would not have detailed knowledge about the structure of the simulation model (which we have of course), then following preliminary conclusions could be drawn from this multiplot:

- The *total population*, the *elderly dependency ratio*, the *male life expectancy*, and *female life expectancy* are positively correlated (not surprising given the structure of the model): hence more attention will be paid to the *elderly dependency ratio* and less to the two other outcomes of interest.
- In the artificial data, there seems to be a positive relationship
  - between the *elderly dependency ratio* & the *health care costs as a fraction of GDP* & the *governmental contribution to AOW as a fraction of GDP* & the *demand for accessible houses* (not surprising),
  - and between the *annual GDP* and the *fraction labour scarcity* (not surprising given the structure of the model).
- In the artificial data, there seems to be an inverse relationship between

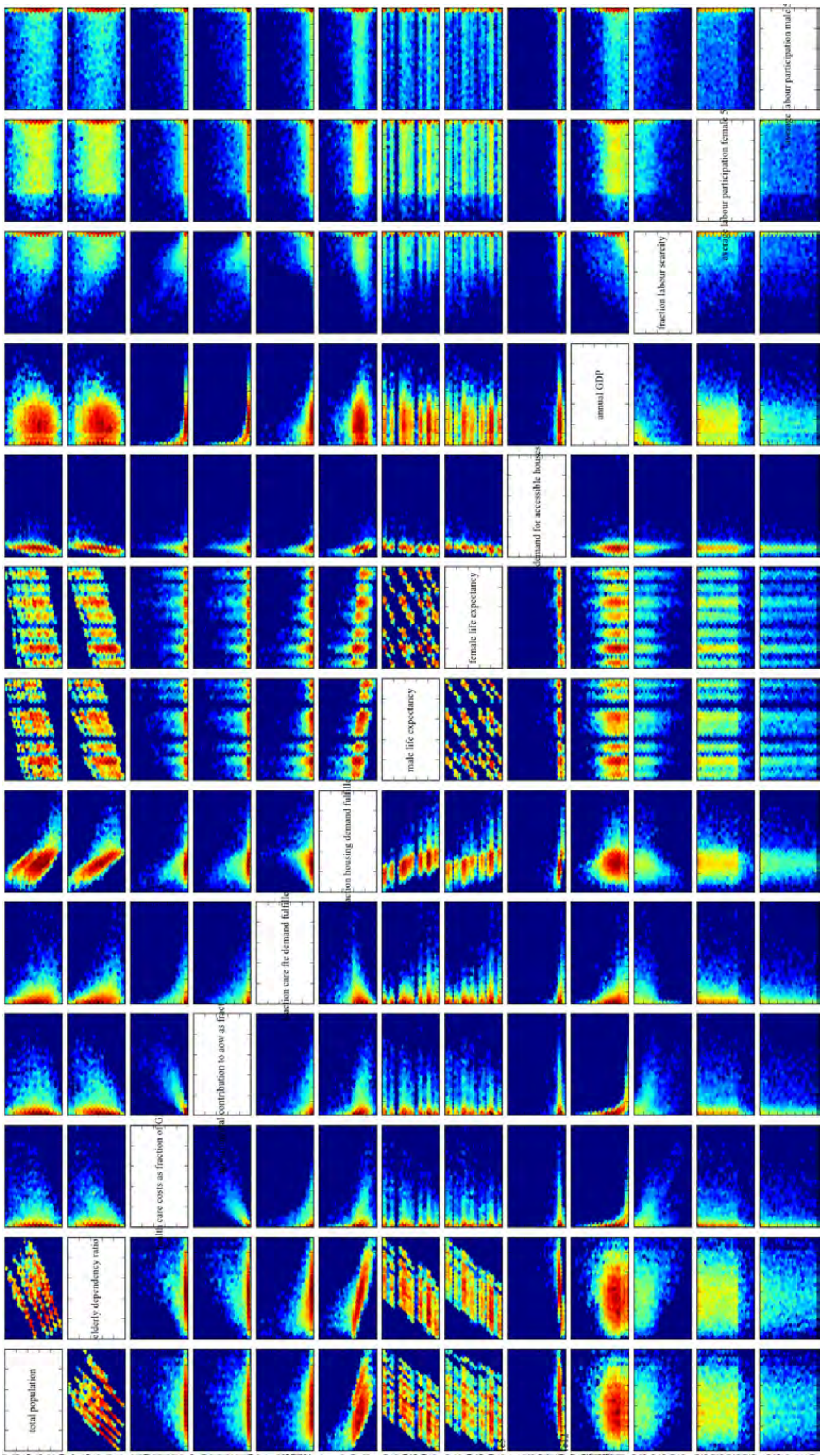


Figure 15: Multiplot with indication of density of 10000 runs for all outcomes of interest

  - the *health care costs as a fraction of GDP* and the *fraction care FTE demand fulfilled*,
  - the *annual GDP* and the *health care costs as a fraction of GDP*,
  - the *annual GDP* and the *governmental contribution to AOW as a fraction of GDP*,
  - the *demand for accessible houses* and the *fraction housing demand fulfilled* & the *fraction care FTE demand fulfilled*.

- In the artificial data, there seems to be a negative relationship between
  - the *fraction care FTE demand fulfilled* and the *governmental contribution to AOW as a fraction of GDP* & *elderly dependency ratio*,
  - the *fraction housing demand fulfilled* and the couple *total population-elderly dependency ratio* (not surprising given the causal relations in the model),
  - the *fraction housing demand fulfilled* and the couple *health care costs as a fraction of GDP-governmental contribution to AOW as a fraction of GDP* (purely correlational).
- There is one outlier in 10000 runs for the *demand for accessible houses* (to be analyzed, not to be removed).
- Several variables seem to be uncorrelated: they are either causally separated and/or require separate analysis and monitoring.

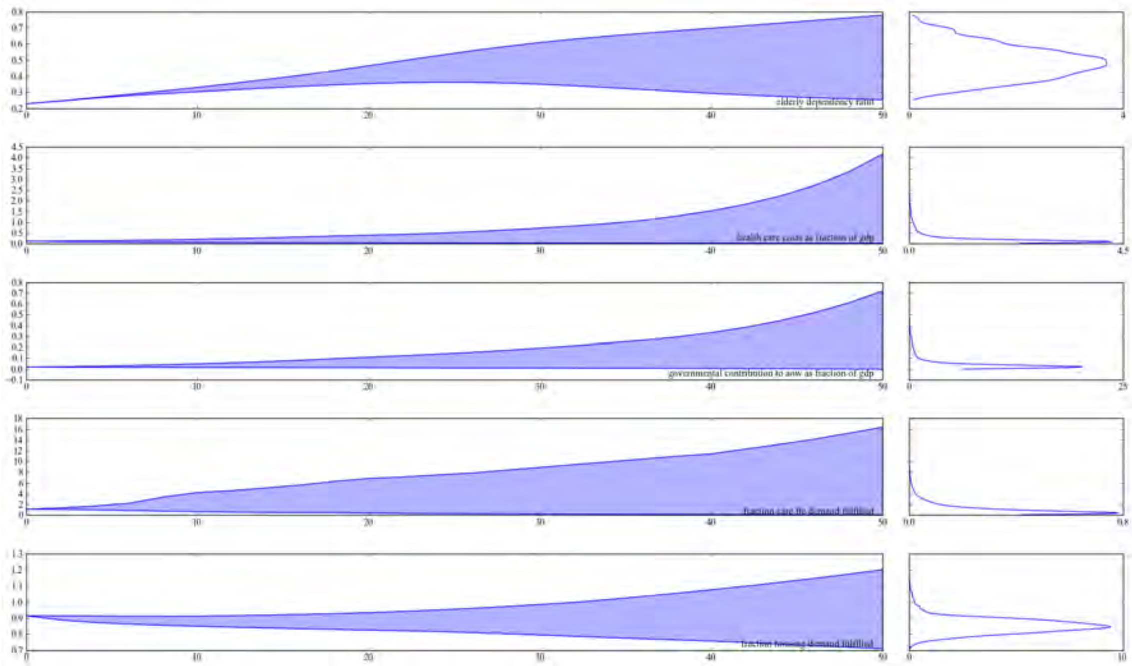


Figure 16: Envelopes of 10000 runs for –topdown– the *elderly dependency ratio*, *health care costs as a fraction of GDP*, *governmental contribution to AOW as a fraction of GDP*, *fraction care FTE demand fulfilled*, *fraction housing demand fulfilled* (KDEs are not to be interpreted as PDFs)

Figure 16 shows the envelopes and end state KDEs of the core KPIs, namely the *elderly dependency ratio*, *health care costs as a fraction of GDP*, *governmental contribution to AOW as a fraction of GDP*, *fraction care FTE demand fulfilled*, and the *fraction housing demand fulfilled*. Note: the Kernel Density Estimates (KDEs) (distributions on the right hand side) are to be interpreted in an exploratory way, NOT in a probabilistic way ('these end states are plausible', not 'these are the probabilities of the future end states'). Visual inspection of these graphs leads to following observations and/or recommendations for further research:

- *elderly dependency ratio*: evolves from 23% to values between 25% and 80% – requires some further investigation (see subsection A.3)
- *health care costs as a fraction of GDP*: requires –given the discrepancy between envelope and KDE– visualization with a lines graph (see Figure 25 on page 30) and further exploration (see subsection A.5)
- *governmental contribution to AOW as a fraction of GDP*: requires –given the discrepancy between envelope and KDE– visualization with a lines graph (see Figure 28 on page 33) and further exploration (see subsection A.6)
- *fraction care FTE demand fulfilled*: requires –given the discrepancy between envelope and KDE– visualization with a lines graph (see Figure 31 on page 36) and further exploration (see subsection A.7)
- *fraction housing demand fulfilled*: requires visualization with a lines graph (see Figure 36 on page 42) and further exploration (see subsection A.9)

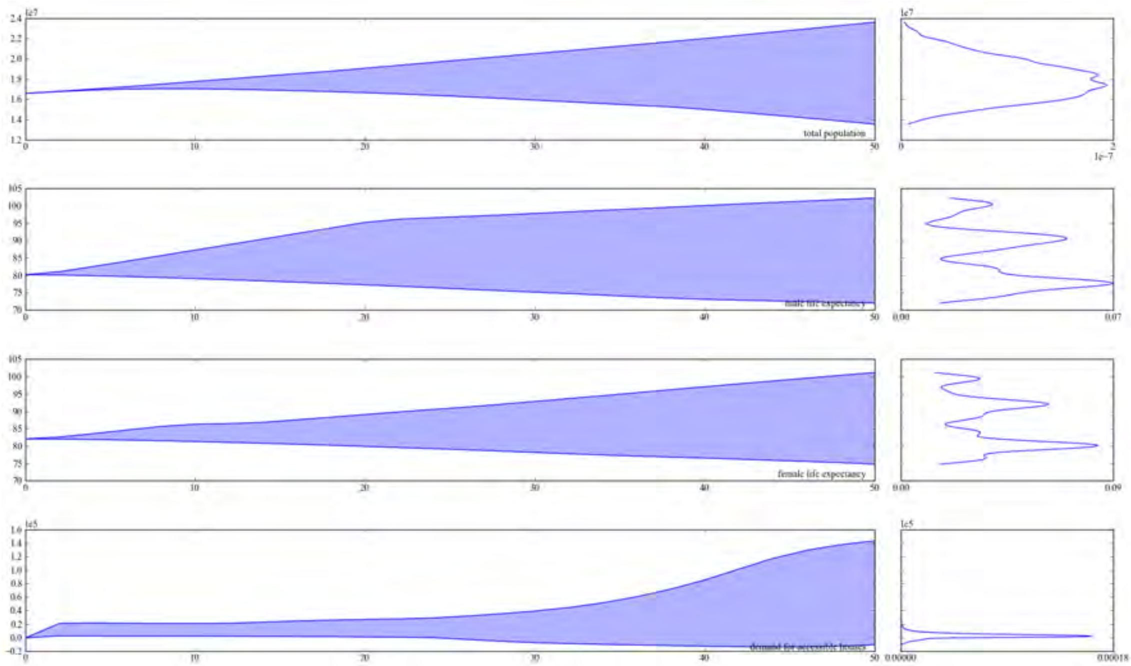


Figure 17: Envelopes of 10000 runs for –topdown– the *total population*, *male life expectancy*, *female life expectancy*, *demand for accessible houses* (KDEs are not to be interpreted as probabilities)

Figure 17 shows the envelopes and end state KDEs of the *total population*, *male life expectancy*, *female life expectancy*, and the *demand for accessible houses*. Visual inspection of these graphs leads to following observations and/or recommendations for further research:

- *total population*: evolves to end state values between (little less than) 14 million and 24 million. Some further exploration may be justified because of (i) the interest of many for this output indicator, and (ii) the causal influence of the population sector on the other sectors of the model (see subsection A.2).
- *male life expectancy*: the remarkable KDE is mainly caused by 6 different evolutions of the male life expectancy provided by workshop participants and activated by the *life expectancy male trend main switch*. This switch-forced evolution is one of the many reasons for refraining from probabilistic interpretations of the outcomes of these analyses.

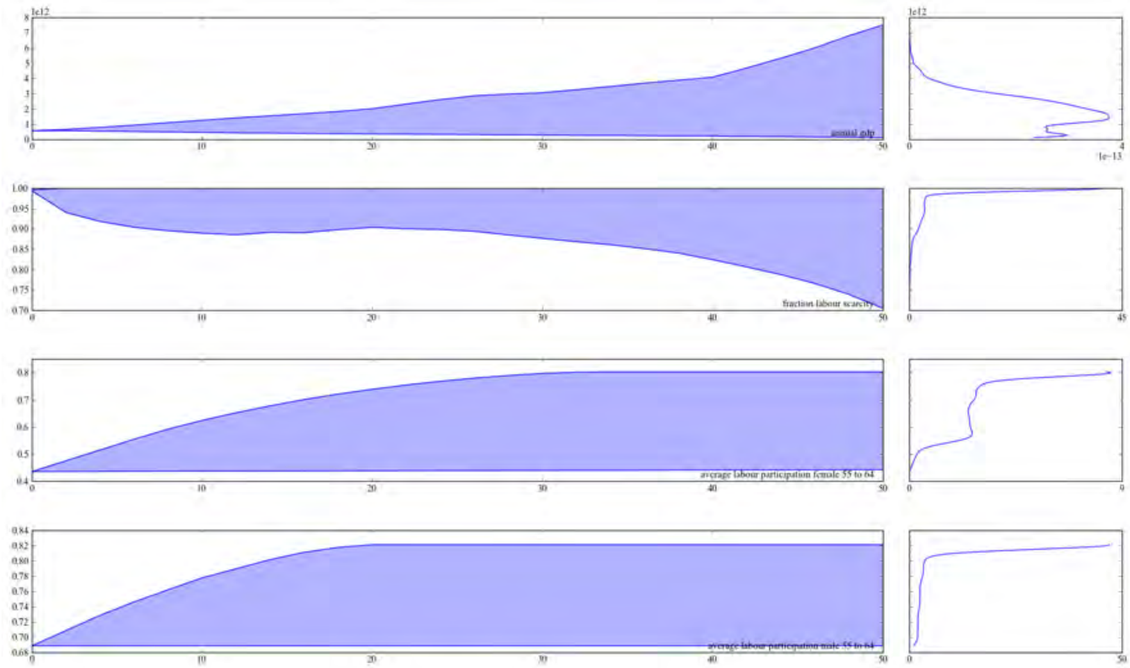


Figure 18: Envelopes of 10000 runs for –topdown– the *annual GDP*, *fraction labour scarcity*, *average labour participation female 55 to 64*, *average labour participation male 55 to 64* (KDEs are not to be interpreted as probabilities)

- *female life expectancy*: the remarkable KDE is mainly caused by 6 different evolutions of the female life expectancy provided by workshop participants and activated by the *life expectancy female trend main switch*. This switch-forced evolution is one of the many reasons for refraining from probabilistic interpretations of the outcomes of these analyses.
- *demand for accessible houses*: requires visualization with a lines graph (see Figure 39 on page 44) and further exploration (see subsection A.10)

Figure 18 shows the envelopes and end state KDEs of the *annual GDP*, *fraction labour scarcity*, *average labour participation female 55 to 64*, and the *average labour participation male 55 to 64*. Visual inspection of these graphs leads to following observations and/or further research:

- *annual GDP*: further exploration may be justified because of the general interest in this output indicator and its influence on the rest of the model (see subsection A.4)
- *fraction labour scarcity*: requires some further exploration (see subsection A.8)
- *average labour participation female 55 to 64*: requires some further exploration (see subsection A.11)
- *average labour participation male 55 to 64*: requires some further exploration (see subsection A.12)

A.2 Superficial investigation of the *total population*

Figure 19 shows the *total population* of each of the 10000 runs of the ensemble. The logarithmic 3D KDE plot in figure 20 shows how the population increases at first but then fans out in all directions. Although probabilistic interpretations should be avoided because they cannot be supported by the sampling method and inputs, it remains interesting to see whether and how many runs end up

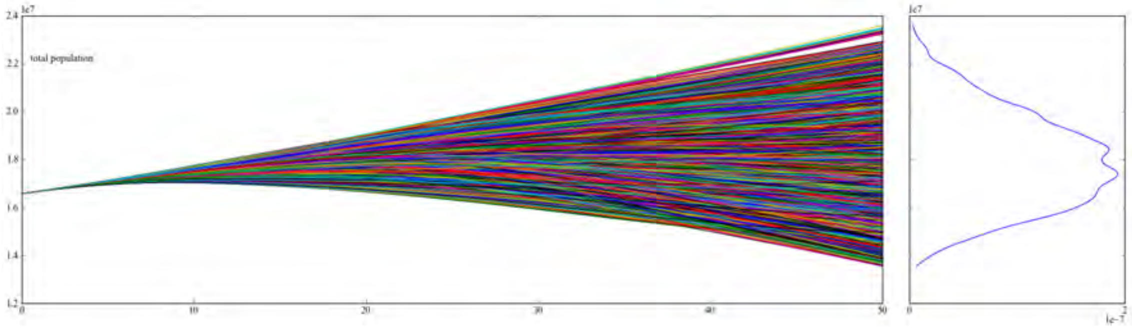


Figure 19: Lines for 10000 runs of the *total population*

in which part of the output space. Out of these 10000 runs, 1697 runs have an end state value below 16 million, 3585 runs have an end state value between 16 and 18 million, 3172 runs have an end state value between 18 and 20 million, 1358 runs have an end state value between 20 and 22 million, and 188 runs have an end state value above 22 million.

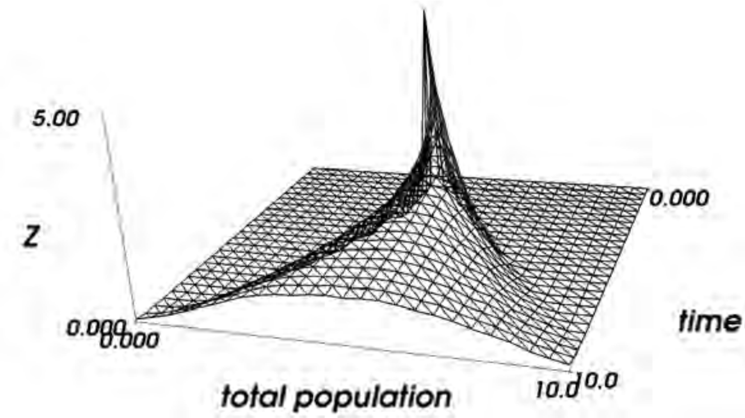


Figure 20: Logarithmic 3D plot of 10000 runs for the total population

The *total population* is not analyzed in depth, because the aim of the paper is to explore the effects of demographic aging, not to predict or analyze demographic evolutions regarding the Dutch population.

A.3 Further investigation of the *elderly dependency ratio*

A.3.1 Visualization and Time-series Clustering

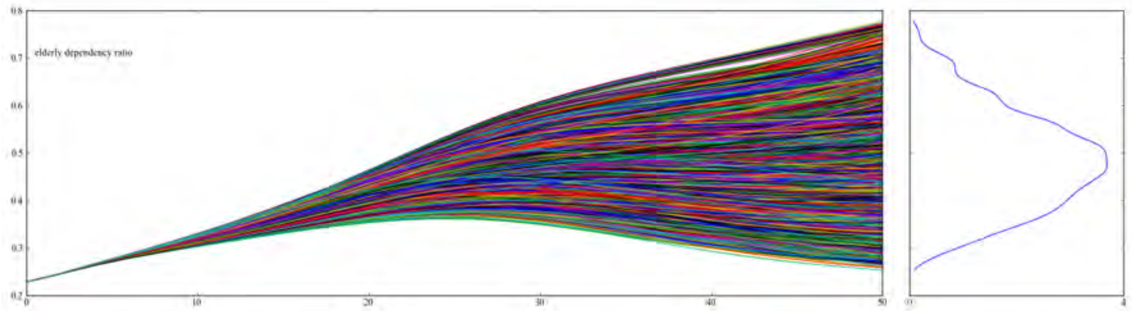


Figure 21: 10000 runs for the *elderly dependency ratio*

Figure 21 shows the trajectory of the *elderly dependency ratio* of each of the 10000 runs of the ensemble: the ensemble is rather homogenous over the first 25 years, but ambiguous thereafter. Different types of behaviors are indeed generated: Figure 22 shows the dendrogram generated by the clustering algorithm (forced to 4 clusters) for the 1000 runs ensemble.

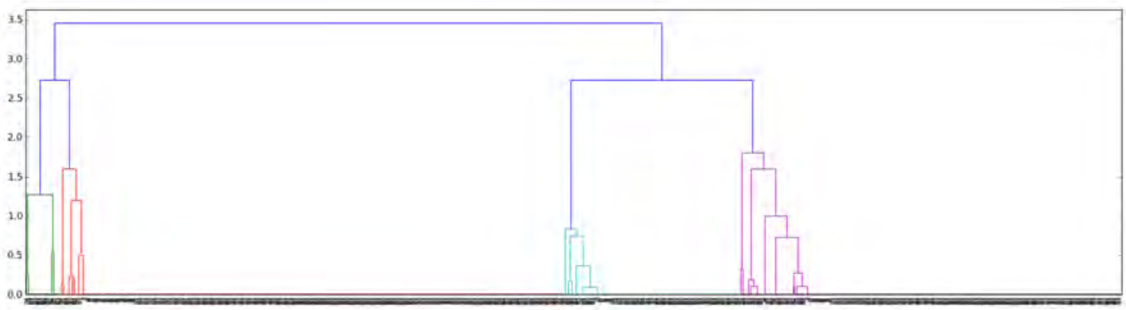
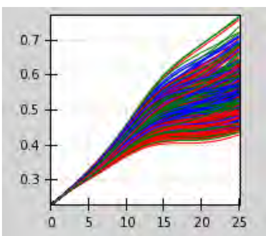
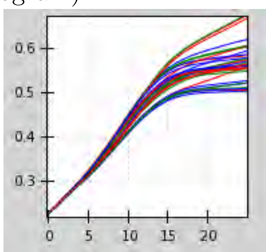


Figure 22: Cluster Dendogram for 1000 runs for the *elderly dependency ratio*

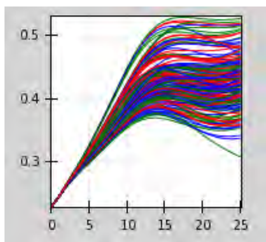
The figures below show the different types of behaviors if forced to only 4 clusters (∼ dark blue top-most spit in the dendrogram):



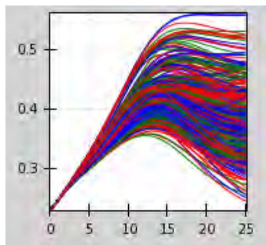
Cluster 1



Cluster 2



Cluster 3



Cluster 4

A.3.2 Reference Outcomes, Classification, Feature Selection & Random Forest, and PRIM

1. **Classification of outcomes (Classify):** Each of the 10000 runs is classified according to its maximum value over the run in one of following 4 classes: (1) an *elderly dependency ratio* below 30%, (2) an *elderly dependency ratio* between 30% and 40%, (3) an *elderly dependency ratio* between 40% and 50%, and (4) an *elderly dependency ratio* above 50%.

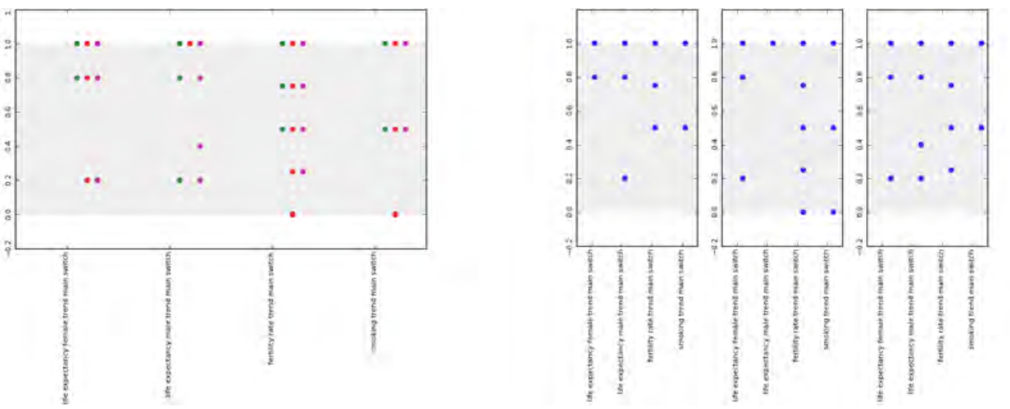
Out of these 10000 runs, 0 have a maximum *elderly dependency ratio* below 30%, 554 have a maximum *elderly dependency ratio* between 30% and 40%, 4912 have a maximum *elderly dependency ratio* between 40% and 50%, and 4534 have a maximum *elderly dependency ratio* above 50%.

2. **Identification of the most important uncertainties (Feature Selection and Random Forest):** The most important uncertainties obtained with the Random Forest and Feature Selection algorithms on the 10000 runs set are the life expectancies, the main trend of the fertility rate, and health related evolutions (smoking and obesity).

3. **Identification of multidimensional boxes in the uncertainty space that cause high outcomes of the *elderly dependency ratio* (PRIM):**

   A PRIM0.3b on the maximum value of each run above 55% with a minimal mass of 5% of the runs and a threshold of 60% positive matches within each box results in 3 boxes with a total of 1831 cases out of 2648 bad cases in total (i.e. with a maximum *elderly dependency ratio* above 55%) out of 10000 runs.

	mass	mean
found box 0:	0,0723	1,00
found box 1:	0,0594	0,94
found box 2:	0,0855	0,64
rest box :	0,7828	0,10



In other words, there are 3 areas in this multidimensional uncertainty space, each containing at least 50 runs and with at least 60% of these runs having a maximum elderly dependency ratio above 55%. The first box even has a 100% positive match ratio: all cases in that box generate a future with a maximum *elderly dependency ratio* above 55%.

The table below –which displays the individual effect of switch values of important categorical uncertainties, more precisely the percentage matches with undesirable behaviour for the three most important switch uncertainties– shows that:

- trend 6 of the *male life expectancy* has a negative influence on the elderly dependency ratio: more than 62% of the 1666 cases with a switch value equal to 6 generate a future with a maximum *elderly dependency ratio* above 55%.
- trend 6 of the *female life expectancy* has a negative influence on the elderly dependency ratio: more than 68% of the 1666 cases with a switch value equal to 6 generate a future with a maximum *elderly dependency ratio* above 55%.
- The third switch uncertainty is weaker and only forces the outcomes to some extent for lower evolutions of the fertility rate.

elderly dependency ratio > 0.55	Avg Nr Runs	matches with undesirable behaviour (EDR>55%)										TOTAL NR UN-
main switch uncertainties:	per switch	1	2	3	4	5	6	7	8	9	10	DESIRABLE RUNS
life expectancy male trend main switch	1666	5,94%	33,61%	12,06%	4,20%	40,76%	62,36%					2648
life expectancy female trend main switch	1666	4,38%	33,73%	9,06%	4,44%	38,90%	68,43%					2648
fertility rate trend main switch	2000	9,50%	17,85%	25,25%	36,70%	43,10%						2648

A.3.3 Conclusions

Regarding the *elderly dependency ratio*, the following preliminary conclusions could be drawn:

- The combination of this model and these uncertainty ranges, generates many scenarios with rather high *elderly dependency ratio*s, either peaking in 25 years or rising at a slower rate after 25 years. 25 years from now, the average babyboomer reaches –first– his –and then her– life expectancy. Following three different modes –caused by different evolutions of the third uncertainty, the fertility rate– occur after this breakpoint: S-shaped decrease, stabilization or linear to slightly exponential growth.

- Major uncertainties determining the maximum *elderly dependency ratio* are –not surprisingly– the evolutions of the male and female life expectancy, of the fertility rate, and of smoking.

- Several subspaces of the uncertainty space have a tendency to generate high *elderly dependency ratio*s. Although a high *elderly dependency ratio* is not a problem in itself, it may in many futures cause serious problems on other KPIs.

A.4 Superficial investigation of evolutions of the *annual GDP*

This output indicator is not analyzed in detail because GDP is not the focus and aim of this paper. The only reason it deserves a subsection dedicated to it, is that it has a direct influence on two key output indicators –the *health care costs as a fraction of GDP* and the *governmental contribution to AOW as a fraction of GDP*– and an indirect influence on the system and other key output indicators.

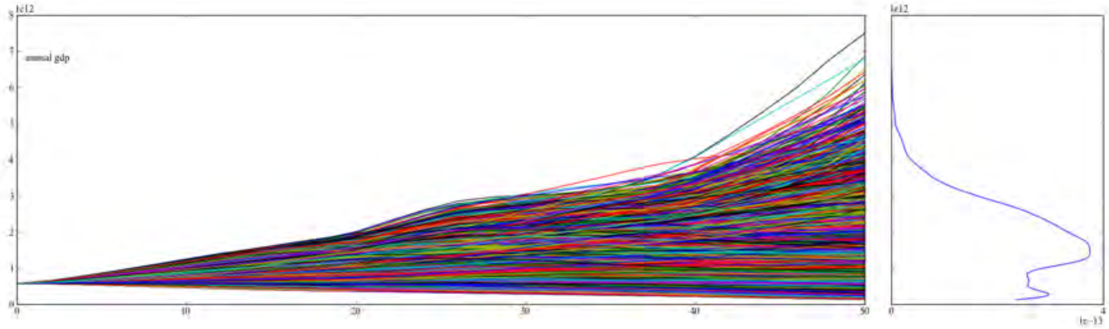


Figure 23: Lines of 10000 runs for the annual GDP

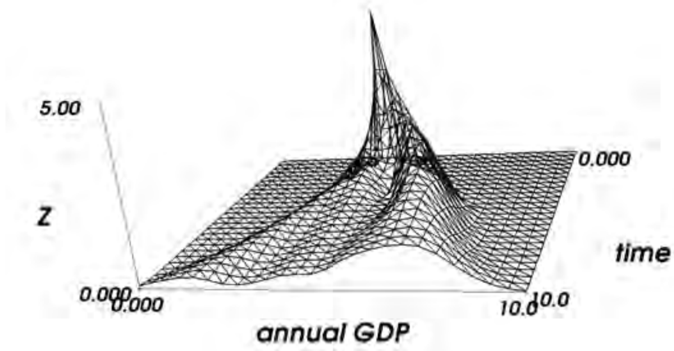


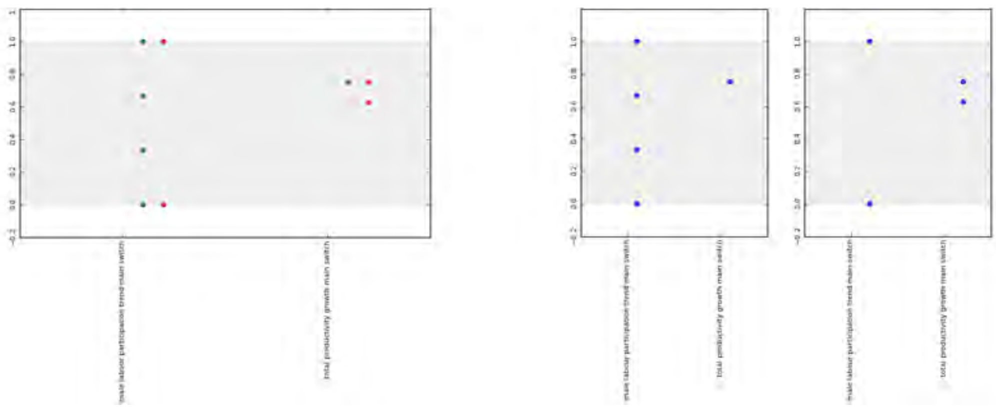
Figure 24: 3D plot of 10000 runs for the annual GDP

Figures 23 and 24 show the trajectories and log 3D KDE plots of 10000 runs of the *annual GDP*. In 2010, the annual GDP was in the order of €570 billion per year. Hence following classification on the end state values in 2060 is made: (1) below €570 billion per year (below the 2010 GDP), (2) between €570 and €1140 billion per year, (3) between €1140 and €2280 billion per year, (4) between €2280 and €4560 billion per year, and (5) above €4560 billion per year.

Out of 10000 runs, 1501 runs have an annual GDP end state value at the end of a 50 years run time below €570 billion per year, 1561 runs have an end state value between €570 and €1140 billion per year, 4096 runs have an end state value between €1140 and €2280 billion per year, 2734 runs have an end state value between €2280 billion and €4560 billion per year, and 108 runs have an end state value above €4560 billion per year.

A PRIM0.3b on end value of each run below the 2010 GDP with a minimal mass of 5% of the runs and a threshold of 60% positive matches within each box results in 2 boxes with a total of 1473 cases out of 1501 bad cases in total out of 10000 runs.

	mass	mean
found box 0 :	0,1112	1,00
found box 1 :	0,0544	0,66
rest box :	0,8344	0,00



Again, the first box has a 100% positive match ratio and contains more than 11% if all runs. The main uncertainty driving this classification is –not surprisingly– the *total productivity growth*. All cases with the seventh lookup function (∼ decreasing productivity) have end state values below the 2010 GDP (see following table).

annual GDP < 570000000000	Avg Nr Runs				matches with undesirable behaviour						TOTAL NR UN-
main switch uncertainties:	per switch	1	2	3	4	5	6	7	8	9	10	DESIRABLE RUNS
total productivity growth main switch	1112	0,00%	0,00%	0,00%	0,00%	0,00%	34,98%	100,00%				1501
total productivity growth in care main switch	2000	14,85%	15,20%	15,40%	14,55%	15,05%						1501

The remainder of the runs ending with values below the 2010 GDP, use lookup function 6 (∼ constant productivity at the 2010 level) or 7 and are characterized by decreases in the male labour participation. The second most important uncertainty as suggested by the Feature Selection and Random Forest algorithms, the *total productivity growth in care*, adds little more than noise.

A.5 Further investigation of the *health care costs as a fraction of GDP*

A.5.1 Visualization and Time-series Clustering

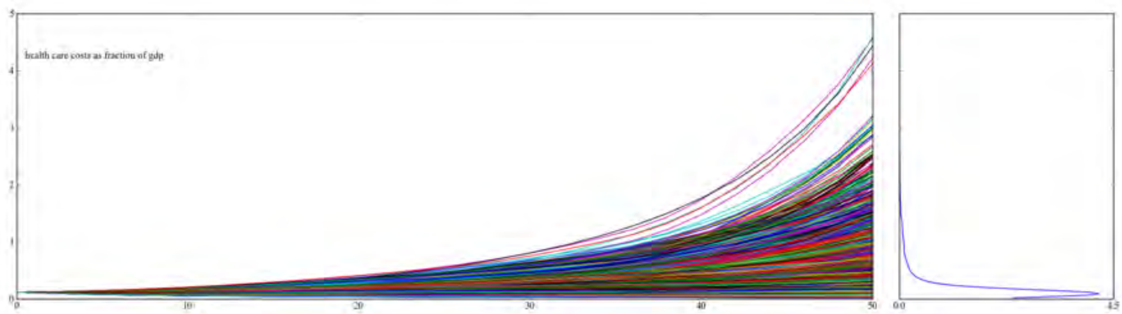


Figure 25: Lines for 10000 runs of the *health care costs as a fraction of GDP*

Figure 25 shows the *health care costs as a fraction of GDP* of an ensemble of 10000 runs. Although the kernel density estimates of the higher values of the *health care costs as a fraction of GDP* are extremely low, especially above 2, these very unlikely outcomes should be analyzed because they represent a serious risk (threat to the financial sustainability of health care). This is confirmed by the log 3D KDE plot of the *health care costs as a fraction of GDP* in Figure 26.

A.5.2 Reference Outcomes, Classification, Feature Selection & Random Forest, and PRIM

1. **Identification of undesirable outcomes:** In 2009, the *health care costs as a fraction of GDP* amounted to 10.1% of GDP (CBS 2009 – see (Logtens 2011) site links 12 and 34)). *Health care costs as a fraction of GDP* of more than 100% are impossible.

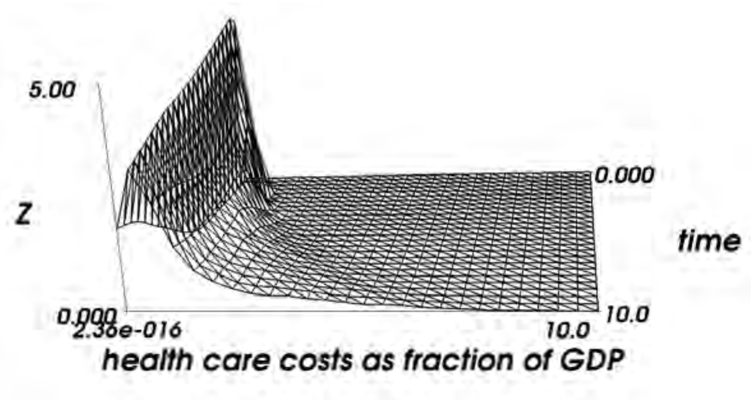


Figure 26: Logarithmic 3D KDE of 10000 runs for the *health care costs as fraction of GDP*

2. **Classification of desirable versus undesirable outcomes (Classify):** The set is classified on the maximum value per run in 4 arbitrary classes based in the 2010 value and the outcome space generated: (1) between 0% and 10%, (2) between 10% and 25%, (3) between 25% and 100%, and (4) above 100%. The idea behind this classification is that *health care costs as a fraction of GDP* between 0% and 10% are better than today, costs between 10% and 25% are worse than today but still feasible and acceptable, costs between 25% and 100% are unacceptable, and costs above 100% are infeasible.

   Out of these 10000 runs, not a single run has a *health care costs as a fraction of GDP* between 0% and 10%, 7858 runs have a *health care costs as a fraction of GDP* between 10% and 25%, 1701 runs have a *health care costs as a fraction of GDP* between 25% and 100%, and 441 runs have a *health care costs as a fraction of GDP* above 100%. Although these values should not be interpreted probabilistically, the last two values are disturbingly high.

3. **Identification of the most important uncertainties (Feature Selection and Random Forest):** The uncertainties in terms of importance obtained with the Random Forest and Feature Selection algorithms on 10000 runs using this classification function on the maximum value per run are: the total productivity growth trends, and trends of the life expectancies.

   Figure 27 shows envelopes for different values of the most important uncertainty (total productivity growth main switch). This Figure shows that the two lowest productivity growth scenarios (constant productivity and decreasing productivity) are almost fully responsible for all runs with *health care costs as a fraction of GDP* higher than 50%, and to a very large extent for those above 25%. Figure 27 is, in this sense, very similar to Figure 30 showing the influence of *total productivity* on the *governmental contribution to AOW as a fraction of GDP*.

4. **Identification of multidimensional boxes in the uncertainty space that cause the undesirable outcomes (PRIM):** Here, an arbitrary value of 25% of *health care costs as a fraction of GDP* (vs. Tom's value of 40%) is assumed to be undesirable.

   A PRIM0.3b on the maximum value of the health care costs as a fraction of GDP for each run above 0.25 with a minimal mass of 5% of the runs and a threshold of 60% positive matches within each box results in 2 boxes with a total of 1571 cases out of 2142 bad cases in total out of 10000 runs.

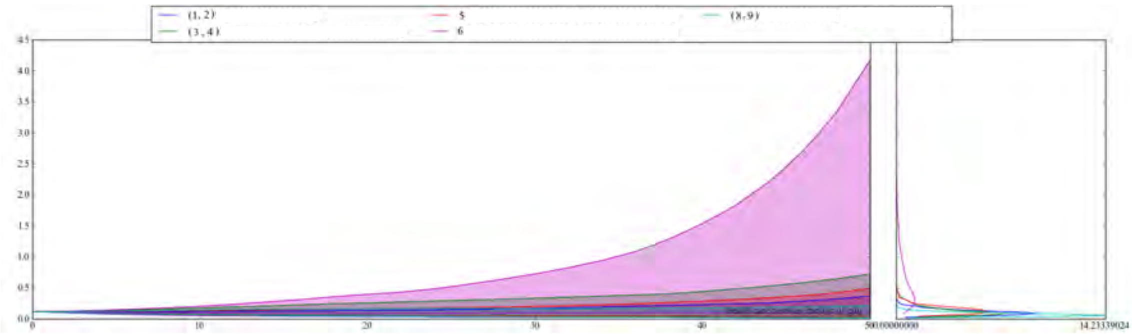
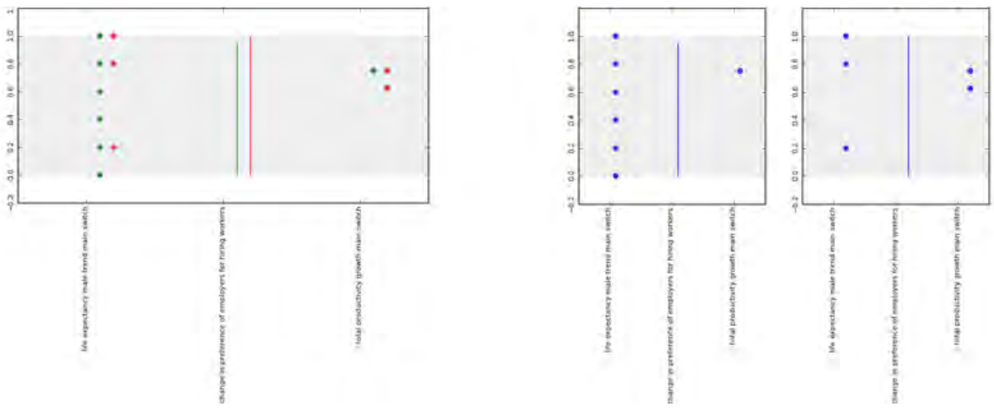


Figure 27: Envelopes of the *health care costs as a fraction of GDP* for different values of the most important uncertainty (total productivity growth main switch)

	mass	mean
found box 0:	0,1054	1,00
found box 1:	0,0592	0,87
rest box :	0,8354	0,07



The first box contains more than 10% of the cases and has a 100% match ratio (!), the second box contains more than 5% of all cases and has a positive match ratio of 87%. The first box is characterized by a decreasing productivity and a change in the preferences of employers for hiring workers. The second box consists of cases without productivity increases and seriously increasing male life expectancies.

The following table displays the percentage matches with undesirable behavior for each of the most important switch uncertainties. Not surprisingly, trend 7 of the *average labour productivity* has a very strong negative influence on the *health care costs as a fraction of GDP*: more than 99% of the 1112 cases with a switch value equal to 7 generate a future with a maximum *health care costs as a fraction of GDP* above 25%. The same is true but to a lesser extent for trend 6 of the *average labor productivity* which leads to more than 62% of the 1112 cases with that switch rising at some point above *health care costs as a fraction of GDP* of 25%. The second and third switch uncertainties are far less determinant.

health care costs as fraction of GDP > 0.25	Avg Nr Runs	matches with undesirable behaviour										TOTAL NR UN-
main switch uncertainties:	per switch	1	2	3	4	5	6	7	8	9	10	DESIRABLE RUNS
total productivity growth main switch	1112	0,09%	1,71%	0,09%	20,14%	8,36%	62,50%	99,73%				2142
life expectancy male trend main switch	1666	15,79%	21,73%	19,27%	14,95%	23,05%	33,79%					2142
life expectancy female trend main switch	1666	17,41%	22,75%	19,03%	15,13%	24,85%	29,41%					2142

A.5.3 Conclusions:

Regarding the *health care costs as a fraction of GDP*, the following conclusions could be drawn:

- *Health care costs as a fraction of GDP* will not go down and may increase a lot, even to unsustainable levels. Unsustainable *health care costs as a fraction of GDP* are mainly caused by insufficient productivity growth in the Dutch economy at large.
- The combination of this model and these uncertainties, generates two major boxes with more than 60% of the runs with a *health care costs as a fraction of GDP* above 25%. The two major boxes are caused by constant or decreasing total productivity of the Dutch economy at large. For this KPI it is therefore recommended to stimulate productivity increases in the Dutch economy at large. The other boxes could –if further analyzed– be used for policy design or testing.
- Given the main uncertainties, it looks as though the *total productivity growth* (and to a lesser extent the productivity growth in care) needs to accommodate increases in life expectancies to keep health care costs of GDP sustainable.

A.6 Further investigation of the *governmental contribution to AOW as a fraction of GDP*

A.6.1 Visualization and Time-series Clustering

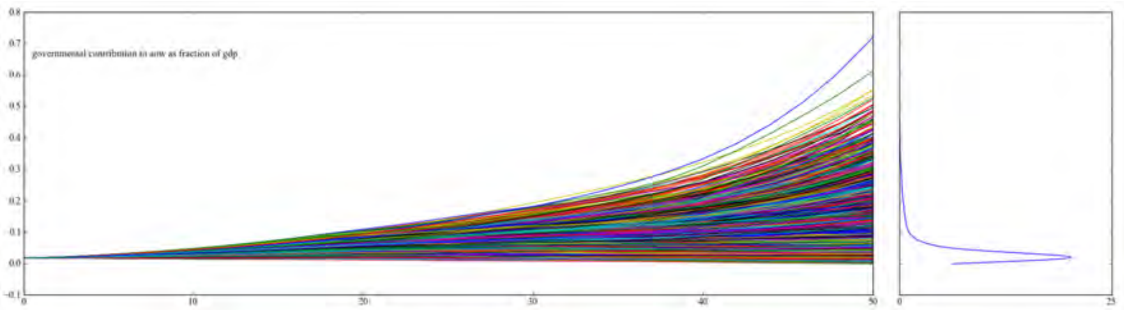


Figure 28: Lines for 10000 runs of the *governmental contribution to AOW as a fraction of GDP*

Figure 28 shows the *governmental contribution to AOW as a fraction of GDP* of each of the 10000 runs of the ensemble. The end state KDE in Figure 28 seems to suggest the general tendency of the *governmental contribution to AOW as fraction of GDP* remains within reasonable bounds, which is confirmed by the logarithmic 3D KDE graph in Figure 29.

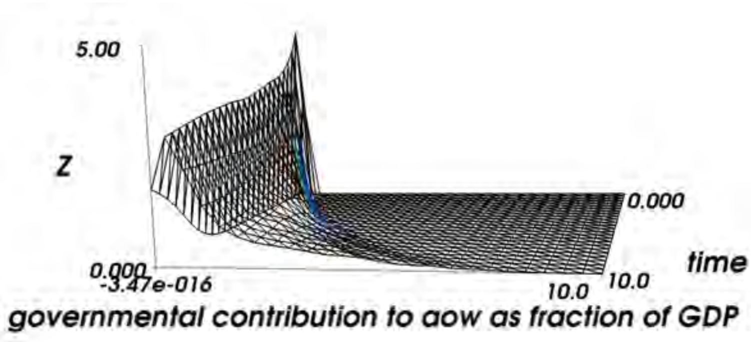


Figure 29: 3D KDE graph of 10000 runs for the *governmental contribution to AOW as a fraction of GDP*

A.6.2 Reference Outcomes, Classification, Feature Selection & Random Forest, and PRIM

1. **Identification of undesirable outcomes:** The *governmental contribution to AOW as a fraction of GDP* amounted in 2009 to 1.7% of GDP.

2. **Classification of desirable versus undesirable outcomes (Classify):** Here, following four arbitrary classes are used to classify 10000 runs on their maximum over the run: (1) a *governmental contribution to AOW as a fraction of GDP* of less than 2%, (2) between 2% and 10%, (3) between 10% and 20%, and (4) above 20%.

   Out of 10000 runs, 904 runs have a *governmental contribution to AOW as a fraction of GDP* of less than 2%, 7912 runs between 2% and 10%, 693 runs between 10% and 20%, and 491 runs above 20%.

3. **Identification of the most important uncertainties (Feature Selection and Random Forest):**

   The two most important uncertainties forcing this classification are the total productivity growth and the fraction of average wage increase to increase in AOW. Figure 30 shows envelopes for different values of the most important uncertainty (total productivity growth main switch). The next three uncertainties show a difference too, but less pronounced. Figure 30 is very similar, in this sense, to Figure 27 showing the influence of *total productivity* on the *health care costs as a fraction of GDP*.

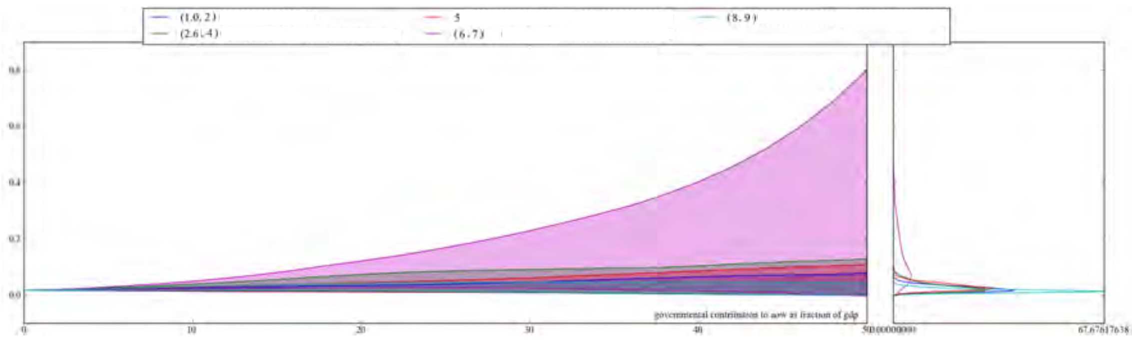
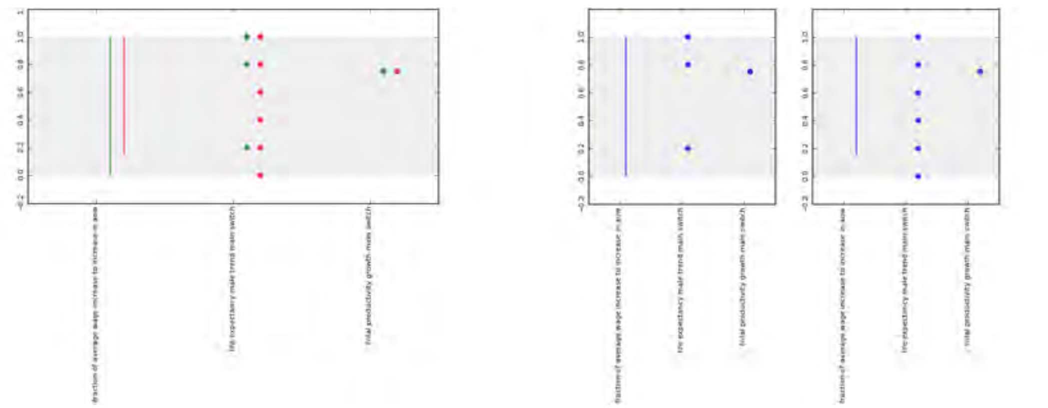


Figure 30: Envelopes for 10000 runs of the *governmental contribution to AOW as a fraction of GDP* for different values of the main uncertainty (*total productivity growth main switch*)

4. **Identification of multidimensional boxes in the uncertainty space that cause undesirable outcomes (PRIM):** A *governmental contribution to AOW as a fraction of GDP* of more than 10% at any point in time is considered highly undesirable here (arbitrary choice).

   A PRIM0.3b on the maximum value of the *governmental contribution to aow as fraction of GDP* of each run greater than 0.1 with a minimal mass of 5% of the runs and a threshold of 60% positive matches within each box results in 2 boxes with a total of 931 cases out of 1184 bad cases.

	mass	mean
found box 0:	0,0520	0,98
found box 1:	0,0508	0,83
rest box :	0,8972	0,03



Highly undesirable values of the *governmental contribution to AOW as fraction of GDP* are to a large extent caused by the seventh *total productivity growth main switch* (i.e. decreasing productivity), either together with high-end evolutions of the male life expectancy or a high fraction of wage increase to increase in aow.

The following table displays the percentage matches with undesirable behavior for the most important switch uncertainty. Trend 7 of the *average labor productivity* (∼ decreasing productivity) has a dramatic influence on the *governmental contribution to AOW as fraction of GDP*: more than 87% of the 1112 cases with a switch value equal to 7 generate a future with a maximum *governmental contribution to AOW as fraction of GDP* above 10%. These cases constitute 82% of all the 1184 cases with a *governmental contribution to AOW as fraction of GDP* above 10%.

governmental contribution to aow as fraction of GDP >0.1	Avg Nr Runs	matches with undesirable behaviour (EDR>55%)										TOTAL NR UN-
main switch uncertainties:	per switch	1	2	3	4	5	6	7	8	9	10	DESIRABLE RUNS
total productivity growth main switch	1112	0,00%	0,00%	0,00%	0,54%	0,00%	18,79%	87,14%				1184

A.6.3 Conclusions:

Regarding the *governmental contribution to AOW as a fraction of GDP*, the following conclusions could be drawn:

- The combination of this model and these uncertainty ranges, generates close to 700 out of 10000 cases with *governmental contribution to AOW as a fraction of GDP* between 10% and 20%, and close to 500 cases with *governmental contribution to AOW as a fraction of GDP* above 20%. The other cases remain below 10%. There are also cases in which the *governmental contribution to AOW as a fraction of GDP* decreases.
- The uncertainty determining the maximum value of *governmental contribution to AOW as a fraction of GDP* most is the *total productivity growth*: almost all undesirable cases are the result of decreasing or constant productivity. The second
- *Productivity growth* could thus both be the cause for problems and the cure for problems related to the *governmental contribution to AOW as a fraction of GDP* too.

A.7 Further investigation of the *fraction care FTE demand fulfilled*

A.7.1 Visualization and Time-series Clustering

Figure 31 shows the lines graph of the *fraction care FTE demand fulfilled*, and Figure 32 a logarithmic 3D KDE plot of the 10000 runs ensemble. Without additional policies (e.g. regarding migration or productivity growth) there is a very high risk of labour scarcity in care.

Figure 33 shows the dendrogram of the 1000 runs ensemble. The figures below show the different types of behaviors if a split is enforced at hight 5:

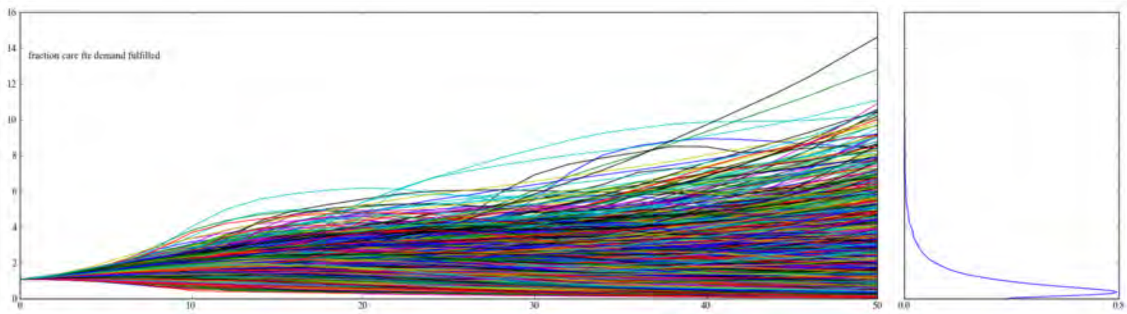


Figure 31: Lines for 10000 runs of the *fraction care FTE demand fulfilled*

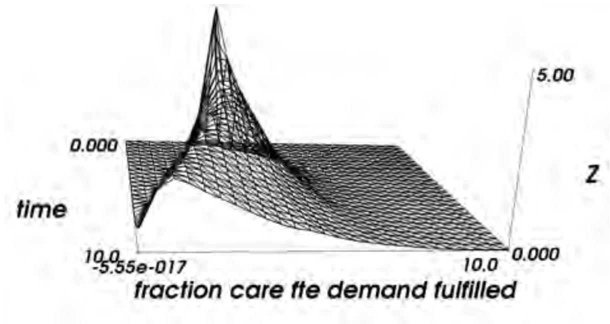
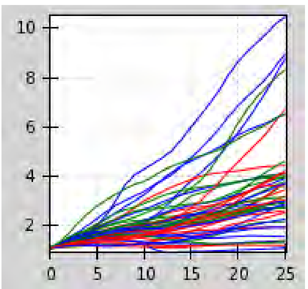
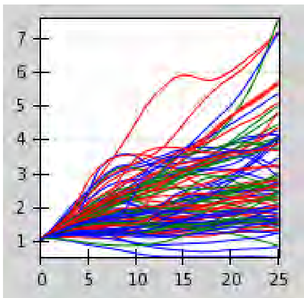


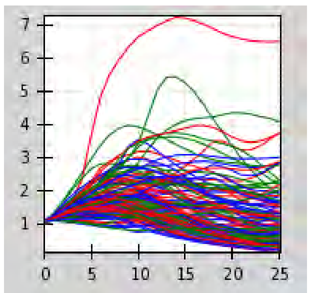
Figure 32: 3D 10000 runs *fraction care FTE demand fulfilled*



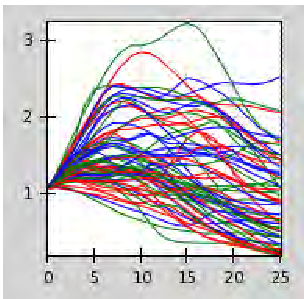
Cluster 1



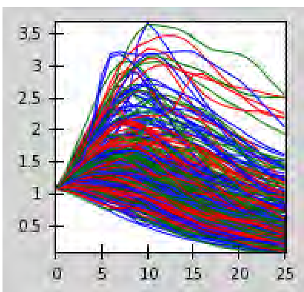
Cluster 2



Cluster 3



Cluster 4



Cluster 5

A.7.2 Reference Outcomes, Classification, Feature Selection & Random Forest, and PRIM

1. **Identification of undesirable outcomes:** Shortages of labor in care are undesirable. Abundance of labor is –from a care point of view (not from a societal point of view)– to be preferred. Hence, the minimum value of *fraction care FTE demand fulfilled* should remain

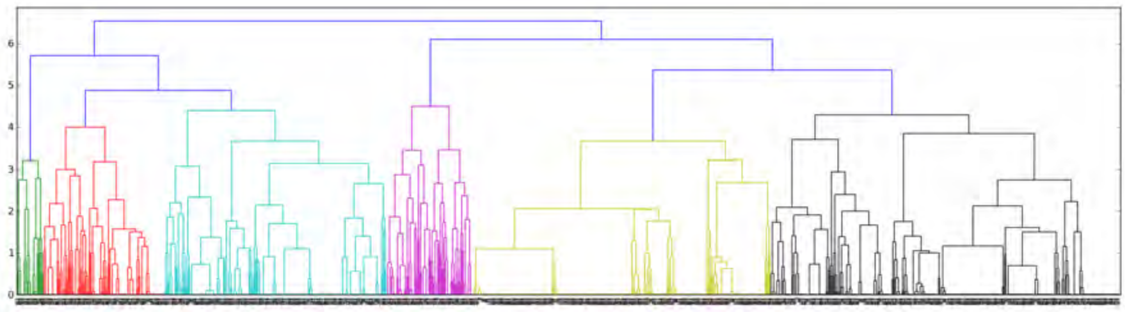


Figure 33: Cluster Dendogram for 1000 runs for the *fraction care FTE demand fulfilled*

close to 1.

2. **Classification of desirable versus undesirable outcomes (Classify):** Each run is classified according to the minimum value over each 50 year run time. Following (arbitrary) classes are used: (1) *fraction care FTE demand fulfilled* below 0.5, (2) *fraction care FTE demand fulfilled* between 0.5 and 0.75, (3) *fraction care FTE demand fulfilled* between 0.75 and 1, (4) *fraction care FTE demand fulfilled* between 1 and 1.25, and (5) *fraction care FTE demand fulfilled* and above 1.25.

   Out of 10000 runs, 3964 runs have a *fraction care FTE demand fulfilled* below 0.5, 1547 runs have a *fraction care FTE demand fulfilled* between 0.5 and 0.75, 1138 runs have a *fraction care FTE demand fulfilled* between 0.75 and 1, 3351 runs have a *fraction care FTE demand fulfilled* between 1 and 1.25, and 0 runs have a *fraction care FTE demand fulfilled* and above 1.25. There are even 1640 runs with a *fraction care FTE demand fulfilled* below 0.25.

   In spite of the fact that these numbers should not be interpreted in a probabilistic or predictive sense, it is very disturbing that so many runs have a *fraction care FTE demand fulfilled* below 0.25, and between 0.25 and 0.75. Appropriate policies are therefore urgently needed to solve plausible labor shortages in care.

3. **Identification of the most important uncertainties (Feature Selection and Random Forest):**

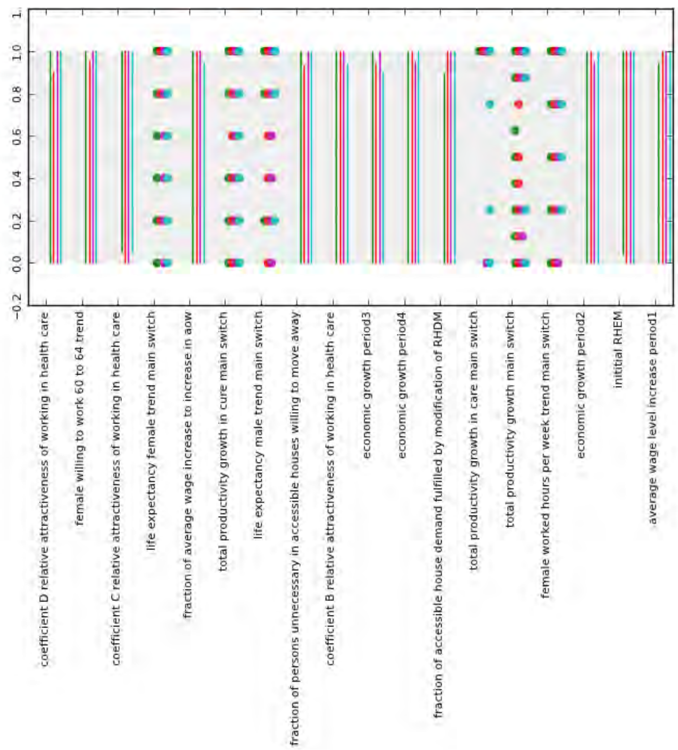
   The main uncertainties forcing this classification for the minimum values over the runs are again the evolutions of productivity in general and of productivity in care (and to a lesser extent in cure), and the male and female life expectancies.

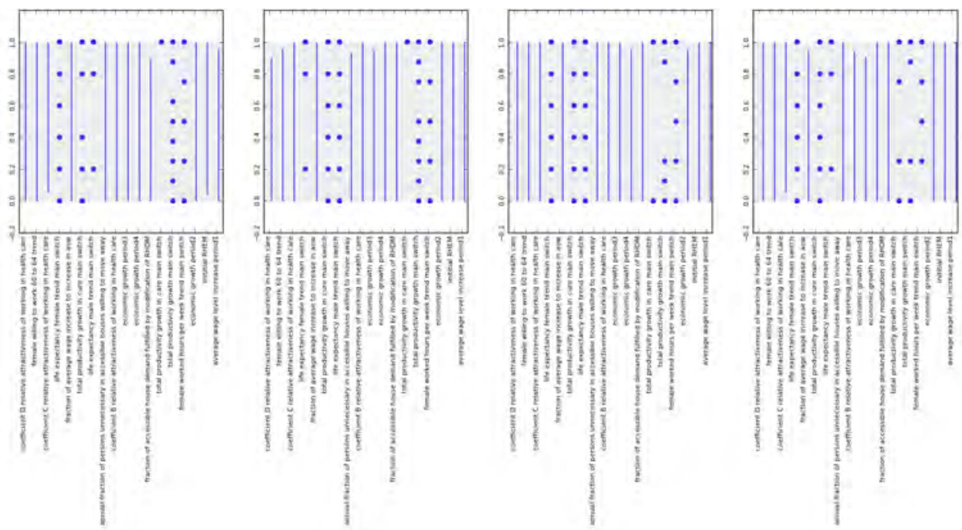
4. **Identification of multidimensional boxes in the uncertainty space that cause the undesirable outcomes (PRIM):**

   A PRIM0.3b on the minimum value of each run below 0.5 with a minimum mass of 5% of the runs and a threshold of 60% positive matches within each box results in 4 boxes with a total of 2532 cases out of 3964 bad cases in total.

	mass	mean
found box 0:	0,0577	1,00
found box 1:	0,0500	0,99
found box 2:	0,1406	0,73
found box 3:	0,0592	0,72
rest box :	0,6925	0,21

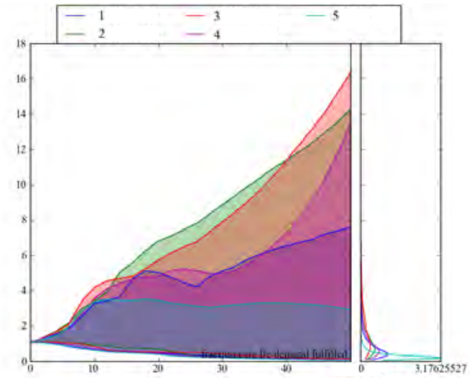
   Again, the first two boxes have (almost) 100% positive match ratios. However, these multivariate boxes are because of the number of uncertainties and values much more difficult to understand and use than many other boxes in this paper.



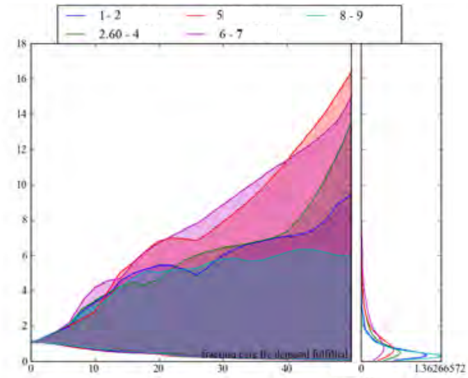


The following table displays the percentage matches with undesirable behavior of the *fraction care FTE demand fulfilled* for the three most important switch uncertainties: More than 86% of the 2000 runs with a switch value 5 for the *total productivity growth in care main switch* – and consequently following a very small increase and later a decrease of the *total productivity growth in care*– generate a future with a minimum *fraction care FTE demand fulfilled* below 50%.

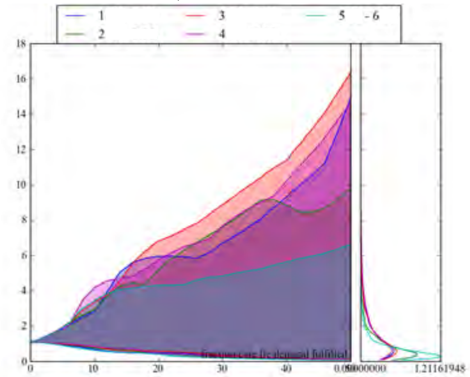
The graphs below show envelopes for different values of the top-most uncertainties.



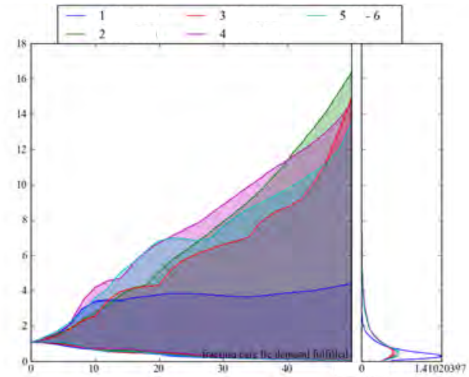
total productivity growth in care main switch – switch 5 (decreasing product. in care) is problematic



total productivity growth main switch switches 8,1,2 (increasing productivity) are somewhat problematic



life expectancy male trend main switch



total productivity growth in cure switch

The light blue envelope in the first graph confirms that the fifth lookup function for the *total productivity growth in care* is problematic: this switch corresponds to a very low growth of productivity in care for the first 20 years followed by a continued decrease to 40% of its 2010 value in 50 years. This development may be unrealistic but it nevertheless points to the need for continued growth of productivity in care. The inverse is true for the total productivity (switch values 8, 9, 1, 2, 3): higher productivity in the rest of the economy seems to drain potential care workers away from care towards the rest of the economy.

fraction care fte demand fulfilled < 0.5	Avg Nr Runs	matches with undesirable behaviour										TOTAL NR UN-
main switch uncertainties:	per switch	1	2	3	4	5	6	7	8	9	10	DESIRABLE RUNS
total productivity growth in care main switch	2000	47,00%	21,50%	13,10%	30,20%	86,40%						3964
total productivity growth main switch	1112	61,51%	42,00%	48,02%	23,65%	26,62%	19,42%	19,69%	56,92%	58,63%		3964
life expectancy male trend main switch	1667	27,95%	44,03%	29,99%	26,33%	49,13%	60,35%					3964

A.7.3 Conclusions:

Regarding the *fraction care FTE demand fulfilled*, the following preliminary conclusions could be drawn:

- The combination of this model and these uncertainty ranges generates many cases with serious scarcity of labor supply in care.
- Consequently, there are several relatively large uncertainty subspaces with high concentrations of runs that have a tendency to generate labor scarcity in care at one or more moments during the 50 year runs.
- Major uncertainties determining the split in abundance/scarcity are the developments of growth in productivity –especially the productivity growth in care. It should be noted that

–at least in this model– productivity in care and productivity in general have opposite effects on labour shortage in care.

- Although productivity growth in the Dutch economy at large seems to be –given the model, uncertainties, and omissions (e.g. specific migration policies)– good for several key output indicators, it may also cause labour scarcity in care.

  Hence, productivity growth in care should –for this KPI– be stimulated more than productivity growth in the rest of the economy, except for productivity growth in the rest of the economy that does not pinch employees needed in care. Hence, in the absence of immigration of qualified workers in care, priority should be given to productivity growth in care and in those sectors of the economy that are not competing for the same (potential) employees.

A.8 Further investigation of the *fraction labour scarcity*

A.8.1 Visualization and Analysis

Figures 34 and Figure 35 show the trajectories and the logarithmic 3D KDE plot of the *fraction labour scarcity*, both for 10000 runs. This fraction relates to the labour scarcity in the Dutch economy at large, i.e. the total work force (minus the structural unemployment) divided by total labour demand. Looking at these Figures, it looks as though many runs are characterized by long term waves.

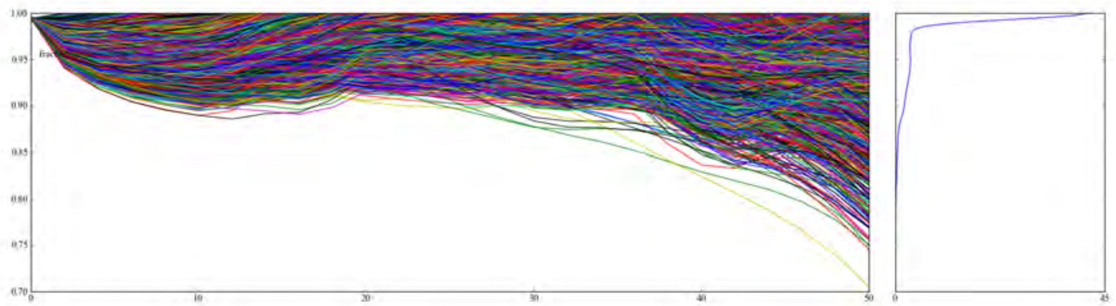


Figure 34: Lines of 10000 runs for the *fraction labour scarcity*

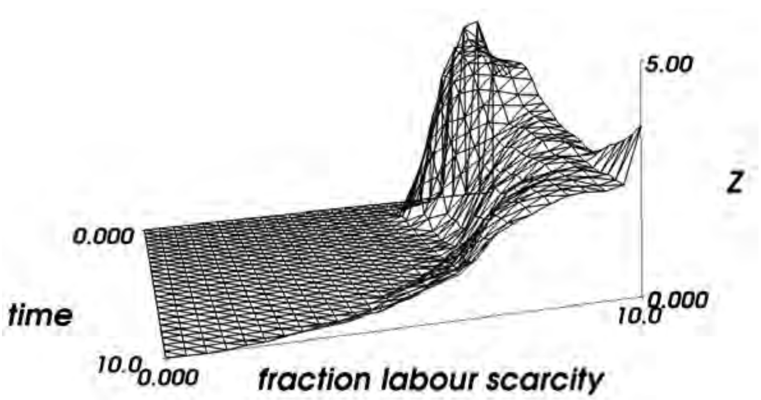


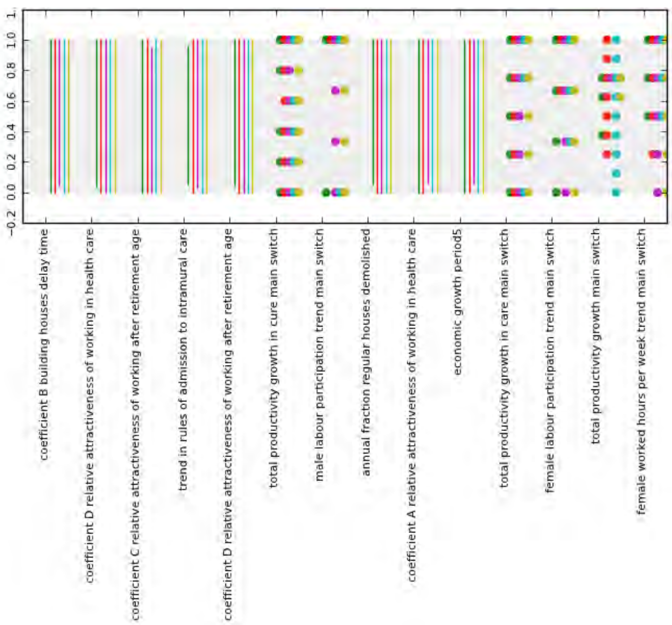
Figure 35: 3D KDE plot of the Fraction Labour Scarcity for 10000 runs

Following arbitrary classification –based on the end state KDE in Figure 34– was used for Feature Selection and Random Forest algorithms on the large ensemble: (1) *fraction labour scarcity* between 1 and 0.975, (2) between 0.975 and 0.95, (3) between 0.95 and 0.9, (4) between 0.9 and 0.8, and (5) below 0.8. Out of 10000 runs, 7177 runs have end state values lying between 1 and 0.975, 794 runs between 0.975 and 0.95, 1356 runs between 0.95 and 0.9, 628 runs between 0.9 and 0.8, and 45 runs below 0.8.

The most important uncertainties according to the Random Forest and Feature Selection algorithms applied to the end state values of 10000 runs using this classification are the total productivity growth and the total productivity growth in care.

A PRIM0.3b on the minimum value of each run below 0.95 with a minimal mass of 5% of the runs and a threshold of 60% positive matches within each box results in 5 boxes with a total of 3083 cases out of 4115 bad cases in total out of 10000 runs.

	mass	mean
found box 0:	0,0663	1,00
found box 1:	0,0593	0,96
found box 2:	0,0624	1,00
found box 3:	0,0645	0,78
found box 4:	0,1011	0,72
rest box :	0,6464	0,16



Again, these boxes are difficult to interpret. The switch uncertainties contributing most to this classification are again the *total productivity growth main switch* and the *total productivity growth in care main switch*. About 90% of the 1112 runs with a *total productivity growth main switch* equal to 7 (∼ decreasing productivity) generates end state values for the *fraction labour scarcity* of less than 0.95:

end state (fraction labour scarcity) < 0.95	Avg Nr Runs					matches with undesirable behaviour						TOTAL NR UN-
main switch uncertainties:	per switch	1	2	3	4	5	6	7	8	9	10	DESIRABLE RUNS
total productivity growth main switch	1112	5,40%	9,80%	5,67%	14,03%	13,67%	35,16%	89,39%	4,77%	4,59%		2029
total productivity growth in care main switch	2000	14,40%	13,00%	12,85%	13,10%	48,10%						2029
female labour participation trend main switch	2500	17,36%	16,92%	21,04%	25,84%							2029

Note however that this arbitrary threshold value (0.95) is relatively high compared to other threshold values, in other words, not so bad.

A.8.2 Conclusions:

Low productivity growth (and productivity growth in care) are major drivers for many cases with labour shortages. In those cases, more people are needed to do the same amount of work. And future labour scarcity may well come in waves.

A.9 Further investigation of the *fraction housing demand fulfilled*

A.9.1 Visualization and Time-series Clustering

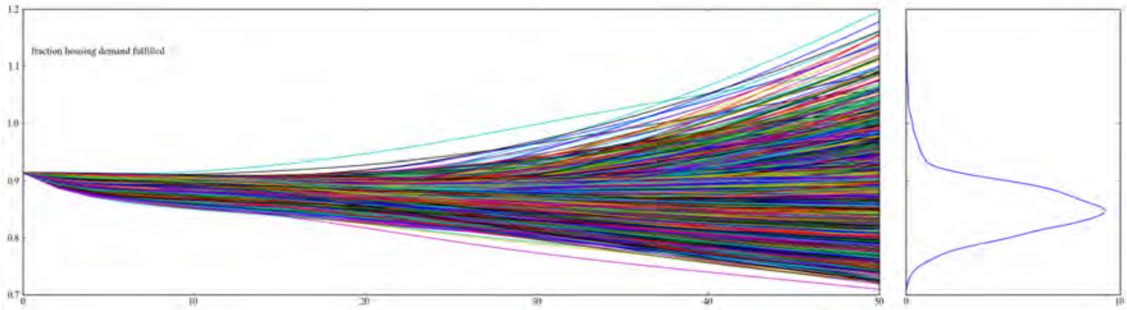


Figure 36: Lines for 10000 runs of the *fraction housing demand fulfilled*

Figure 36 shows the *fraction housing demand fulfilled* of an ensemble of 10000 runs, and Figure 37 the logarithmic 3D KDE plot of 10000 runs. In the absence of additional policies, the ensemble shows a continuous evolution –overall towards slightly more unfulfilled (special) housing demand, but in some cases towards less unfulfilled (special) housing demand.

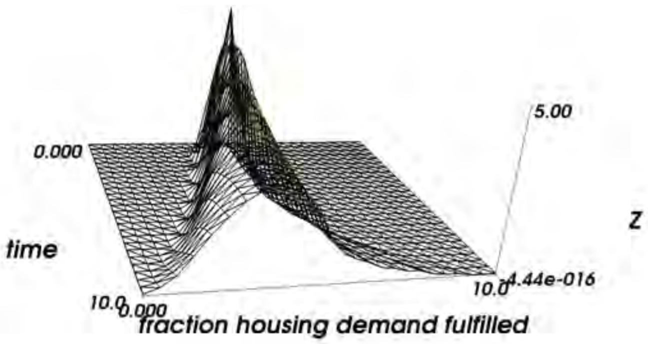


Figure 37: 3D plot of the fraction Housing Demand fulfilled

A.9.2 Reference Outcomes, Classification, Feature Selection & Random Forest, and PRIM

1. **Identification of undesirable outcomes:** Starting from slightly over 90% fulfilled, it is assumed here that worse than today is undesirable.

2. **Classification of desirable versus undesirable outcomes (Classify):** Hence the minimum values of each run is classified by means of the following classification: (1) less than 70% housing demand fulfilled, (2) between 70% and 80% housing demand fulfilled, (3) between 80% and 90% housing demand fulfilled, and (4) above 90% fulfilled at any time over a run.

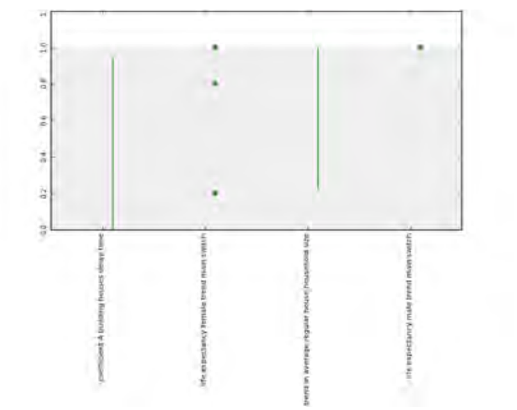
   Out of the 10000 runs, 0 runs have a *fraction housing demand fulfilled* of less than 70% housing demand fulfilled, 1159 runs have a *fraction housing demand fulfilled* between 70% and 80% housing demand fulfilled, 8786 runs have a *fraction housing demand fulfilled* between 80% and 90% housing demand fulfilled, 55 runs have a *fraction housing demand fulfilled* above 90% fulfilled at any time over a run. Hence, active housing policies seem to be needed, at least, in/according to this model.

3. **Identification of the most important uncertainties (Feature Selection and Random Forest):** The most important uncertainty for this classification of the minimum value of each run, based on a feature selection and a random forest, is the *trend in male life expectancy*, followed at some distance by the *trend in female life expectancy*, and the *trend in average regular house household size.*

4. **Identification of multidimensional boxes in the uncertainty space that cause the undesirable outcomes (PRIM):**

   A PRIM0.3b on the minimal value of each run below 0.8 with a minimal mass of 5% of the runs and a threshold of 60% positive matches within each box results in 1 boxes with a total of 539 cases out of 1159 bad cases in total out of 10000 runs.

	mass	mean
found box 0:	0,061	0,88
rest box :	0,939	0,07



Especially the highest evolution of the *male life expectancy* (to 90+ years) shifts the envelope and KDE of the housing demand fulfilled downwards (see Figure 38).

The table below displays the effects on the *fraction housing demand fulfilled* of the evolutions of the two most important switch uncertainties. The first uncertainty –the evolution of the male life expectancy– has a match ratio (i.e. a minimum *fraction housing demand fulfilled* below 0.80) of close to 50%. Hence, for special housing, it may be important to monitor average male life expectancy.

fraction housing demand fulfilled < 0.8	Avg Nr Runs				matches with undesirable behaviour							TOTAL NR UN-
main switch uncertainties:	per switch	1	2	3	4	5	6	7	8	9	10	DESIRABLE RUNS
life expectancy male trend main switch	1667	0,12%	7,50%	0,90%	0,18%	12,42%	48,41%					1159
life expectancy female trend main switch	1667	2,52%	13,62%	4,98%	2,94%	15,96%	29,51%					1159

A.9.3 Conclusions:

The combination of this model and these uncertainty ranges, generates –in the absence of additional housing policies– many runs with a larger unfulfilled fraction of housing demand than today. Shortages are especially the case for high-end evolutions of the *male life expectancy*, and to a lesser extent for high-end evolutions of the *trend in female life expectancy*, strongly decreasing *house household sizes*, and longer delays in house construction.

A.10 Further investigation of the *demand for accessible houses*

Figure 39 shows the trajectories and the (non-logarithmic) 3D KDE plot of *demand for accessible houses*, both for 10000 runs. There is one outlier.

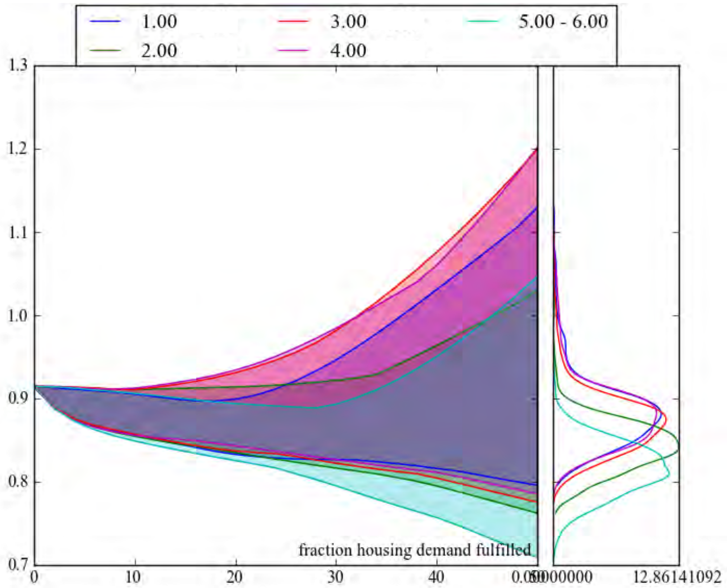


Figure 38: Envelopes and KDEs of the *fraction of housing demand fulfilled* for different evolutions of the main uncertainty, the *male life expectancy*

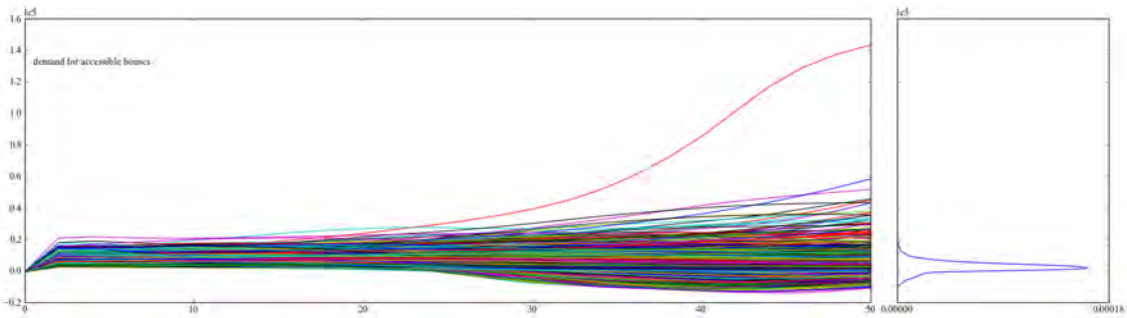


Figure 39: Lines for 10000 runs of the *demand for accessible houses*

A.11 Superficial investigation of the *average labour participation female 55 to 64*

Figure 40 shows the trajectories and the logarithmic 3D KDE plot of the *average labour participation female 55 to 64*, both for 10000 runs.

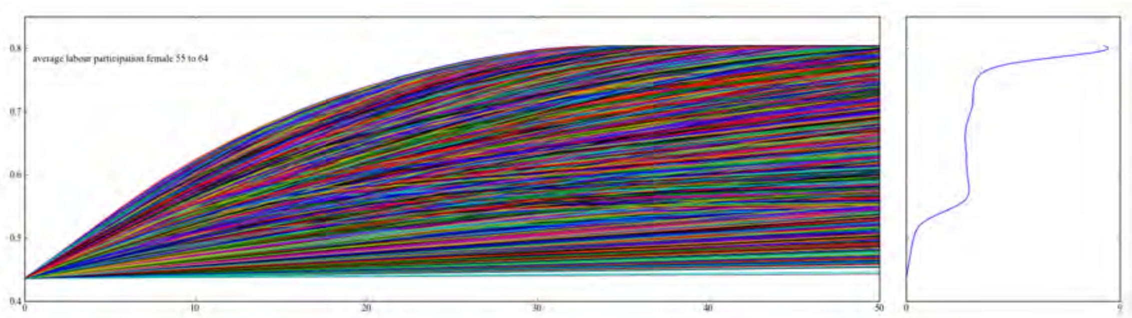


Figure 40: Lines graph of 10000 runs for Labour Participation of Females betwn 55-64

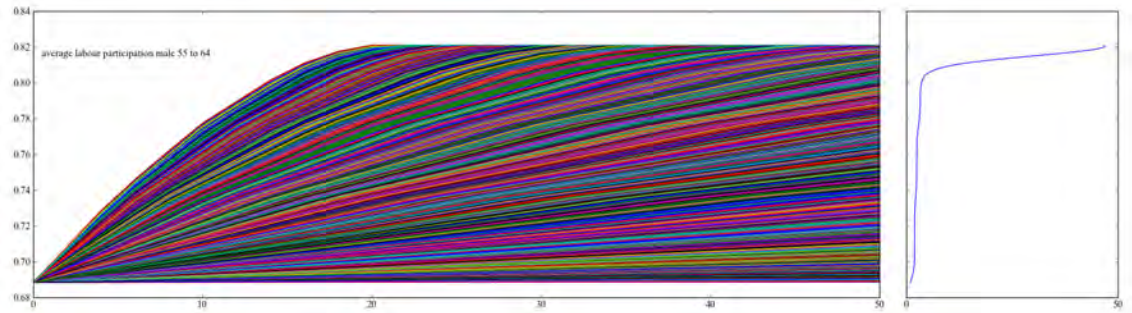
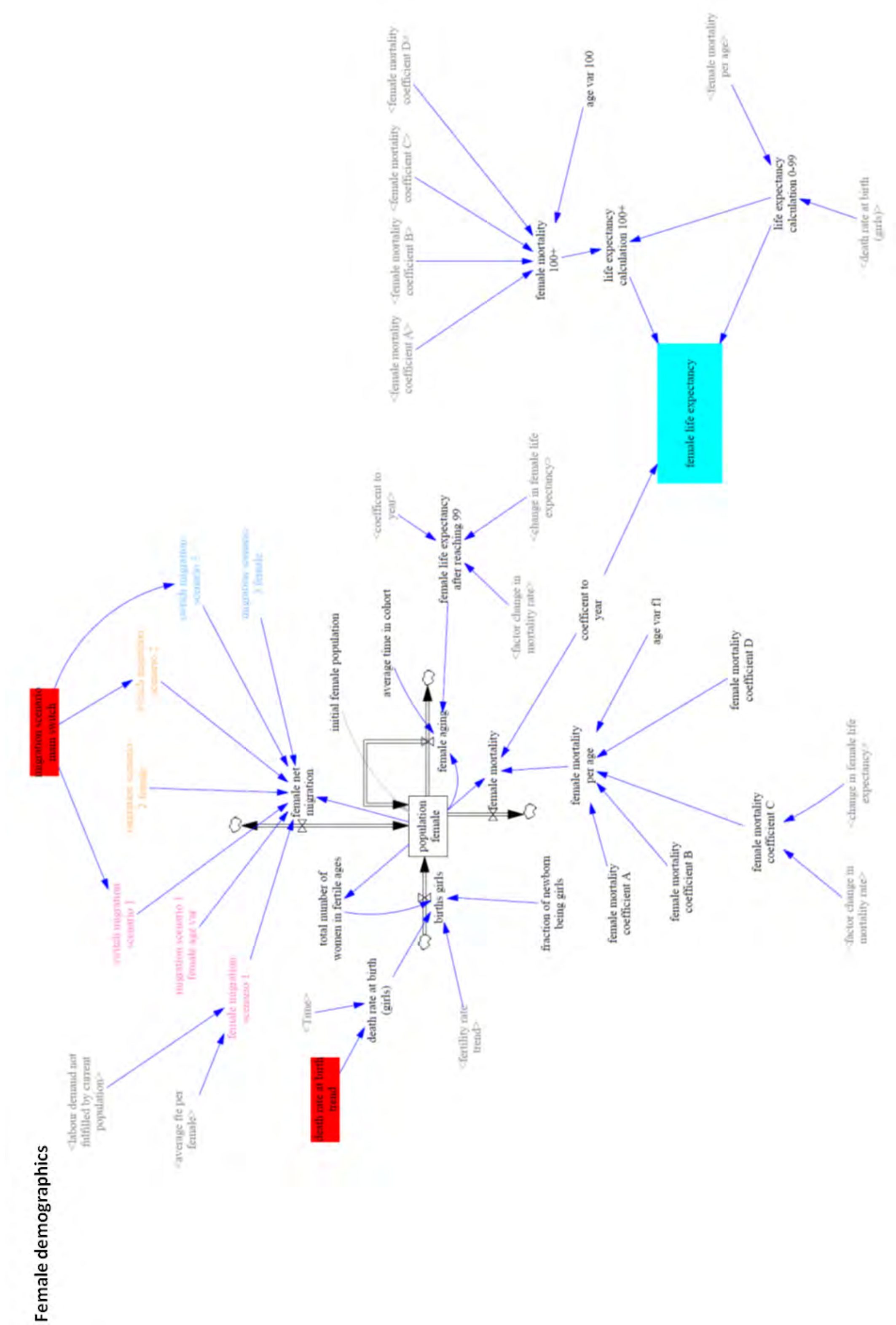


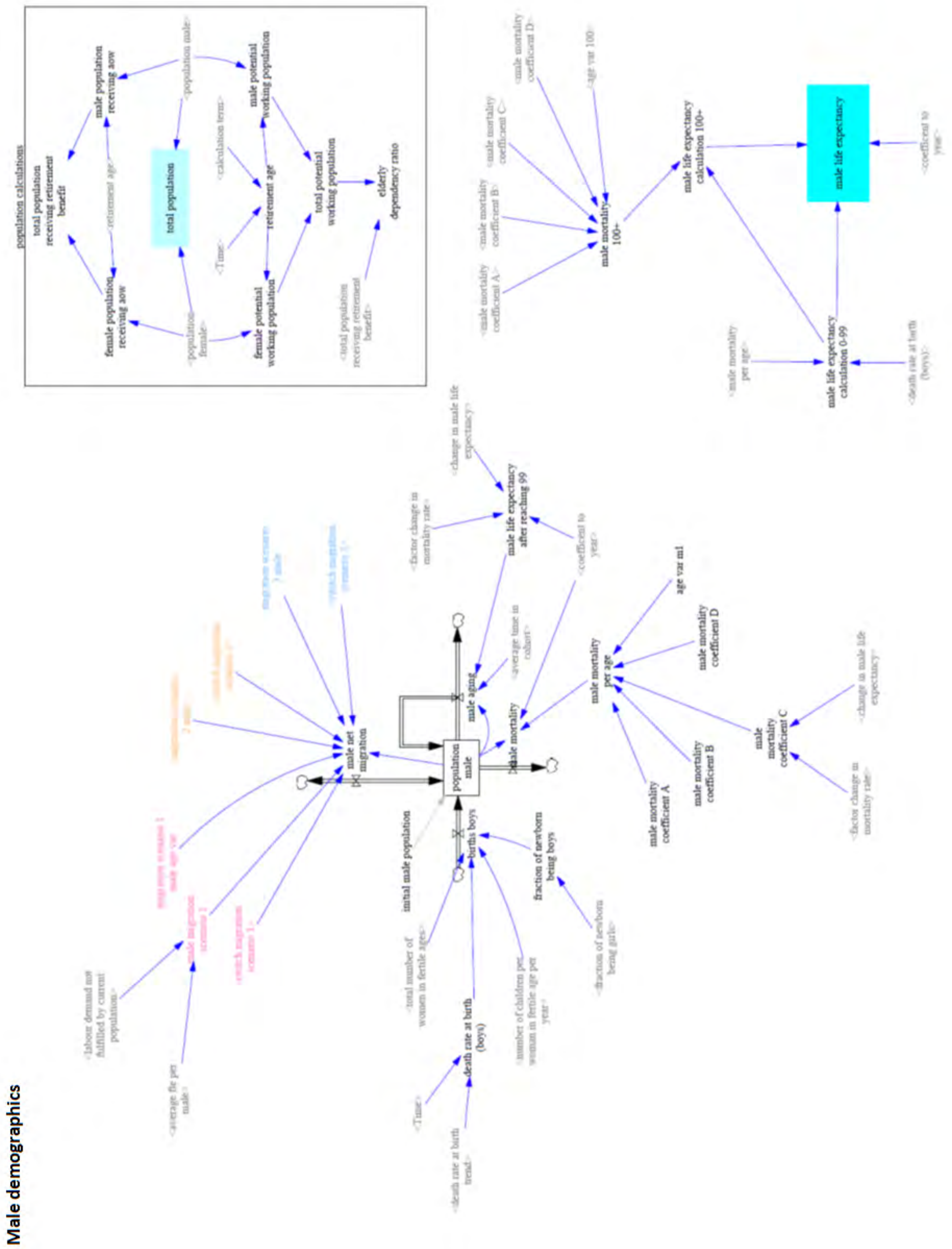
Figure 41: Lines graph of 10000 runs for the Labour Participation of Male Employees betwn 55-64

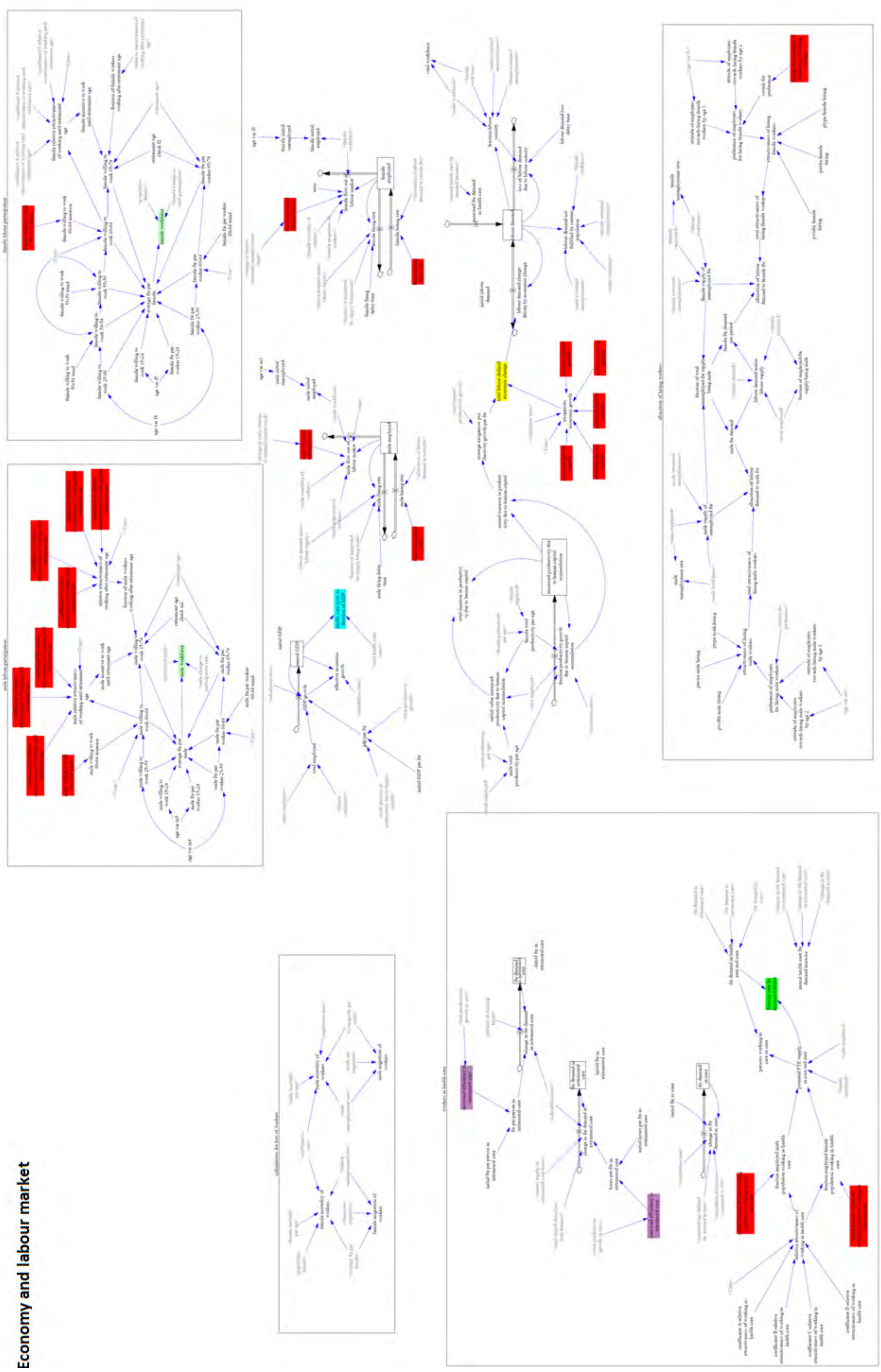
A.12 Superficial investigation of the *average labour participation male 55 to 64*

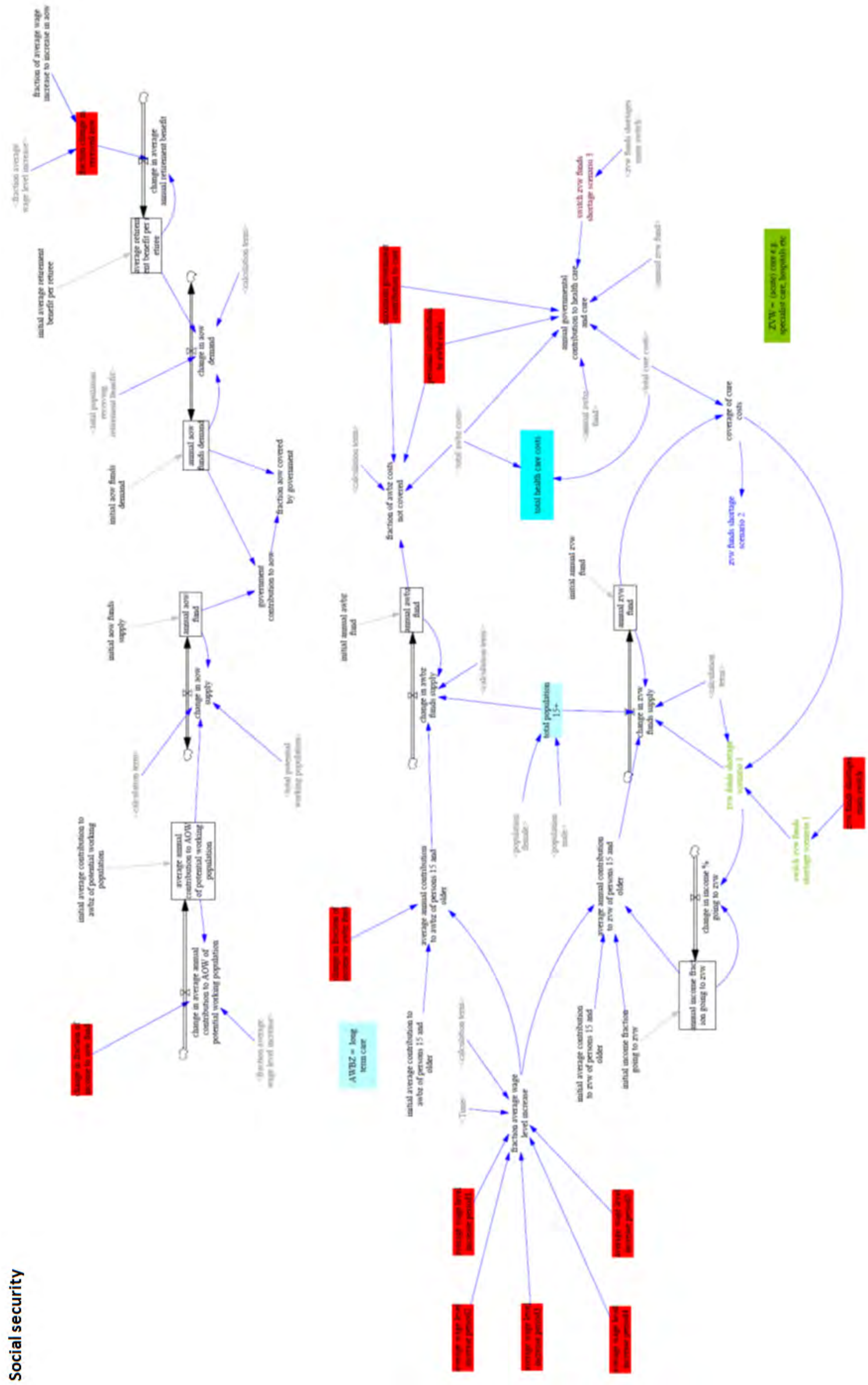
Figures 41 show the trajectories and the logarithmic 3D KDE plot of the *average labour participation male 55 to 64*, both for 10000 runs.

B Stock-Flow Diagrams of the Sectors in the Model

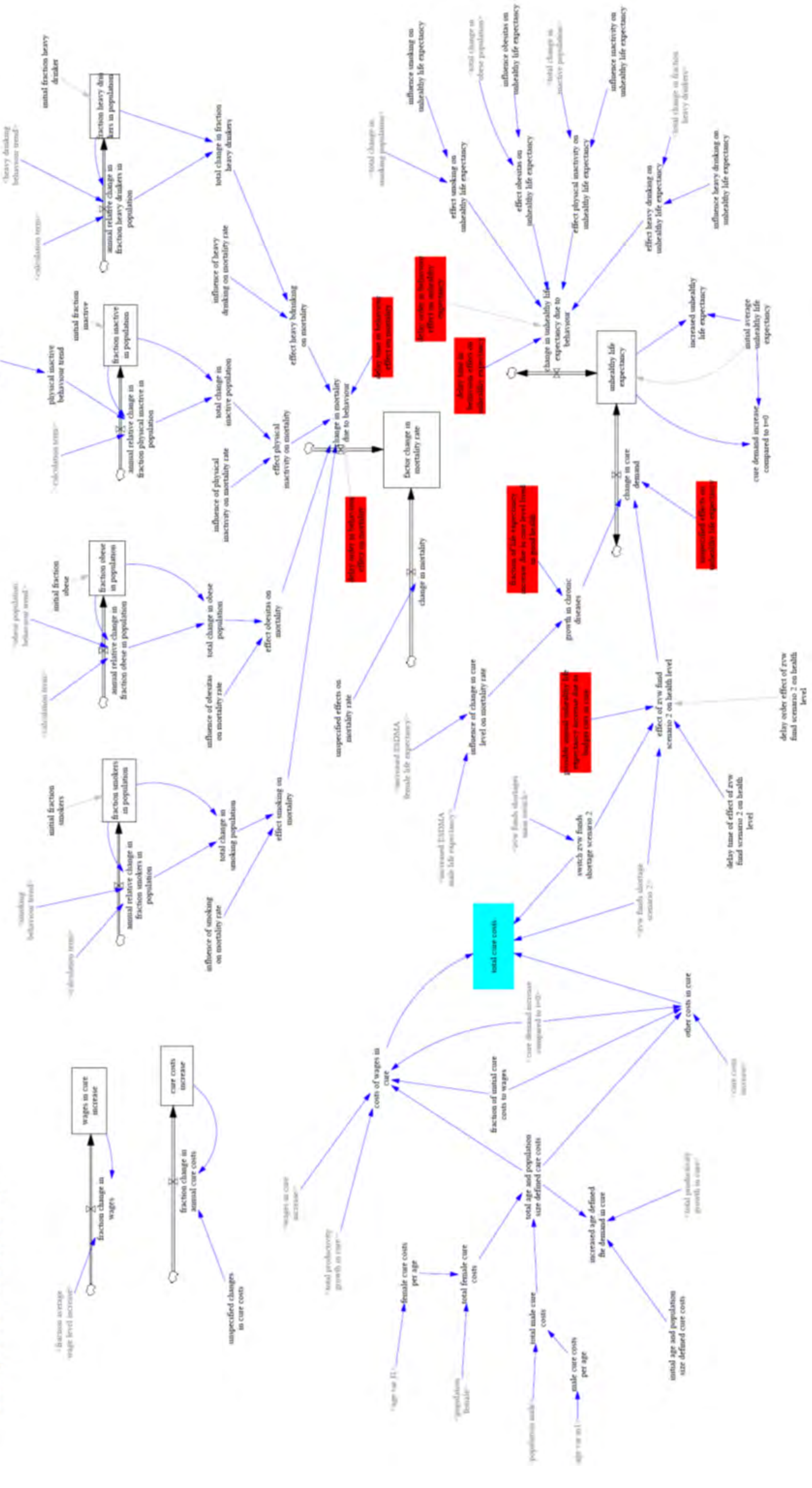


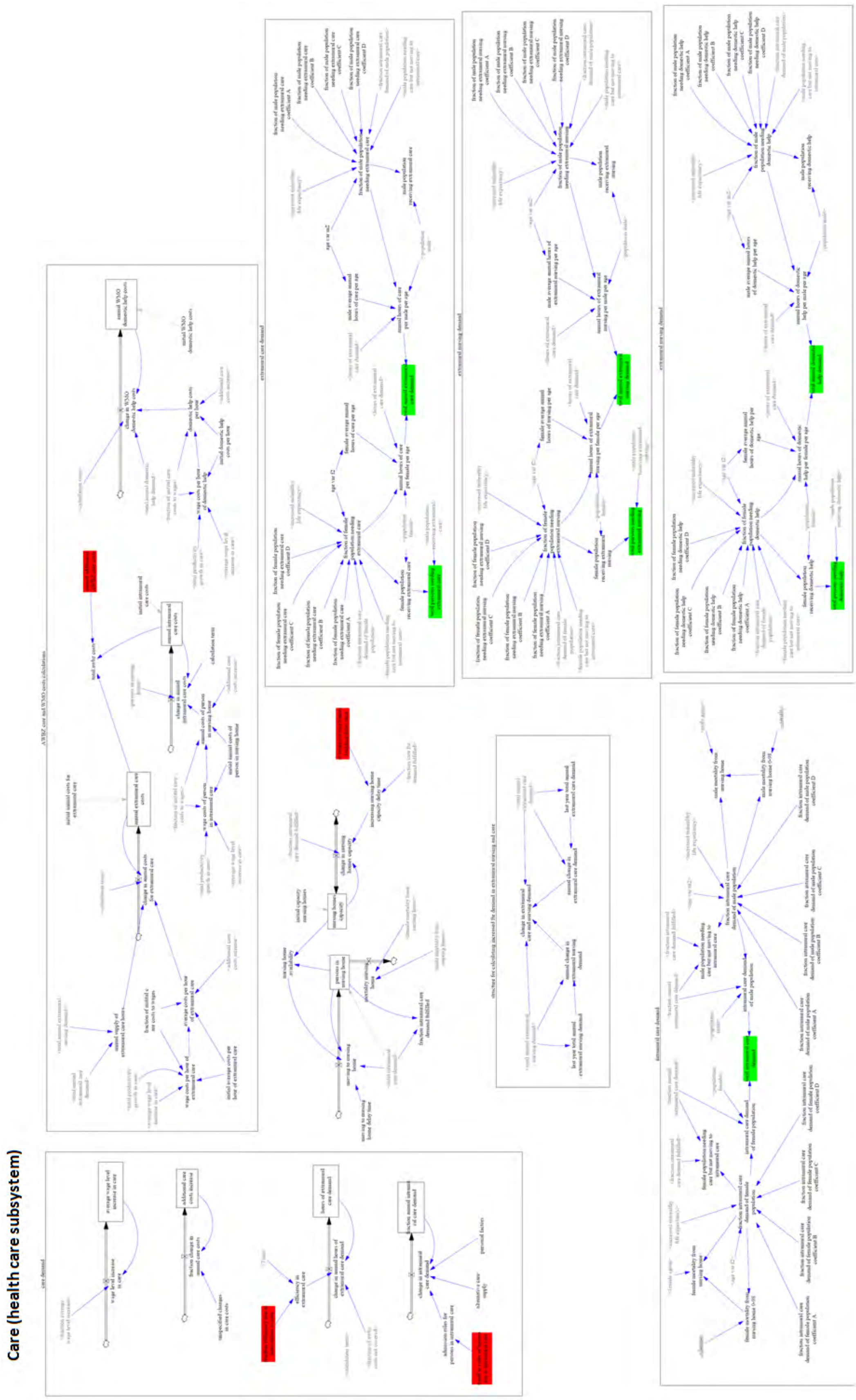

Male demographics
population calculations
total population receiving retirement benefit
male population receiving aow
female population receiving aow
total population
male potential working population
female potential working population
retirement age
total potential working population
elderly dependency ratio
male mortality 100+
male life expectancy calculation 100+
male life expectancy
male life expectancy calculation 0-99
male life expectancy after reaching 99
population male
male net migration
male aging
male mortality
births boys
initial male population
fraction of newborn being boys
death rate at birth (boys)
male mortality per age
age var m1
male mortality coefficient A
male mortality coefficient B
male mortality coefficient C
male mortality coefficient D

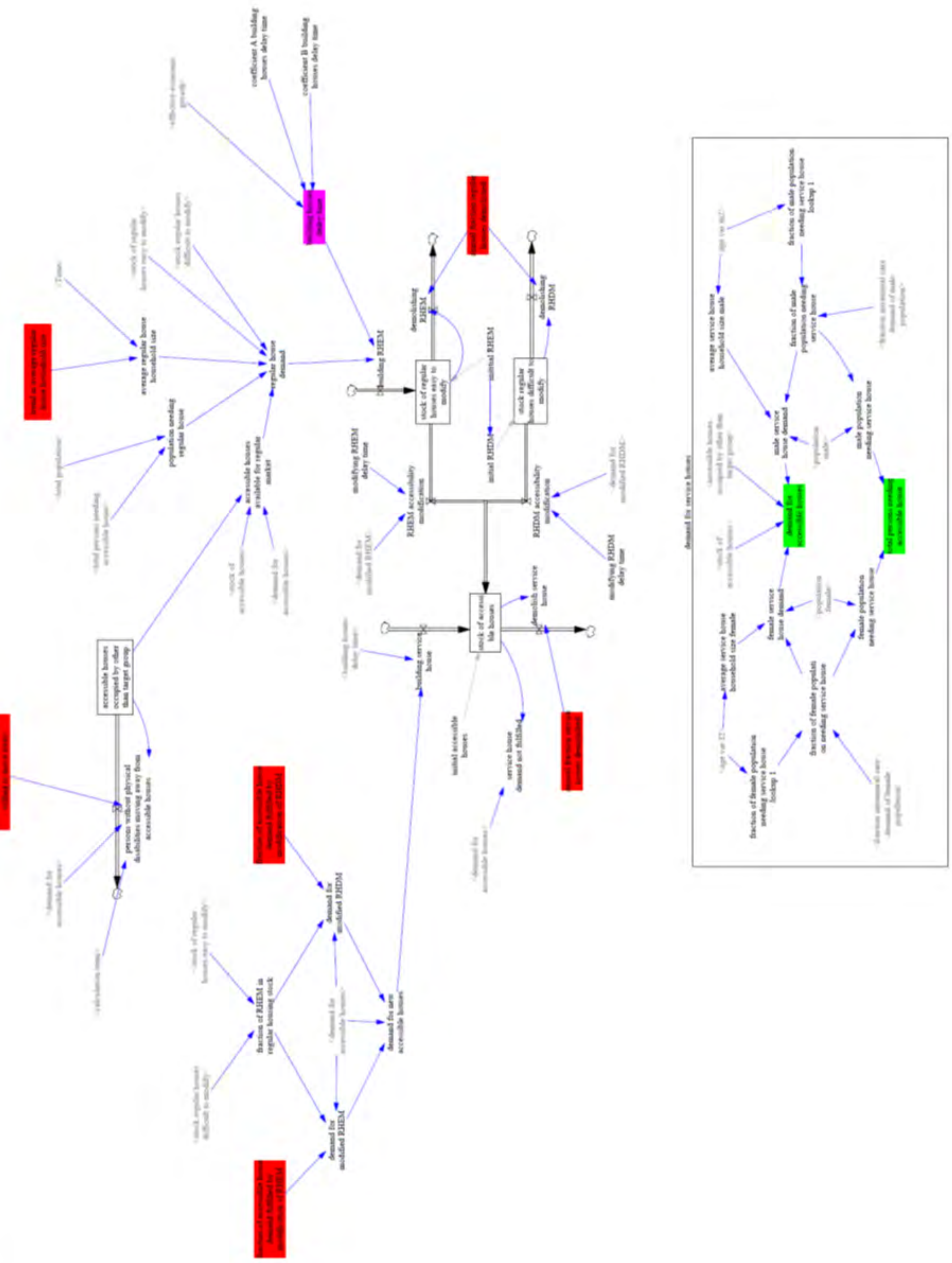

Economy and labour market


Social security

Cure (health care subsystem)




Care (health care subsystem)



Housing