IARIW 2022

IARIW 2022

Monday 22 - Friday 26 August

Wealth Dynamics of Households; Linking Micro and Macro

Ingber Roymans
(Banque centrale du Luxembourg)
ingber.roymans@bcl.lu

Paper prepared for the 37th IARIW General Conference

August 22-26, 2022

Session 6C-2, Reducing Gaps between Micro and Macro Statistics on Household Income, Consumption, and Wealth in Compiling Distributional National Accounts II

Time: Friday, August 26, 2022 [11:00-12:30 CEST]

Wealth dynamics of households: linking micro and macro

Ingber Roymans[1]

Abstract. Distributional National Accounts (DNA) integrate micro data, for example from a household survey, with the macro totals of the National Accounts (NA). This paper focusses on the wealth of households, equated with the NA "Net worth" concept. By integrating the non-financial and financial parts of the DNA, one obtains a picture of the wealth dynamics (changes in wealth over time) of different household groups.

However, the NA describe only the household sector as a whole. Transfers of wealth between different household groups are in general ignored in the NA, while movements of households (including their wealth) between groups cancel out at the macro level. As a result, the NA, and therefore DNA, cannot give a complete picture of the changes in wealth of a single household group. To get some insight into these inter-group flows, we have to simulate individual households and their wealth over time (micro simulation).

We discuss a basic methodology for simulating the wealth of a household group over time, based on a household survey sample, in this case the Luxembourg Household Finance and Consumption Survey (HFCS). To account for population growth, we have to simulate the immigration and emigration of households and deaths and births. As a bonus, this gives us an estimate for the capital transfers between households due to inheritances. It turns out that not all relevant wealth dynamics are easily captured, for example: mergers and splits due to marriages and divorces are difficult to simulate. In the end, we thus obtain a complete, but not necessarily correct, estimate of the changes in wealth of each household group due to six components: Savings, Capital transfers, Holding gains and losses, Net migration, (social) Mobility (i.e. the movement of households between groups) and Other changes.

The method gives the dynamics of household wealth at a single point in time. But these changes can also be extrapolated forwards and backwards in time to create a small time series. Such time series segments from different surveys can be stitched together into a longer time series. This stitching process will naturally create another wealth flow, a statistical artifact, which can be shown separately or absorbed into the components mentioned above.

Keywords: Distributional National Accounts; Household Finance and Consumption Survey; micro simulation; inequality

Disclaimer: The results in this paper are preliminary and aimed to stimulate discussion and critical comments. References in publications should be cleared with the author. This paper should not be reported as representing the views of the BCL or the Eurosystem. The views expressed are those of the author and do not necessarily reflect those of other staff or policymakers in the BCL or the Eurosystem.

[1] Banque centrale du Luxembourg,
2, boulevard Royal, L-2983 Luxembourg,
Tel. : +352 47744318, E-mail : ingber.roymans@bcl.lu

Acknowledgements: I am grateful for the valuable comments received from Paolo Guarda, Michael Ziegelmeyer and Jean-Pierre Schoder.

1. Introduction

This paper focusses on the wealth dynamics of households, i.e. the changes in wealth of a household or a household group over time. The distribution of wealth among households is of high interest for economic analysis and several international initiatives exist to measure this distribution and it's change over time, see for example [Coli et al. 2022; Garbinti et al. 2021; Zwijnenburg et al. 2021]. The joint OECD-Eurostat Expert Group on Measuring Disparities in a National Accounts Framework (EG-DNA) and the European Central Bank (ECB) Expert Group on Linking Macro and Micro Data for the Household Sector (EG-LMM) both are developing a methodology to link micro data from surveys with macro data from the NA.

This paper has the same objectives. We use data from a sample of individual households at micro level. Such micro data is typically collected through a survey, in this case the Household Finance and Consumption Survey (HFCS) [ECB 2016]. These micro data are then linked with macroeconomic National Accounts concepts. By linking economic micro and macro data within a National Accounts framework, we obtain so-called Distributional National Accounts (DNA) [Coli et al. 2022; Garbinti et al. 2021; Zwijnenburg et al. 2021].

The National Accounts provide a consistent and complete picture of the "Net worth" (B.90) of the "Household sector" (S.14) and its changes over time [United Nations 2009; Eurostat 2013]. We will therefore conveniently equate our "wealth" concept with Net worth, which is equal to all household financial and non-financial assets minus household financial liabilities as defined by the NA. The non-financial part of the NA and DNA describes how the savings and capital transfers of households are the result of their income, consumption and capital transfers. The financial part of the NA and DNA describes the assets and liabilities of a sector, as well as the transactions, holding gains and losses and other changes in these assets. Thus, by combining non-financial and financial DNA we obtain a complete picture of the change in wealth of the household sector.

However, NA describe only the household sector as a whole, and therefore often ignore wealth transfers between households, while the movement of households or people (including their wealth) between household groups cancel out at the macro level. Thus, the NA, and necessarily also the DNA, do not provide a complete picture of the changes in wealth of a household group. To gain some insight into these additional wealth flows we will have to perform a simulation at micro level. The main advantage of our method is that, by design, it guarantees consistency in four areas:

- the change over the year equals the sum of the changes over the quarters,
- the sum of the flows equals the change in stocks,
- the net lending/net borrowing is the same between financial and non-financial distributional accounts
- group totals equal the micro data aggregates.

The theory behind is described in section 2, where we derive a micro level flow equation. The main challenge is to identify the wealth flows at micro level with the familiar wealth flows in the NA. As it turns out, this is non-trivial and a suitable simulation approach has to be found that allows for a straightforward one-to-one link. Section 3 describes a minimum, baseline simulation methodology, how to construct a time series and some possible ways to present the results. Section 4 concludes.

2. Theory

2.1 Household groups over time

Because equation (10) is central to our discussion, we derive it here. However, the reader might skip this technical part and go directly to section 2.2.

We write $X_i(t)$ as the value of variable X for an individual household i at time t. Here i indexes all the households that existed in the population P during the time interval under consideration. A typical time interval would be a year or a quarter. Below, X stands for wealth, but it could stand for any numerical variable associated with a household. Its change between time t_0 and time t_1 is given by

(1) $\nabla X_i(t_0, t_1) = X_i(t_1) - X_i(t_0)$

which is called the backward difference at t_1 in the calculus of finite differences. Any change in a stock type of variable, like wealth, is called a "flow".
The aggregate of X over a household group Q, i.e. a specific set of households i, at time t is written

(2) $X_Q(t) = \sum_{i \in Q(t)} X_i(t)$

where the summation is now only over those households i that fall within group Q at time t. In addition, we can define the population total X by simple setting Q equal to P.

(3) $X(t) \equiv X_P(t) = \sum_{i \in P(t)} X_i(t)$

We shall define the change in X_Q over time as

(4) $\nabla X_Q(t_0, t_1) \equiv X_Q(t_1) - X_Q(t_0) = \sum_{i \in Q(t_1)} X_i(t_1) - \sum_{i \in Q(t_0)} X_i(t_0)$

This definition is preferable to the common alternative

(5) $\nabla X_Q(t_0, t_1) \equiv \sum_{i \in Q(t_1)} \nabla X_i(t_0, t_1)$

because (4) has the desirable property

(6) $\nabla X_Q(t_0, t_2) = \nabla X_Q(t_0, t_1) + \nabla X_Q(t_1, t_2)$; where $t_0 < t_1 < t_2$

In words, using definition (4) guarantees that the change over a year equals the sum of the changes over the four quarters. This is not true for the alternative definition (5). Notice that (4) is path independent, i.e. it only depends on the values at times t_1 and t_0, and not on any times in between. It can be decomposed into two components:

(7) $\nabla X_Q(t_0, t_1) = C1_Q(t_0, t_1) + C2_Q(t_0, t_1)$

with

$$C1_Q(t_0, t_1) = \int_{t_0}^{t_1} \sum_{i \in Q(t)} \frac{dX_i(t)}{dt} dt$$

and

$$C2_Q(t_0, t_1) = \sum_i \left([t_1 > t_{iQ}^+ \geq t_0] X_i(t_{iQ}^+) - [t_1 \geq t_{iQ}^- > t_0] X_i(t_{iQ}^-) \right)$$

where t_{iQ}^+ denotes the time when household i becomes a member of group Q and t_{iQ}^- denotes the time when it ceases to be a member. The [ ] brackets are Iverson brackets, defined for any statement S as: [S] = 1 if S is true, otherwise [S] = 0.

Component C1 represents the effect from the changes in the variable X_i experienced by the individual households, while component C2 represents the effect from changes in the composition of the group. It represents the amount brought into the group by households entering, minus the amount taken out by households leaving the group[2]. The challenge is that although the sum of C1 and C2 is path independent, C1 and C2 themselves are not: they depend on the values of X_i and the composition of group Q at every point in time. Because such detailed information is generally not available, we have to find approximations for C1 and C2 that are path independent. Common practice is to use averages for X_i and Q. We can write

(8) $C1_Q \cong \frac{1}{2}\left(\sum_{i \in Q(t_1)} \nabla X_i + \sum_{i \in Q(t_0)} \nabla X_i\right)$

$$C2_Q \cong \frac{1}{2}\left(\sum_{i \in Q(t_1)} \left(X_i(t_1) + X_i(t_0)\right) - \sum_{i \in Q(t_0)} \left(X_i(t_1) + X_i(t_0)\right)\right)$$

(For convenience, and because we consider a fixed time interval, below we will drop the (t_0, t_1) from our notation for the changes ∇.) Such a linear approximation is exact whenever the variables X_i change linearly over time and changes in the composition of the group are evenly spread over time. It is the best decomposition we can make in the absence of more information. Note that (7) still holds, i.e. the sum of C1 and C2 corresponds to the total change; it is only decomposition (8) that is an approximation.

Formula (8) is an application of the product rule within the calculus of finite differences:

(9) $\nabla(f \cdot g) = \nabla f \cdot \overleftarrow{g} + \overleftarrow{f} \cdot \nabla g$

were the backward average of f at t_1 is defined as

$\overleftarrow{f} \equiv \frac{f(t_1) + f(t_0)}{2}$; where $t_0 < t_1$

This rule will be used several times in what follows.

---

[2] Although this expression is already complicated, it is still a simplification because it ignores that households can enter and leave multiple times. For intervals that are sufficiently short such an approximation is justified.

When we are dealing with one or more samples S of households, i now indexes the different households in all the samples available for the time interval. Each household in the sample has to be properly weighted so that it represents a number $w_i(t)$ of households in the population. In general, this is done with a weighting procedure, like calibration or post-stratification, using selected auxiliary variables; see for example [Nguyen 2011]. In this case, the composition of a group Q has to be re-written as

$$\sum_{i \in Q(t)} \quad \rightarrow \sum_{i \in Q(t)} w_i(t)$$

Thus, one now also has to take account of any changes in the weights and re-write (8) as

$$\text{(10)} \qquad C1_Q \cong \frac{1}{2}\left(\sum_{i \in Q(t_1)} w_i(t_1)\nabla X_i + \sum_{i \in Q(t_0)} w_i(t_0)\nabla X_i\right)$$

$$C2_Q = C2A_Q + C2B_Q$$

with

$$C2A_Q \cong \frac{1}{4}\left(\sum_{i \in Q(t_1)} \big(w_i(t_1) + w_i(t_0)\big)\big(X_i(t_1) + X_i(t_0)\big)\right) - \frac{1}{4}\left(\sum_{i \in Q(t_0)} \big(w_i(t_1) + w_i(t_0)\big)\big(X_i(t_1) + X_i(t_0)\big)\right)$$

$$C2B_Q \cong \frac{1}{4}\left(\sum_{i \in Q(t_1)} \nabla w_i\big(X_i(t_1) + X_i(t_0)\big) + \sum_{i \in Q(t_0)} \nabla w_i\big(X_i(t_1) + X_i(t_0)\big)\right)$$

where we used

$$\nabla w_i \equiv w_i(t_1) - w_i(t_0)$$

2.2 The simulated sample

The previous subsection decomposed the measured change in the household variable X for the household group Q into:

- C1: changes in the individual households' variable X_i
- C2: changes in the composition of the group
  - C2A: changes in the members of the group or sample
  - C2B: changes in the sample weights

Alternatively, these three components can be seen as different approaches to construct any desired change in X. We will use them to construct a "simulated" sample of households at time t_0 from the observed sample at t_1 that will describe the changes in X over time correctly. This can be useful because the HFCS survey is only conducted every three to four years, which is too long to construct an annual time series.

To give an example, consider the birth of a child within a two-member household. This can be described as a +1 increase over time in the number of household members N_i (approach C1). Alternatively, one could increase the weight of the three-member household by +1 over time and decrease the weight of a similar two-member household by -1 (approach C2B). Similarly, one could remove one original three-member household with weight 1 from the simulated sample at t_0 and remove a similar two-member household with weight 1 from the original sample at t_1 (approach C2A). All three approaches yield the same result. However, because such an exactly similar two-member household will in general not exist in the sample, it has to be constructed by removing the newborn from the original three-member household. Thus, the three approaches are equivalent in practice. Where they differ is in the decomposition of the flows according to (10).

The C1 micro simulation is the preferred approach because it most closely mimics reality. However, C1 alone cannot account for all changes. Approach C2 is necessary to account for households appearing and disappearing from the population, either by making changes to the sample or to the weights. In practice, re-weighting unavoidably leads to weight changes, while changing between samples is unavoidable over longer time intervals. In these cases, the resulting C2 flows will be a composite of different NA flows with some statistical uncertainty mixed in to complicate things further. Therefore, to obtain flows that can easily be identified and linked one-to-one to the familiar flows in the NA, one has to find a suitable combination of approaches. The central question is whether a sample at time t_1 can be adjusted to represent the population at an earlier point in time t_0. We do not provide a general answer to this question, but limit ourselves to the case of the HFCS survey. For a population of households evolving over relatively short intervals, one may assume that the populations at times t_1 and t_0 highly overlap each other. This is especially the case when the population is in a demographic "steady state" situation. This is a state where the demographic characteristics change only relatively slowly at the population level even when they change more rapidly at the individual household level. Thus, the sample at t_1, when adjusted properly, should be able to represent the population at t_0 rather well, except that we have to account for the small number of households that appeared and disappeared over time.

Normally, for a sample to be representative it must be generated by a properly randomized selection process with an associated weighting procedure. Therefore, the simulation must not destroy the random selection process. For example, consider simulating the age of household members one year into the past. This is easy to do. However, the average age in the simulated sample will be one full year below the average age in the original sample[3]. This counterfactual result indicates that the simulated sample is biased. The reason is obvious: individuals who died during the year are missing from the original sample, and therefore also from the simulated sample. The simulated sample is no longer a random selection from the entire population.

2.3 Flows

We will now discuss the three flow components separately, focusing on the variable wealth and the case of the HFCS survey.

[3] It decreases by less than one year if we remove newborns with a negative age at the beginning of the year.

C1: Changes in the variable

The change ∇X_i in the variable X_i of the individual household can be computed if we collect the data to compute ∇X_i directly or if we collect X_i at different times, for example, with a panel survey. In addition, one can benchmark the micro data at different times to the macro totals X, provided these are known, to obtain a rough estimate for ∇X_i. By summation of ∇X_i over the average group composition, we obtain component C1.

Similar to a household group, changes in X_i are the combined result of changes at the level of individual household members and changes in the composition of the household. While survey data often sufficiently covers the changes for the individual household members, they in general do not cover changes in household composition. It is therefore difficult to simulate changes in the composition of the individual household. In addition, is it difficult to estimate the impact of such changes on both X_i and ∇X_i.

C2: Changes in composition

Changes in the composition of a group can occur for two reasons: changes in the households that are a member of the group and changes in the sample weights, giving rise to components C2A and C2B.

C2A: Changes in the members of the group or sample

Mobility of households
To simplify things, we will assume every household is member of the resident population P at all points in time

$$\sum_{i \in P(t)} \quad \rightarrow \sum_i$$

Therefore, changes to population composition have to be captured solely through changes in the weights. In this case, the composition of different groups within the population only changes because certain households move between groups.

A change in an individual household variable may lead the household to move from one group to another. For example, by calculating changes in individual household wealth, one can simulate the changes in the composition of a wealth group. We will call these changes “social mobility”. Note that social mobility is relative: a household can change groups even without experiencing a change itself. For example, a middle-income household can become a higher-income household if there is a household in the higher-income group whose income drops below that of the household in the middle-income group.

Changes in the sample
Because the sample can describe the population only accurately for a limited time interval, it might be necessary to change samples over time. This can be done gradually or abruptly. The simplest way is to keep the sample S_1 collected at t_1 constant over the entire interval and switch abruptly at t_0 to another sample S_0 if necessary. In that case, formula (10) applies to the current sample S_1 only, for the entire interval, but we have to record an extra instantaneous flow as the result of the switch. This extra flow shall be called a “statistical” flow. It is also given by (10), but instead of using backward averages one uses only the values at t_0

(11) $\quad C2A_{Q}^{statistical} = \sum_{i \in Q(t_0), i \in S_1} w_i(t_0)X_i(t_0) - \sum_{i \in Q(t_0), i \in S_0} w_i(t_0)X_i(t_0)$

There are clear advantages to this approach. If the sample does not change, component C2A only accounts for mobility and can be denoted $C2A^{mobility}$. The effect of this mobility on the variable X cancels out at the level of the total population.

(12) $\quad C2A_{P}^{mobility} = \sum_i \overleftarrow{w}_i \overleftarrow{X}_i - \sum_i \overleftarrow{w}_i \overleftarrow{X}_i = 0$

In addition, if the two samples are both benchmarked at t_0 to the same population total $X(t_0)$, then the $C2A^{statistical}$ component will also cancel out at the total population level. This makes formula (10) easier to interpret.

C2B: Changes in the weights

Changes in the weights result naturally from the weighting procedure because re-weighting to another point in time will result in weight changes whenever the values of the auxiliary variables change over time. The auxiliary variables are often demographic variables, like age, number of people, number of households and household composition. The weighting procedure is in general a mathematical optimization method. It can be interpreted as an attempt to represent the entire population through appropriate combinations of sampled households. Such combinations of households cannot be expected to behave over time in a realistic way and formula (10) will in such case generate flows that cannot be easily interpreted or linked to the NA. Therefore, the re-weighting method is preferably used only for variables of interest that cannot be easily simulated with method C1, for example, household composition or income.

Only those weight changes that occur in reality, i.e. when the number of a certain type of household in the group changes, can be unambiguously interpreted. Such possible compositional changes are depicted in figure 1:

- appearance/disappearance of households,
- mobility between groups,
- immigration/emigration (mobility between countries),
- splits/mergers of households.

The dynamics depicted here consist of single events. Dynamics consisting of multiple consecutive events may also occur, for example, a household entering and subsequently leaving a group or a merger between three households. We may ignore such dynamics if their probability is sufficiently small or the interval is sufficiently short.

The biggest challenge pose those households that cannot be observed at t_1 because they disappeared or emigrated. Ideally, such missing households should be imputed and counted as changes to the sample. We will not do this here. Instead, we will keep the same sample and try to account for the missing households by adjusting the weights. For this to work, we have to assume that for each household that disappears a representative household remains in the sample.

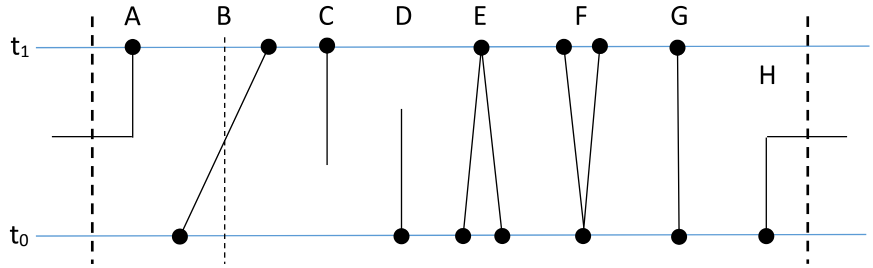

Figure 1. Changes to the members of a group or population over time. Black dots are households existing at times t_1 and t_0; solid black lines represent time-lines. Bold dashed lines represent the borders of the country; the thin dashed line separates two household groups. A immigration; B mobility between groups; C appearance of a new household (Normally, households do not appear out of nowhere but come into existence through a split.); D disappearance of a household; E a merger between two households; F a split; G no event; H emigration.

Natural population change due to births and deaths of people as well as immigration and emigration (net migration) lead over time to a change in the number of households, while movements within the country between household groups, which will be called "social mobility", lead to changes in the number of households within a group. Thus, one can decompose weight changes into four components plus a residual

(13) $$\nabla w_i = \nabla w_i^{\text{natural change}} + \nabla w_i^{\text{net migration}} + \nabla w_i^{\text{mergers/splits}} + \nabla w_i^{\text{mobility}} + \nabla w_i^{\text{residual}}$$

and, if desired, one could write

$$\nabla w_i^{\text{natural change}} = \nabla w_i^{\text{births}} + \nabla w_i^{\text{deaths}}$$

$$\nabla w_i^{\text{net migration}} = \nabla w_i^{\text{immigration}} + \nabla w_i^{\text{emigration}}$$

$$\nabla w_i^{\text{mergers/splits}} = \nabla w_i^{\text{mergers}} + \nabla w_i^{\text{splits}}$$

Each of these will make their own contribution to C2B.

Natural change

This is the change in number of households (not individuals) due to natural population change, i.e. births and deaths of individuals (not households). Population change can be decomposed into natural population change and net migration.

(14) $$\nabla N = \sum_i \nabla(w_i N_i) = \nabla N^{\text{natural population change}} + \nabla N^{\text{net migration}}$$

Focusing on natural change, we can write

(15) $$\nabla N^{\text{natural population change}} = \nabla N^{\text{births}} + \nabla N^{\text{deaths}}$$

In general, the birth of a person does not result in the appearance of a household; instead, a person is born within an already existing household. Therefore, births do not have a direct impact on the number of households, and therefore should, if properly accounted for, not have an impact on the weights. Births are better accounted for by micro simulating the number of household members over time, i.e. approach C1 rather than approach C2B, or they could be ignored altogether in which case

$$(16) \quad \nabla w_i^{\text{births}} = 0$$

Deaths only affect the number of households if all members in the household die. Therefore, to account for the effect of deaths we will have to use a mixed approach. Combining (16) and (15) with (14) gives

$$(17) \quad \nabla N^{\text{natural population change}} = \sum_i (\nabla w_i^{\text{deaths}} \overleftarrow{N}_i + \overleftarrow{w}_i (\nabla N_i^{\text{births}} + \nabla N_i^{\text{deaths}}))$$

Here $\nabla N_i^{\text{births}}$ and $\nabla N_i^{\text{deaths}}$ are the change in the number of household members over the period due to births and deaths, using approach C1. This approach can only be used for households observed in the sample. To properly calculate these changes in the number of households, one must be able to identify households with newborn children in the sample. Aggregate data on mortality rates by age and gender are also needed. Since households that disappeared between t_0 and t_1 can no longer be observed, they are accounted for by a change in weight $\nabla w_i^{\text{deaths}}$ of similar observed households, i.e. approach C2B.

Normally, births have no immediate impact on the wealth of a household. We will also assume that the wealth of a deceased member remains within the household. This is a simplification, but if any wealth flows associated with a death occur before the disappearance of the household, they could in theory be observed in the HFCS. Here we try to simulate only the impact of the disappearing households because they cannot be observed. Since under normal conditions wealth cannot disappear, the disappearance of a household must give rise to a wealth transfer to other households, i.e. an inheritance. Such transfers are recorded as “Other capital transfers” (D.99) in the NA. Normally, a survey cannot directly observe inheritances paid, only the ones received. Using $\nabla w_i^{\text{deaths}}$ in formula (10) will give us an estimate of the inheritances paid. The inheritances received do not have to be estimated because they are already recorded in the HFCS.

Net migration
Net migration is the change in the number of households due to the difference between immigration and emigration. Approach C2B can be used to simulate this change. One needs to know the characteristics of immigrant and emigrant households, so that the appropriate weight changes can be applied to similar households in the sample.

$$(18) \quad \nabla N^{net\ migration} = \sum_i (\nabla w_i^{immigration} + \nabla w_i^{emigration}) \overleftarrow{N}_i$$

Unfortunately, emigrant households will be missing from the sample. A workaround is to assume that such households, since they are the immigrants of another country, have similar characteristics as immigrant households. However, this ignores asymmetries between countries. For example, pensioners tend to migrate to warmer, cheaper countries and poor households to more prosperous countries. Once we identified immigrant and emigrant households in the sample (see section 3.5) we must assign them the correct changes in weights so that (18) holds. This requires knowledge of the total number of immigrants and emigrants.

Mergers/splits
Individuals are born, but households normally come into existence through a split. A household can cease to exist when its last member dies or when it merges with another household. Mergers and splits, for example children moving out, marriages and divorces, affect the number of households in the population. However, other variables like the number of individuals or the wealth in a group are in general not affected. If the newly created households belong in another group than the original households; then the

result is mobility between groups. Strictly speaking, the newly created households have not moved between groups; it is only the people that have moved. We will still call it mobility, though.

For any extensive population variable X, i.e. a numerical variable that is additive for subsystems, in this case individuals, mergers or splits will not affect the population total.

(19) $$\sum_i \nabla w_i^{\text{mergers/splits}} \overleftarrow{X}_i = 0$$

Because mergers/splits have the potential to affect both the composition of a group and the wealth of the household, ideally they should be properly accounted for. One possible way to do this is with transition matrixes [Prais 1955].

A split might result in income or capital transfers between the two separated households, for example alimony. To simulate such transfers is difficult. Fortunately, one does not have to bother with these as long as one simulates only one year into the past because income transfers paid and received between households over the past year are properly collected in the HFCS.

Mobility
Households moving between groups can also result in a change in weights. Note that the $C2A^{mobility}$ component already records mobility resulting from micro simulation of the variable defining the group, for example, the micro simulation of wealth in combination with a wealth quintile. In addition, $C2B^{mobility}$ records any remaining mobility when the characteristic group variable is not simulated and its change at micro level is not known. For example, the micro simulation of wealth in combination with an income quintile. If the change in the characteristic variable is known, it may be assumed that $C2A^{mobility}$ takes account of all mobility. If in such a case $C2B^{mobility}$ is not zero, this is an indication that something is wrong, either in the simulation, the recalibration or in the assumptions.

Mobility between groups cannot change the total number of households. Therefore, the following should hold

(20) $$\sum_i \nabla w_i^{mobility} = 0$$

Mobility should also not change the population total of our variable of interest X_i

(21) $$\sum_i \nabla w_i^{mobility} \overleftarrow{X}_i = 0$$

When this equation does not hold even approximately, the weight changes do not represent pure mobility and should be classified under “residual”.

Residual
The final term $\nabla w_i^{\text{residual}}$ includes all weight changes that occur because of statistical uncertainty or other problems. For this term, (19), (20) and (21) will not hold exactly. If desired, the resulting flows could be isolated by classifying them as “statistical” flows and grouped together with (11).

3. Methodology

3.1 The HFCS survey

We want to apply the theory described above to the Luxembourg HFCS. For a general description of the HFCS see [ECB 2016]; for a description of the Luxembourg HFCS in particular see [Chen et al. 2020]. Data availability is summarized as follows:

- Survey data for a sample of households is available for time t_1, but not for time t_0 one year earlier. The sample weights $w(t_1)$ are already calibrated to certain population totals at time t_1.

- The HFCS data can provide estimates for the wealth of each individual household observed at time t_1. Since there is no panel component available, we need a conceptual link with the NA to compute changes in the individual household wealth over the year using equation (25). The National Statistical Institute (NSI) also publishes the aggregate number of deaths and births and the number of immigrants and emigrants. NA data for the year under consideration can also be useful to serve as benchmark.

3.2 A baseline approach

Our baseline simulation method is deterministic. It is also "closed" since we only use the existing sample. However, enhancing the base line method with stochastic simulation or with "open" methods that add imputed households is possible.

Linking the survey to the NA concepts and micro simulating wealth
First, one has to calculate the variable of interest, in our case net worth $B.90_i$, for each household in the sample at t_0 from the observed variables at t_1. This is done using formula (22). The $\Delta B.90_i$ can be computed when we link the survey to the NA concepts and use (25). For example, based on the income, capital transfers and consumption reported in the HFCS we can estimate a household's "Savings and capital transfers" $B.101_i$.

Micro simulation of population growth
Second, one has to calculate the weights at time t_0 from the original weights at t_1. The original weights will be kept as much as possible untouched. We will not perform a re-weighting because the resulting flows will be difficult to interpret. Instead, we simulate only weight changes due immigration, emigration and the disappearance of households due to the deaths of all its members. There is no need to fully simulate natural population growth because we assumed that births and deaths have no impact on the wealth of a household. A full micro simulation of population growth is only necessary if one re-weights to the changes in demographics and wishes to avoid a change in the weights as a result. Only when a household disappears entirely through the death of all its members we have to account for a wealth flow.

Mergers and splits, like marriages and divorces, are difficult to simulate and their impact on the wealth and wealth flows of the household difficult to estimate. Use of transition matrixes would be necessary. This is, however, difficult to do and since we lack the data for it, we will not try it here and instead ignore this component; we therefore set

$$\nabla w_i^{\text{mergers/splits}} \cong 0$$

Because splits and mergers can have a significant impact on the wealth of a household, we make an error by ignoring them. Thus, our description will necessarily be incomplete.

We also take the following shortcuts:

- For simplicity, we keep the sample constant over the year and, if necessary, switch abruptly to another sample at the beginning of the year.

- Because we simulate only the variable wealth, we can observe mobility flows $C2A^{mobility}$ between wealth groups only. Therefore, we will considers only wealth groups, in which case all mobility is captured within component $C2A^{mobility}$ and we can set the $C2B^{mobility}$ component equal to zero

$$\nabla w_i^{\text{mobility}} \cong 0$$

The result of the above two simulation steps is the simulated sample for t_0.

Re-weighting
One could re-weight to the auxiliary variables at t_0.We choose not do this because we would not know how to interpret the resulting flows. Thus, we can set

$$\nabla w_i^{\text{residual}} \cong 0$$

Benchmarking to the NA
To calculate a time series of wealth it helps to benchmark to the macro totals of the NA. We use the "proportional allocation" method to benchmark to the NA.

Time series computation
Time series are computed by benchmarking the nearest HFCS wave to the NA for each year. These benchmarked samples are then micro-simulated one year back in the past. The resulting annual time series segments are joined together by the instantaneous flow $C2A^{statistical}$ (11). This instantaneous end-of-year flow can be attributed 100% to the preceding year, to the ensuing year, or 50%/50% to both. We choose to attribute it fully to the ensuing year for simplicity.

These steps are discussed in more detail below.

3.3 Linking the survey with the National Accounts

Micro and macro variables are linked at two levels: the conceptual and the numerical level. First, we have to establish a conceptual link between the variables in the HFCS and in the NA. Although not always straightforward, there are some satisfactory methods [ECB 2020; Zwijnenburg 2016]. The conceptual link allows us to compute numerical values for the relevant NA variables at micro level, for every individual household. Aggregating over households (using appropriate weights) will then produce an aggregate value for the NA household sector based on micro data. We can compare this estimate to official NA figures. This will generally reveal a "gap" because the NA are based on other data sources. By correcting the individual households data or their weights appropriately, one can remove the gap and thus benchmark micro to macro data. However, we will not discuss this last step here but only refer to some of the literature on this subject [Coli et al. 2022; ECB 2020; Zwijnenburg 2016].

Our wealth concept aligns with the "Net worth" (B.90) concept from the NA, which is equal to all household financial and non-financial assets minus household financial liabilities.

(22) B.90 = Assets – Liabilities

In the NA, "Changes in net worth" (B.10) are the result of the "Savings and capital transfers" (B.101), "Other changes in the volume of assets" (B.102) and "Holding gains and losses" (B.103) over the time interval $[t_0, t_1]$; $t_0 < t_1$.

(23) $B.90(t_1) – B.90(t_0) = B.10 = B.101 + B.102 + B.103$

Item B.101 "Savings and Capital transfers" is roughly the difference between Income and Consumption and is a non-financial NA concept. The link between the non-financial NA and the HFCS, as described in [Zwijnenburg 2016, Zwijnenburg et al. 2021], allows us to calculate this item for each individual household from it's income, consumption and capital transfers. The link between the financial part of the NA and the HFCS, as described in [ECB 2020], allows us to estimate individual household balance sheets of financial and non-financial assets and thus "Net worth" (B.90), see formula (22). Individual household "Holding gains and losses" (B.103) can be estimated by applying appropriate revaluation indexes to each instrument on the balance sheet. Such revaluation indexes can be directly computed from NA data or market indexes can be used. Estimating "Other changes in the volume of assets" (B.102) at the micro level is more challenging. This item contains unusual flows like write-offs or statistical reclassifications. Since we lack any other information, we can simply distribute the macro totals from the NA across households in proportion to their holdings. The only component of B.102 that we have to estimate directly is "Changes in classification" K.6 since this is linked to net migration and therefore reflects changes in population composition. K.6 is often ignored in the NA, but an estimate could be provided by our micro estimates of net migration. We return to this later.

Difficulties arise mainly from variables that exist in the NA but not in the HFCS, or vice versa. One example is house sales: although the HFCS contains a question about house purchases, it does not ask about sales. Fortunately, house sales do not change households' net wealth. NA also contain many imputed flows that are not observable in real life and therefore do not exist in the HFCS. One example is Financial Intermediation Services Indirectly Measured (FISIM); another is the production and consumption of housing services associated with owner-occupied dwellings. Fortunately, these imputed flows cancel out at the end of the non-financial accounts and have no effect on Net worth[4].

[4] The only imputed flow that might have an effect is "Change in pension provisions" (D.8). If there is a problem linking this item to the HFCS, it could always be removed from the National Accounts, i.e. removed from the "Net savings" (B.8n). If "Pension provisions" AF.63 are also removed from Net worth, the NA framework remains fully consistent. In our case, we have kept pension provisions as part of wealth.

Because the NA provide a complete picture of wealth, there are not many variables relevant to wealth in the HFCS that are not in the NA. Any such variables involve flows or financial positions between households that cancel out at the macro level. One example is the sale of second-hand consumer goods. Because goods held by households for consumption are not assets in the NA, sales and purchases of second-hand consumer goods are recorded in the NA as negative or positive consumption. Therefore, they may affect the wealth of the individual household, but at the aggregate level of the household sector these flows cancel out, so they are generally ignored. A similar example is "Other capital transfers" (D.99) between households (mainly inheritances). We will return to these later.

3.4 Micro simulation of wealth

From the sample at t_1 we have to create a sample at t_0, one year earlier. This is done by simulating for each household values for Net worth (B.90) at time t_0. In this way, we create a "simulated sample".

(24) $\mathrm{B.90_i(t_0) = B.90_i(t_1) - \nabla B.90_i}$

In the case of wealth, our NA framework and the conceptual link with the HFCS provides us with estimates for all the necessary components of stocks and flows at the micro level:

(25) $\mathrm{\nabla B.90_i} = \mathrm{B.101}_i\ (excluding\ inheritances\ paid\ D.99) + \mathrm{B.102}_i\ (excluding\ K.6) + \mathrm{B.103}_i$

The HFCS does not include inheritances paid, recorded in the NA under "Other capital transfers" D.99, because deceased households cannot be surveyed. However, there is a question in the HFCS on inheritances received. The NA generally assume that inheritances paid and received cancel out at the population level.

Summing (25) over the group yields component C1 for the group; summing over the entire population yields an estimate for the NA aggregate.

(26)
$$\mathrm{C1_P} = \frac{1}{2}\sum_\mathrm{i}\left(\mathrm{w_i(t_1) + w_i(t_0)}\right)\mathrm{\nabla B.90_i}$$
$$= (\mathrm{B.101 - inheritances\ paid\ D.99}) + (\mathrm{B.102 - K.6}) + \mathrm{B.103}$$

3.5 Micro simulation of population growth

Immigration

The Luxembourg NSI publishes data on the total number of persons that immigrated over the year. This total should be equal to

(27) $\mathrm{\nabla N}^{immigration} = \sum_\mathrm{i} \mathrm{\nabla w_i}^{immigration}\overleftarrow{\mathrm{N}}_\mathrm{i}$

The HFCS question "In which year did you arrive in Luxembourg" helps to identify immigrant households. Let IM denote the group of households that reported they immigrated during the period. One could expect the weighted sample of immigrant households to match aggregate immigration

(28) $\nabla N^{immigration} \cong \sum_{i \in IM} w_i(t_1) \overleftarrow{N}_i(t_1)$

where we have set $w_i(t_0)=0$ for immigrants. In practice, this will not be the case. The following is a simple way to re-calibrate the weights, depending on the ratio *f*

(29) $f = \frac{\nabla N^{immigration}}{\sum_{i \in IM} w_i(t_1) \overleftarrow{N}_i(t_1)}$

$$
\begin{aligned}
f = 1: &\quad \nabla w_i^{immigration} = f w_i(t_1);\ w_i(t_1) \to \nabla w_i^{immigration};\ w_i(t_0) = 0 \\
f < 1: &\quad \nabla w_i^{immigration} = f w_i(t_1);\ w_i(t_1) \to w_i(t_1);\ w_i(t_0) = w_i(t_1) - \nabla w_i^{immigration} \\
f > 1: &\quad \nabla w_i^{immigration} = f w_i(t_1);\ w_i(t_1) \to \nabla w_i^{immigration};\ w_i(t_0) = 0
\end{aligned}
$$

If f > 1, it is also possible to replace IM with a larger group of households that immigrated during the previous two or three years; this will decrease *f*. After re-calibration, (27) should hold. If one is comfortable with negative weights, the f<1 case could be applied everywhere.

Emigration

Emigrated households cannot be surveyed and therefore are missing from the observed sample. In the absence of other sources of information, one may assume that emigrants and immigrants are similar. Then one could apply the same method as in (29) but now with total net migration (immigration minus emigration). Because net migration is much smaller than gross immigration, *f* will probably be smaller than one. This is preferable because then the original weights $w_i(t_1)$ are not affected.

Wealth flows due to net migration can now be computed for each group and for the population. In the NA this flow is recorded under "Changes in classification" K.6.

(30) $C2B_P^{\text{external growth}} = \frac{1}{2} \sum_i (\nabla w_i^{immigration} + \nabla w_i^{emigration}) (B.90_i(t_1) + B.90_i(t_0)) = K.6$

Normally, K.6 is not recorded in the NA, probably because annual net migration is relatively small in most countries. In Luxembourg, however, it is roughly 2% of the population. Our method provides a first estimate for this aggregate NA figure.

Births

Since the HFCS collects the ages of all household members, this allows us to identify new-born children. Let NB stand for the group of households with newborns and $\nabla N_i^{\text{births}}$ for the number of births in household i. Then, with (17) one expects the following to hold

(31) $\nabla N^{births} \cong \sum_{i \in NB} \overleftarrow{w}_i(t_1) \nabla N_i^{\text{births}}(t_1)$

One could calibrate weights with the same method (29) as for immigration. However, since births have no direct impact on household wealth, this is not necessary.

Deaths
Deaths shift households into groups with fewer household members. The probability for a given household member to survive past a given age is given by survival rate sr(a), which mostly depends on the age and gender of the individual. Some NSI publish survival rates, which can otherwise be derived from the age distribution of the observed sample, by assuming it is at a steady state. This requires the sample's age distribution to decline smoothly. Otherwise, it might help to combine several age groups. From survival rates we can compute mortality rates both backwards and forwards as

(32) $$\overleftarrow{mr}(a) = \frac{sr(a-1)-sr(a)}{sr(a)}$$
$$\overrightarrow{mr}(a) = \frac{sr(a)-sr(a+1)}{sr(a)}$$

where a is the age. Summing mortality rates for all N_i members in every household gives an estimate of the change in the population due to deaths over a given year

(33) $$\nabla N^{deaths} = -\sum_{i \in P(t_0)} w_i(t_0) \sum_{j=1}^{N_i} \overrightarrow{mr}_{ij}(t_0)$$

where the mortality rates depend on time through the age $a_j(t)$ of the member. For this, we need to know the age and gender of every household member. The sample also needs to be large enough so that most age and gender combinations are present. In addition, since mortality rates are statistical averages, results are only accurate for a large population. Comparing the (33) estimate to independent macro data on the number of deaths allows us to calibrate either the weights or the mortality rates. However, as long as differences are within the expected statistical variation, calibrating is not advisable.

Formula (33) gives an estimate of the number of individual deaths. However, we are interested in the number of households that disappear due to these deaths. Multiplying forward mortality rates by the individual members can yield an estimate of the number of a particular household i with N_i members disappearing.

(34) $$\nabla w_i^{\text{deaths}} = -w_i(t_0) \prod_{j=1}^{N_i} \overrightarrow{mr}_{ij}(t_0)$$

The problem with (34) is that we only observe households at t_1. From this, we could compute the composition of the entire sample at t_0, but this is tedious, especially if we take account of mergers and splits. Such a procedure injects a lot of statistical uncertainty and works only for large populations, so it may not work well with a limited sample. Instead, we will make a demographic "steady state" assumption. This means that the number of households of the particular type i disappearing is constant over time. In this case, one can approximate (34) using the values observed at t_1

(35) $$\nabla w_i^{\text{deaths}} \approx -w_i(t_1) \prod_{j=1}^{N_i} \overrightarrow{mr}_{ij}(t_1)$$

For shorter periods and low mortality rates, (35) is only significant for single-member households above a certain age.

This change in the weights seems to contradict the “steady state” assumption. The steady state can only exist if we also account for all the changes in the age and composition of the households. The households that disappear are exactly replenished by the ageing, decay, mergers and splits of other households. Here we ignore such dynamics, either because they do not affect the wealth of the household or because they are too difficult to simulate. It would be much simpler to keep the weights constant in the steady state, but then one misses all the wealth flows associated with the mentioned household dynamics.

Plugging (35) into (10) yields an estimate of the inheritances paid by a household group or by the total population

$$(36) \qquad \mathrm{C2B}_{\mathrm{P}}^{\text{internal growth}} = \frac{1}{2}\sum_{\mathrm{i}} \nabla \mathrm{w}_{\mathrm{i}}^{\text{deaths}} \left(\mathrm{B.90}_{\mathrm{i}}(\mathrm{t}_1) + \mathrm{B.90}_{\mathrm{i}}(\mathrm{t}_0)\right) = \text{ inheritances paid D.99}$$

Inheritances paid should match inheritances received at the aggregate level. However, differences can be attributed to transfers from or to non-resident households and/or charity organizations, so we decided not to match inheritances paid and received.

3.6 Re-weighting

Re-weighting is an optional step in the process. By re-weighting the simulated sample to match auxiliary data on population totals at time t_0 we obtain weights $w(t_0)$. These re-calibrated weights should be as close to the original weights $w(t_1)$ as possible. The weighting procedure of the HFCS is described in [Girshina 2017]. We opted not to do any re-weighting because we would not know how to classify the resulting flows. More research is needed on this subject.

3.7 Benchmarking to the National Accounts

The sum of components (26), (30) and (36) provides an independent estimate for the NA item “Changes in net worth” (B.10). One could benchmark these estimates to the official NA figures by making corrections at micro level. This has the same advantage as calibrating the sample weights to known macro totals. It is especially useful to produce a time series because the NA are available on a quarterly basis, while the HFCS surveys are only available every 3 years or so. Possible methods for benchmarking the HFCS to the NA are described in detail in [ECB 2020; Zwijnenburg et al. 2021; Chakraborty et al. 2019]. We choose the “proportional allocation” method to remove the gap.

3.8 Time series computation

Because there is survey data available for it, the reference year can be readily simulated, while for other years, in the past or in the future, this is more difficult. In contrast, benchmarking a sample to the NA can easily be done for any year. Simulation of other years than the reference year is possible, though, if one assumes that each household’s income and consumption remains relatively unchanged. In this respect, it helps to benchmark the micro data to the NA so that economy-wide changes in income and consumption are accounted for. Alternatively, one could try to micro-simulate income and consumption over time. The distribution of unemployment over time is especially relevant in this respect.

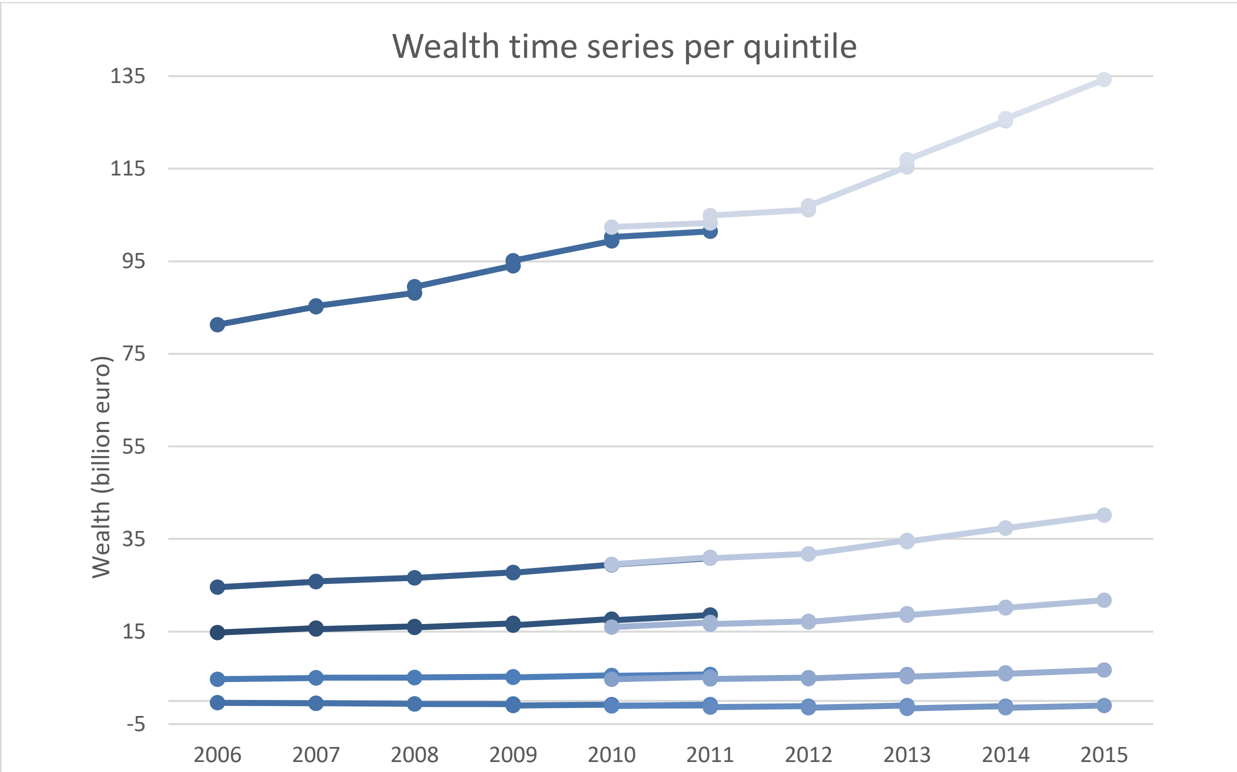

Figure 2. Experimental annual wealth time series estimates for five wealth percentiles. Note that annual segments in general do not match; especially in the middle between the 2009 and 2013 HFCS waves[5].

Forward or backward simulation of one single survey sample will create a short time series. Those time series fragments can be stitched together to form a longer time series. This is shown in figure 2 for the five wealth quintiles. Here, for every year the micro data of the nearest survey have been benchmarked to the macro totals from the NA. This results in a differently benchmarked sample for each year. These samples are then simulated one year into the past. Notice that the resulting annual segments do not match perfectly. Discrepancies are classified as $C2A^{statistical}$ flows that result from switching between benchmarked samples, according to formula (11). A possible improvement would be to change between samples in a more gradual way. One could take the weighted average between two subsequent waves, weighting the first wave's sample with a weight that decreases linearly over time, while weighting the second wave's sample with a linearly increasing weight. Such interpolation would not get rid of the $C2A^{statistical}$ flows but would distribute them more evenly over the years.

[5] Reference data for the year 2009 was collected during the Luxembourg 2010/2011 HFCS (wave 1); reference data for the year 2013 was collected during the 2014 HFCS (wave 2).

When results for a given year differ based on two different survey waves, this normally results from the statistical uncertainty in the survey samples or our simulation might be wrong. Therefore, one could fine-tune the simulation methodology to minimize the statistical discrepancies. One could also choose to eliminate the statistical discrepancy, because it does not have an equivalent in real life, similar in the way statistical discrepancies are often eliminated in the NA. One way is to distribute the discrepancy evenly over the other components. We prefer to report it separately under "statistical residual", see figure 3, because it gives an indication of the uncertainty in the figures. Note that if the micro wealth stocks and flows are benchmarked to the NA then this statistical artefact disappears at the macro level.

3.9 Wealth growth rates

Another way to present the results is as annual wealth growth rates per wealth quintile, see figure 3.

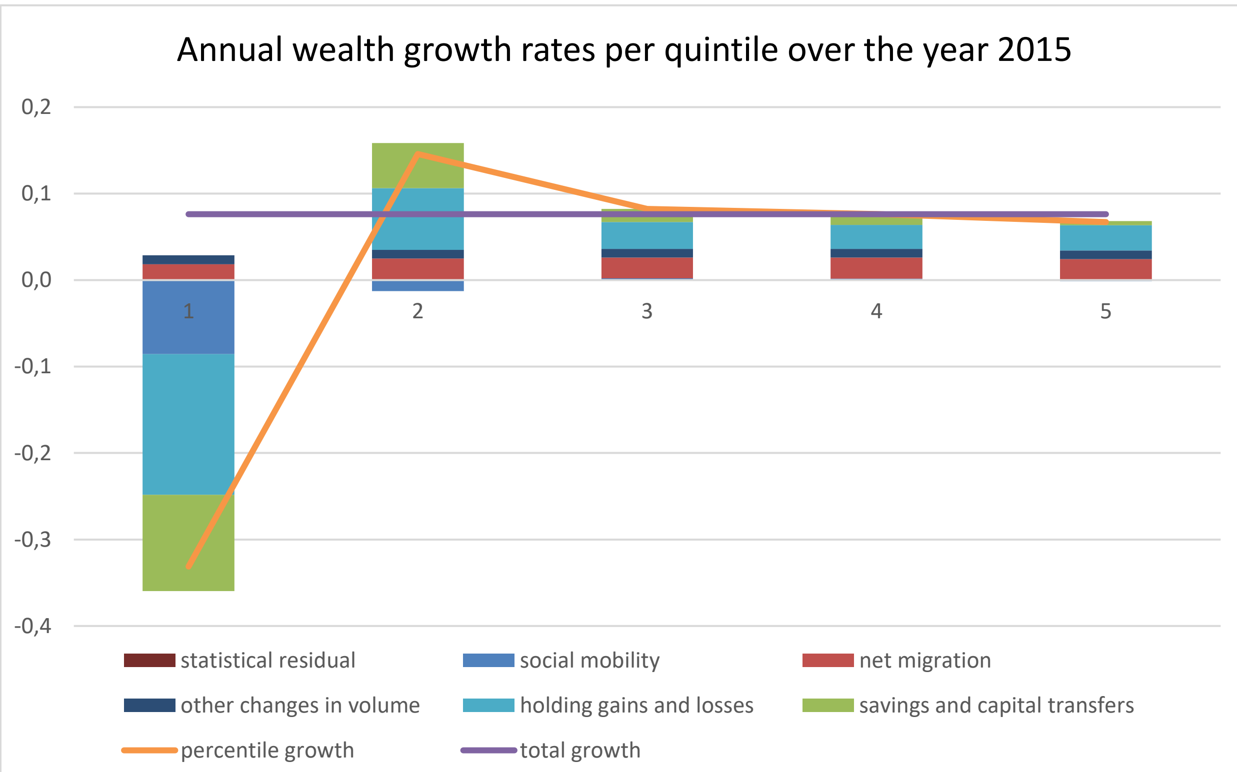

Figure 3. Experimental annual wealth growth rates per wealth quintile broken down by major components. Note that, because net wealth, i.e. the denominator in formula (37), for the first quintile is negative, a negative growth rate here means actually an increase in net wealth.

Such growth rates, here defined as

$$(37) \qquad r_Q = \frac{\nabla B.90_Q}{B.90_Q(t_0)}$$

can be decomposed into their various components, thus shedding some light on the main drivers behind the dynamics of wealth and wealth inequality. Growth rates can be more robust against the "missing rich" problem, see [Chakraborty et al. 2019], than other indicators.

Also shown in figure 3 is the wealth growth rate for the total population (straight line). The difference between percentile growth and total growth, i.e. the growth rate deviation, for the fifth quintile approximates the growth in the Gini coefficient rather well for Luxembourg, see figure 4.

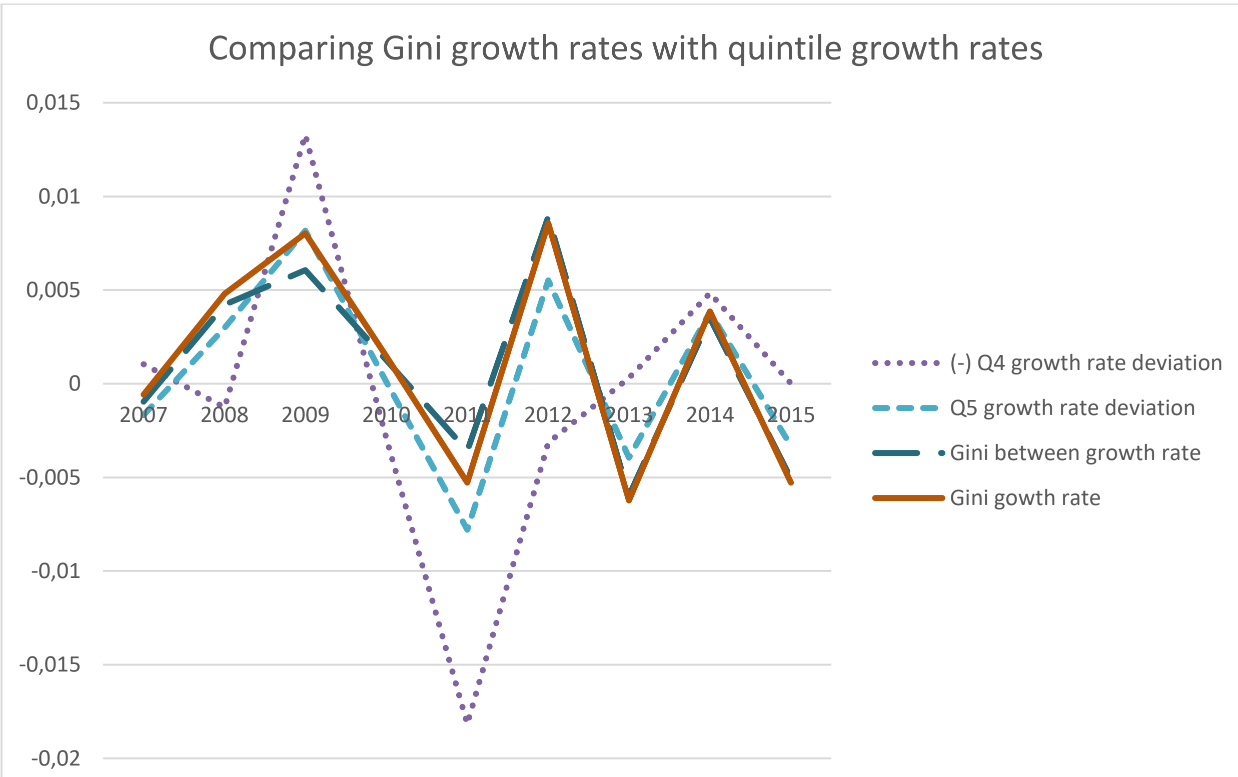

Figure 4. Experimental wealth Gini growth rates for Luxembourg for the years 2007 to 2015 based on our simulation method. Shown for comparison are the $G^{between}$ growth rate and the growth rate deviation of the fourth and fifth quintile. (The first three quintiles show larger and more asynchronous swings than the fourth and have been left out because they would not fit.) The growth rate deviation of the fourth quintile is shown with a minus sign because it correlates negatively with Gini growth. Note how the Gini is dominated by the $G^{between}$ and the fifth quintile.

The explanation is not immediately obvious. One could argue that in general the Lorentz curve follows some family of curves. For example, income and wealth distributions are often approximated by a combination of lognormal and Pareto distribution [Vermeulen 2016]. It appears to be a property of the lognormal and Pareto distribution with a high (above 0,5) Gini coefficient that the growth in the Gini and in the fifth quintile follow each other closely. In addition, the fifth quintile seems to follow the lognormal or Pareto curves more faithfully, while other quintiles, especially the first, are more out-of-line.

Another part of the explanation could be that in Luxembourg the fifth quintile rather dominates the calculation of the Gini wealth coefficient. Consider the usual formula for the decomposition of the Gini coefficient G

(38) $$G = G^{between} + \frac{1}{5}\sum_{Q=1}^{5} w_Q\, G_Q^{within} + G^{residual}$$

where $G^{between}$ is the Gini coefficient between the quintiles and G_Q^{within} is the Gini coefficient within each quintile Q, to be weighted by the share of total wealth w_Q held by the quintile. In the case of wealth groups and a wealth Gini, $G^{residual}$ is zero. In Luxembourg, G_5^{within} is large compared to the other quintiles and its weight w_5 is also high, meaning that the fifth wealth quintile dominates the Gini coefficient. Based on experimental distributional data computed by the ECB, other European countries show a similar close correlation between the Gini and the top percentile. Therefore, a presentation as in figure 3 for such countries would allow to see at a glance the direction in which wealth inequality is moving.

3.10 Problems encountered

This section briefly mentions the most important problems encountered when applying the above in practice.

- HFCS sample design and collection in Luxembourg is subject to some particularities. The sample is drawn from the social security register at the end of year T-1, but it is weighted to match the population composition at the end of reference year T. Then, data is collected over several months during the year T+1. This can create timing issues. For example, households that immigrated during the reference year are missing from the sample and we have to assume that the characteristics of immigrant households remain the same between T and T-1.

- Also, the HFCS target population does not completely align with the resident household sector. People living in institutions and international civil servants are missing from the social security register and therefore from the sample (except when the international civil servant is married to a spouse who is on the social security register). One must rely on the crude assumption that these households have the same characteristics as the rest of the population.

- HFCS reporting of consumption is incomplete. For example, earlier HFCS waves collected only daily consumption and ignored large one-off payments for holidays or insurance. Later waves improved on this situation, but aggregate consumption from the HFCS is typically only half of that in the NA. Thus, the estimated distribution of consumption (and therefore savings) is rather uncertain, even after benchmarking to the NA.

4. Conclusions

The dynamics of a population of households over time are complicated. Even so, the NA provide a complete and consistent description of wealth and changes in wealth for the household sector as a whole. The micro simulation method described in this paper provides a fully consistent description of the change in wealth at the micro level, although its decomposition in different flow components is necessarily an approximation. Our methodology allows us, under certain conditions, to link micro level wealth flows unambiguously to macro flows in the NA and to identify the main drivers behind the dynamics of wealth inequality based on only one single HFCS survey. This also facilitates the computation of a wealth time series. The main advantage of our method is that it guarantees full consistency between stocks and flows at micro level. More common methods, such as [ECB 2020], are consistent only at the group level. In addition, our method guarantees full consistency between the financial and non-financial DNA. Our method also has disadvantages, including its complexity and the fact that it works better for wealth than for other variables such as income. The methodology is still under development; planned future enhancements include: adding a simulation of mergers and splits, adding more HFCS waves and smoothing the time series using interpolation.

References

[1] Chakraborty R., Kavonius, I. K. Pérez-Duarte S. & Vermeulen P. (2019), "Is the top tail of the wealth distribution the missing link between the Household Finance and Consumption Survey and national accounts?", Journal of Official Statistics, 35(1):31-65.

[2] Chen Y., Mathä T.Y., Pulina G., Schuster B. and Ziegelmeyer M. (2020), "The Luxembourg Household Finance Consumption Survey: Results from the third wave", BCL Working Paper 142.

[3] Coli A., Istatkov R., Jayyousi H., Oehler F. and Tsigkas O. (2022), "Distributional national account estimates for household income and consumption: methodological issues and experimental results", Publications Office of the European Union, Luxembourg.

[4] ECB (2016), "The Household Finance and Consumption Survey: methodological report for the second wave", ECB Statistics Paper Series No 17 / December 2016.

[5] ECB (2020), "Understanding household wealth: linking macro and micro data to produce distributional financial accounts", ECB Statistics Paper Series, No 37 / July 2020.

[6] Eurostat (2013), "European System of Accounts 2010 (ESA 2010)", Publications Office of the European Union, Luxembourg.

[7] Garbinti, B., Goupille-Lebret, J. and Piketty, T. (2021), “Accounting for Wealth Inequality Dynamics: methods, estimates and simulations for France”, Journal of the European Economic Association, 19(1):620-663.

[8] Girshina A., Mathä T.Y. and Ziegelmeyer M. (2017), ”The Luxembourg Household Finance Consumption Survey: Results from the 2nd wave”, BCL Working Paper 106.

[9] Nguyen J. H. (2011), “An Introduction to Calibration Estimators”, The Waterloo Mathematics Review Volume I, Issue 2.

[10] Prais, S. J. (1955). Measuring social mobility. Journal of the Royal Statistical Society, A118, 56–66.

[11] United Nations (UN), Eurostat, IMF, OECD, and World Bank (2009), “System of National Accounts 2008”, New York.

[12] Vermeulen, P. (2016), "Estimating the Top Tail of the Wealth Distribution", American Economic Review: Papers & Proceedings, 106(5), 646-50.

[13] Zwijnenburg, J. (2016), “Further enhancing the work on household distributional data: Techniques for bridging gaps between micro and macro results and now-casting methodologies for compiling more timely results”, Paper prepared for the 34th General IARIW Conference Dresden.

[14] Zwijnenburg, J., S. Bournot, D. Grahn and E. Guidetti (2021), “Distribution of household income, consumption and saving in line with national accounts – Methodology and results from the 2020 collection round”, OECD Statistics Working Papers, No 2021/01.