A. Estimation of 2nd pillar pensions

B. Modeling of contribution probabilities

C. Modeling of future retirement probabilities

D. Extended description of the micro simulation model

Online-Appendix

A. Estimation of 2nd pillar pensions

A core question to be answered by the micro-simulation is how pension levels differ before and after the nationalization of the second pillar in 2010-2011. For this evaluation, a projection of future second pillar pensions is applied. Appendix A outlines the method, data and assumptions for this estimation. Initial 2nd pillar pensions B^{SP} of an individual i in a future year f depend on four factors: 1) the funded pension capital PC accrued until the future year of retirement f, 2) the remaining unisex life expectancy $\mathrm{LE}_{\mathrm{s,f}}$ at retirement age s in the future year f, 3) the expected interest rate $\mathrm{r}^{\mathrm{FDC}}$ of the funded scheme as well as 4) the indexation regime pi (reflecting the annual increase in pension benefits).

$$\mathrm{B}^{\mathrm{SP}}_{\mathrm{s,f,i}} = \mathrm{PC}_{\mathrm{s,f-1,i}} * \frac{(\mathrm{r}^{\mathrm{FDC}}-\mathrm{pi})}{1-\left(\frac{1+\mathrm{pi}}{1+\mathrm{r}^{\mathrm{FDC}}}\right)^{\mathrm{LE}_{\mathrm{s,f}}}}. \tag{1}$$

The pension capital PC_{f} itself is a function of $\mathrm{PC}_{\mathrm{f-1}}$ in the previous year indexed by the interest rate $\mathrm{r}^{\mathrm{FDC}}$ to year f (see **Error! Reference source not found.**). Additionally, PC_{f} reflects contributions paid to the funded scheme C in year $\mathrm{f}-1$ deducted by contribution fees F applied in year $\mathrm{f}-1$.

$$\mathrm{PC}_{\mathrm{s,g,f}} = \mathrm{PC}_{\mathrm{s,f-1,i}}\left(1+\mathrm{r}^{\mathrm{FDC}}\right) + \mathrm{C}_{\mathrm{s,f-1,i}} * (1-\mathrm{F}_{\mathrm{f-1}}). \tag{2}$$

Contributions paid during the year $\mathrm{f}-1$ are derived by multiplying gross earnings w with the contribution rate τ of this year (see **Error! Reference source not found.**).

$$\mathrm{C}_{\mathrm{s,f-1,i}} = \mathrm{w}_{\mathrm{s,f-1,i}} * \ \tau_{\mathrm{f-1}}. \tag{3}$$

Various data sources and assumptions have been used for the projection of 2nd pillar benefits. The starting point for the estimation provide 2nd pillar assets gathered from the contributors' database for the end of 2006 (in market values). 2nd pillar pension contributions paid after 2006 are based on the simulation of annual gross earnings w (outlined in section 4.1.1 in the main text). Contribution rates τ_{f} to the 2nd pillar added up to 8% until 2010. Thereafter, τ_{f} depends on the chosen reform scenario. Under the pre-reform scenario, contribution rates of second pillar members add up to 8.5% for the years after 2010. A zero contribution rate is considered in the switchback scenario. Contribution fees F applied in the calculations are set at a level of 5% over the entire projection horizon.[1] Additional costs, which occur when converting accounts into annuities upon retirement, are disregarded.[2] A crucial assumption of the calculations has to be made on the interest rate of 2nd pillar investments $\mathrm{r}^{\mathrm{FDC}}$. In line with recommendations of the European Commission, $\mathrm{r}^{\mathrm{FDC}}$ is set at 3% (net of asset management fees). The indexation parameter pi adds up to 0% (in real terms). In other words, a price indexation of 2nd pillar annuities is applied. Unisex life expectancy at retirement LE is estimated based on EUROPOP2013 assumptions.

[1] This value reflects roughly the average of operational fees in the period 2005-2009. See Hirose (2011), p. 184.
[2] These annuity costs can be sizeable. Murthi et al. (2001) estimate a 10% reduction factor of accounts at retirement for the United Kingdom due to these converting costs.

B. Modeling of contribution probabilities

Appendix B provides detailed insights into the modeling of contribution probabilities.

In 2010, Hungarians had the highest pension contribution rate among EU countries. 33.5% of gross earnings were contributed to the public pension system. The largest proportion, 24% of earnings, was transferred by employers to the PAYG system. The remaining 9.5% of gross earnings were paid directly by employees. Those who participated only in the single pillar channeled their entire contributions to the unfunded system (9.5%). Mixed pillar participants split their contributions into funded pension accounts (8% in 2010) and into the PAYG system (1.5% in 2010) – see **Error! Reference source not found.**. With the 2010 pension reform, the contribution rates for mixed pillar members changed. From 2011 onwards also the remaining mixed pillar participants channel their entire employee contributions (10%) into the 1st pillar.[3]

Table 1: Pension contribution rates for single and mixed pillar members in terms of gross earnings

Contribution rate	Single pillar member (2010)	Mixed pillar member (2010)	All scheme members, after switchback (2011)
Employer - unfunded	24.0%	24.0%	24% *
Employee - unfunded	9.5%	1.5%	10.0%
Employee - funded	0.0%	8.0%	0.0%

* These contribution rates represent the share of "social contributory tax" paid for pensions, based on the 2010 distribution of employer contribution rates. Since 2013 employer contribution rates have be increased to 27 %.

Source: own illustration.

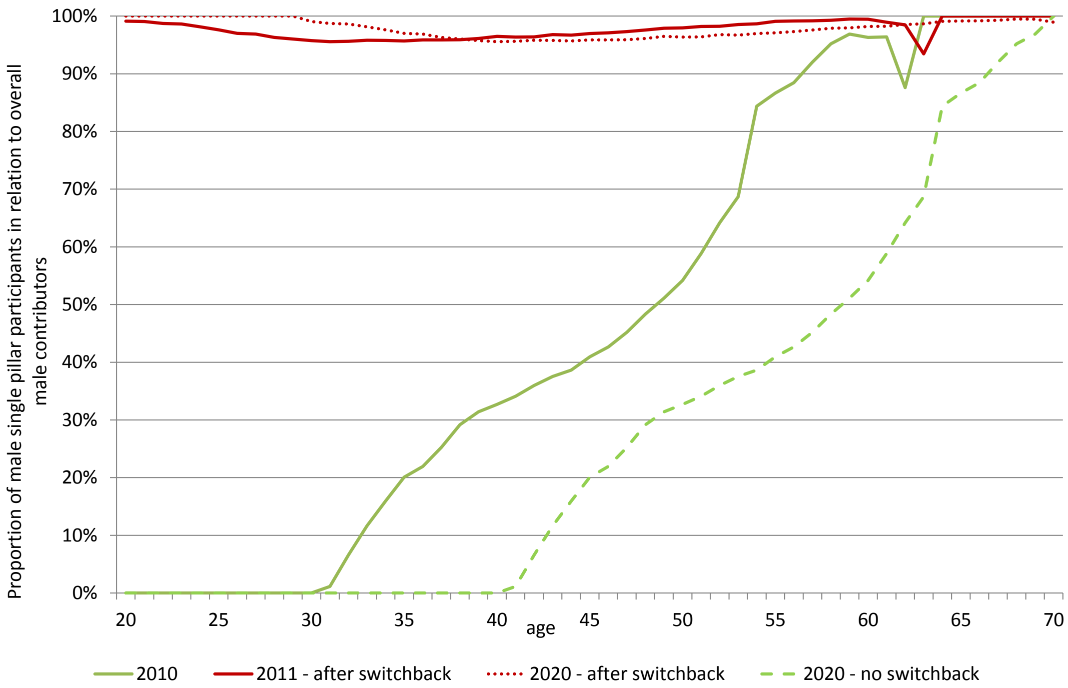
As shown in **Error! Reference source not found.**, contribution rates to the unfunded system differ widely between single and mixed pillar members under pre-2010 pension regulations. Thus, the relative size of these two groups in the contributors' population plays an important role for the projection of future pension finances (in the pre-2010 scenario). Up to 2010, the age specific participation in the single and mixed pillar system mirrored the decisions chosen during the introduction of the 2nd pillar scheme in 1997: Younger cohorts overwhelmingly opted to participate in the two pillar system which explains the relatively low single pillar participation rates of cohorts 30 to 50 in 2010 (see Figure 1).[4] The vast majority of older cohorts (aged 50+), on the contrary, chose to stay

[3] The labeling and composition of contributions changed after 2011. From 2012 onwards all employer contributions, namely pension, health and unemployment contributions, are summarized under the term "social contributory tax". In total they amount to 27% of gross earnings since 2012. In other words, the employer contributions paid before separately for health care and unemployment (in total 3%) are added from 2012 onwards to the former employer pension contributions (24%). In our calculations we neglect health and unemployment contributions.

[4] All new entrants to the pension system were automatically enrolled in the mixed pillar system under pre-2010 rules.

in the monolithic system back in 1997. Therefore, they have higher single pillar participation rates in 2010 (see Figure 1). Step by step, these cohorts would have exited into retirement with the consequence that the Hungarian public pension system would have gradually shifted from a single to a mixed pillar system over the next decades (see exemplarily dotted green line of the year 2020 in Figure 1). The 2010 pension reform reversed this transformation process. 97% of mixed pillar members switched back to the mono-pillar system at the end of 2010 (see red line in Figure 1), given the unbeneficial conditions of staying in the mixed pillar system.

Figure 1: Changing proportion of single pillar members (example of males)

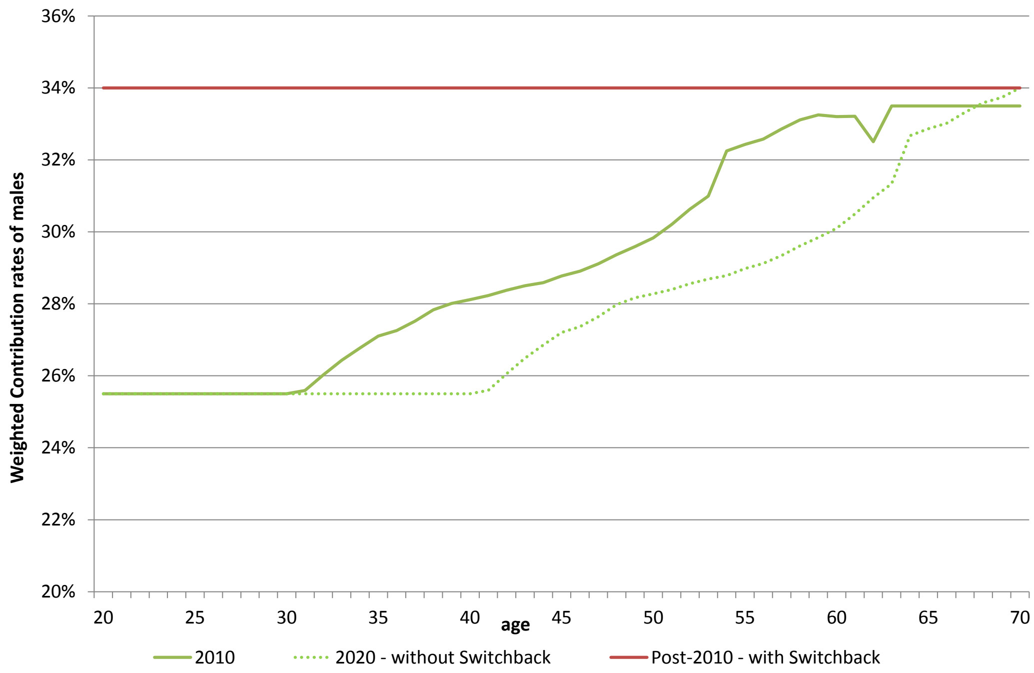


Source: own illustration based on ONYF contributors' data.

The pension model estimates average contribution rates ($\tau^{average}$) based on the participation rate in the single ($1 - p^{part\ mixed}$) and mixed pillar system ($p^{part\ mixed}$) in the respective years as well as based on the mixed pillar member ($\tau^{mixed\ pillar}$) and single pillar member ($\tau^{single\ pillar}$) unfunded contribution rates. The impact of the 2010 switchback reform on $\tau^{average}$is displayed exemplarily in Figure 2. As expected, this reform increases $\tau^{average}$ significantly. The most decisive effect on $\tau^{average}$ is visible for younger age groups. A 25-year-old contributed, for instance, on average 25.5% of his gross earnings to the PAYG system in 2010. After the switchback reform, this proportion increased to 34% in 2011 (see Figure 2).[5] Of course, with rising $\tau^{average}$ average contributions $C^{average}$ increase, too (see **Error! Reference source not found.**1 in the main paper).

[5] Please note that total employee contribution have been increased from 9.5% to 10% of gross earnings in 2011. We consider this change in both the switchback and the non-switchback scenario.

Figure 2: Weighted contribution rates of males – reflecting the single/mixed pillar participation[6]



Source: own calculation.

The starting value of $p_{s,g,f}^{contrib}$ (see Eq. 11 in the main text) reflects the number of contributors Z relative to the overall population sizes P in each cohort and for both genders (see **Error! Reference source not found.**). Up to the age of 45 years we keep these contribution probabilities constant over time.[7]

$$p_{s,g,f}^{contrib} = \frac{Z_{s,g,b}}{P_{s,g,b}} \quad for\ s < 45. \tag{4}$$

For cohorts above the age of 45 years, the impact of changing retirement patterns on $p^{contrib}$ is considered. In this course, four factors which determine future retirement decisions are taken into account: 1) the increase in legal retirement ages, 2) the cut of early retirement channels, 3) the retirement histories and 4) decreasing disability prevalence rates. For a start, the probability to be a contributor constant for cohorts aged 45 and older is kept constant (see dotted line in Figure 2 in the main paper).[8] Thereafter, $p^{contrib}$ is corrected for outflows from the labour market into old age and disability retirement. More precisely, we reflect for older cohorts (s>44) the impact of changing probabilities to be already retired into old-age p^{old} or into disability p^{disab}(see **Error! Reference source not found.** and **Error! Reference source not found.**).

[6] The rise in employer contribution rates in 2013 is disregarded in this figure.

[7] This assumption is reasonable because expected changes in labour market participation of these younger age groups are not considerable. See EC (2014).

[8] As outlined in **Error! Reference source not found.** (main paper) contribution probabilities are relatively high. This observation can be explained by the fact that any sort of contribution classifies a person as a contributor. Even very small and short (one day per year) contribution periods are taken into account.

An example of the resulting changed contribution probabilities for the year 2025 is provided in **Error! Reference source not found.** (see dashed line). As shown, contributors' probabilities are expected to increase considerably up to the year of 2025 due to the increase in retirement ages to 65, the abolishment of early retirement channels and dropping disability incidence rates.

$$p_{s,g,f}^{contrib} = p_{44,g,b}^{contrib} * \varepsilon_{s,g,f} \qquad if\ s > 44\,. \tag{5}$$

$$\varepsilon_{s,g,f} \ = 1 - p_{s,g,f}^{old} - p_{s,g,f}^{disab}. \tag{6}$$

For the estimation of contributors' probabilities, the likeliness to be an old age (p^{old}) or disability beneficiary (p^{disab}) in a future year f is considered. It depends on the respective prevalence rates in the base year b (see **Error! Reference source not found.** and **Error! Reference source not found.**). Additionally, future incidence rates i to enter into old-age ($i_{j,g,f}^{old}$) and disability ($i_{s,g,f}^{disab}$) determine the probability to be an old age or disability beneficiary in a future year f. These inflow probabilities can, of course, change over time, for instance, due to reforms, such as the increase in statutory retirement ages. The estimation of $i_{j,g,f}^{old}$ and $i_{s,g,f}^{disab}$ is outlined in further detail in section **Error! Reference source not found.** in the main paper.

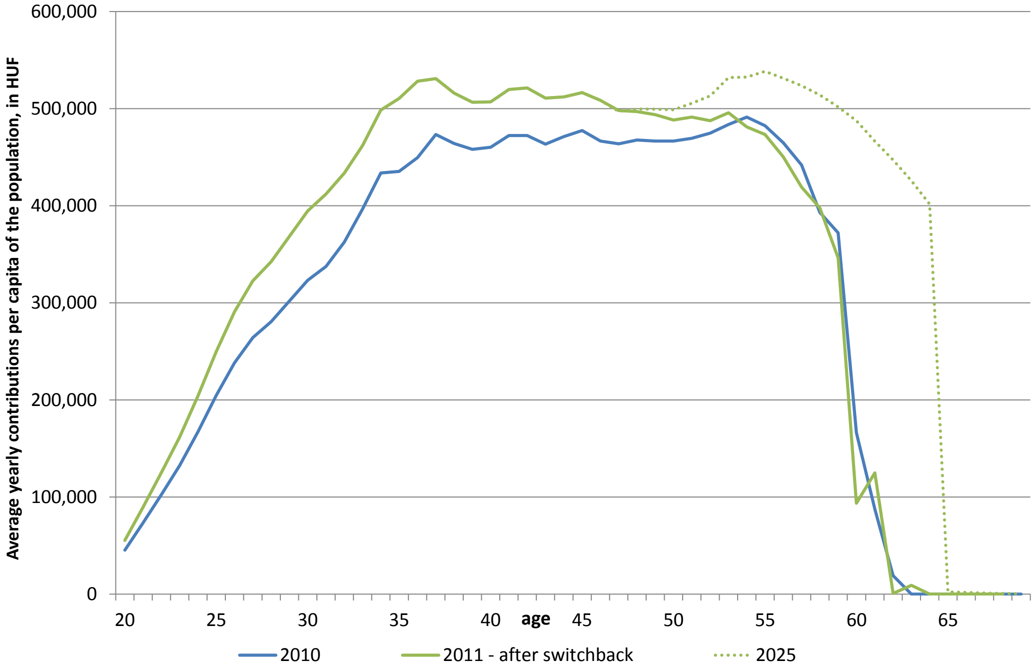
$$p_{s,g,f}^{old} = p_{s-(f-b),g,b}^{old} + \textstyle\sum_{j=45}^{x} i_{j,g,f}^{old} \qquad if\ s > 44. \tag{7}$$

$$p_{s,g,f}^{disab,c} = [p_{s-1,g,f-1}^{disab,c} * (1 - e_{s,g,f}^{disab,c}) + \ i_{s,g,f}^{disab,c}] - p_{44,g,b}^{disab,c}. \tag{8}$$

To calculate disability prevalence rates exit probabilities to leave the disability scheme due to a loss of eligibility are considered via the variable e^{disab}.[9] The application of incidence and exit probabilities for disability allows to reflect feedback loops. For instance, the impact of increasing statutory and early retirement ages on the inflow into disability can be modelled (see section **Error! Reference source not found.** in the main paper).

[9] Additionally, we consider that $p_{44,g,b}^{contrib}$ already reflects to some extent disability retirement, as inflows into disability before the age of 45 is possible, contrary to old age pensions. Therefore, to avoid double counting, we subtract the disability prevalence rate at age 44 $p_{44,g,b}^{disab}$ from the disability prevalence rates at age s>44 (in Eq. 13 in the main paper).

Figure 3: PAYG contributions per capita of population, in HUF and prices of 2010[10]



Source: own calculations.

After we have estimated contribution probabilities and their changes due to retirement behaviour, we derive average contribution ($c^{average}$) profiles, shown in **Error! Reference source not found.**. Interesting is their change over time. With the 2010 pension reform, $c^{average}$ increased from 2010 to 2011 in particular for younger cohorts. The reason for this development lies in the switch of mixed pillar members to the mono-pillar system which raised significantly the average PAYG contribution rates (see **Error! Reference source not found.**). Remarkable is also the impact of the rise in retirement ages to 65 and the cut in early retirement channels. Both measures increase the average contributions from 2011 to 2025 considerably for the age groups 50+ (see **Error! Reference source not found.**).

[10] For illustrative reasons wage growth is set at zero. Only the increase of contributions from 2010 to 2011 is considered which is largely determined by the 2011 rise in average contribution rates. The rise in employer contribution rates in 2013 is disregarded in this figure.

C. Modeling of future retirement probabilities

Appendix C is about to show estimation of future retirement probabilities in greater length than in the main text. The most important task here is to evaluate the possible impact of changes in legal retirement ages (introduction of 40 service years rule, closing the channels of early retirement and retirement age increase to 65 for both genders and the interaction of these measures).

Deriving retirement probabilities

The number of future new retirees is greatly determined by retirement behavior. To reflect the timing of future retirement inflows retirement probabilities are used. They are applied to assess the impact of pension reforms, such as the increase in retirement ages to 65. Additionally, retirement probabilities are applied to model the outflow from the labour market and the resulting changes in contributors' probabilities (discussed in Appendix A).

For the estimation of retirement probabilities, we base on recent retirement data and correct these patterns in line with legal retirement age increases. Thereby, a shift of pension careers from disability to old age is considered, too. More precisely, we take into account that a continuation of the low entrance probabilities to disability, observed in past years, will lead to a higher inflow into the old age system.

Considering changes of legal retirement ages

Base year retirement probabilities i_b^{old} are the starting point to approximate future retirement behavior i_f^{old} (see **Error! Reference source not found.**9).[11]

$$i_{s,g,f}^{old} = i_{s,g,b}^{old}. \tag{9}$$

With the adopted increase in statutory retirement ages until 2022 and the cut in early retirement channels after 2012, retirement patterns will change remarkably in the years to come. Usually, the minimum retirement age represents the lower boundary to retire. Therefore, increases of this pension age adopted with the 2009 reform translate directly into shifts of pension behaviour.[12] With the cut in early retirement possibilities legislated in 2011, no entrance can generally be expected before the

[11] They are estimated by dividing the number of new retirees $R_{s,g,b}^{new}$ in the base year with the number of scheme participants, in line with the following equation: $i_{s,g,b}^{old} = R_{s,g,b}^{new} / \left(p_{s,g,b}^{part\ pen} * P_{s,g,b}\right)$. We assume that the entire population P is participating in the general public pension system (i.e. the participation probability $p_{s,g,b}^{part\ pen}$ equals to unity). The same approach is applied for disability retirement.

[12] We assume that age groups which so far retired before or at the minimum retirement age will postpone their retirement in line with increases of early retirement ages.

(increasing) statutory retirement age. As a result, the median effective retirement age of males is expected to shift from 60 to 65 in the period 2013-2022.[13]

Early retirement is, however, still possible for a significant group of women: All females who have accrued 40 or more service years at the point of retirement are allowed to retire before statutory retirement ages.[14] To evaluate the fiscal impact of this early retirement privilege (40 years rule), it is crucial to approximate the number of women who can benefit from it. In 2013, about 30 percent of women aged 60 received a 40-service-year pension. According to our estimates, also in future years roughly every third woman can take use of the generous 40 years rule. **Error! Reference source not found.** demonstrates the age-specific probabilities to retire newly under the 40 years rule from 2013 onwards. It is estimated based on the micro-simulation model and on new retirees' data. As shown, the cumulated life-cycle probability of women to retire under the 40 service rule is expected to increase from about 30% in 2013 to 37% in 2022 (see bottom row). More women reach 40 service years after the retirement age increase becomes effective. In the period 2023-2030, the proportion declines to about 25% due to less beneficial contribution careers. After 2030, these probabilities are kept constant.

Table 1: Age specific probabilities of women to retire early under the 40 service years rule in a given year

age/year	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**	**2024**	**2025**	**2026**	**2027**	**2028**	**2029**	**2030+**
55	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	0%	0%	0%	0%
56	2%	2%	2%	2%	2%	2%	2%	2%	2%	2%	1%	1%	1%	1%	1%	1%	1%	1%
57	4%	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%	2%	2%	2%	2%	2%	2%
58	6%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	4%	4%	4%	4%	4%	3%
59	11%	8%	8%	8%	8%	8%	8%	8%	8%	8%	8%	8%	7%	7%	7%	6%	6%	6%
60	7%	8%	8%	8%	8%	8%	8%	8%	8%	8%	8%	8%	7%	7%	7%	6%	6%	6%
61	0%	3%	3%	3%	3%	3%	3%	3%	3%	3%	2%	2%	2%	2%	2%	2%	2%	2%
62	0%	0%	1%	1%	3%	3%	3%	3%	3%	3%	2%	2%	2%	2%	2%	2%	2%	2%
63	0%	0%	0%	0%	1%	1%	3%	3%	3%	3%	2%	2%	2%	2%	2%	2%	2%	2%
64	0%	0%	0%	0%	0%	0%	0%	1%	1%	3%	2%	2%	2%	2%	2%	2%	2%	2%
Sum in a given year	**30%**	**29%**	**31%**	**31%**	**33%**	**33%**	**34%**	**36%**	**36%**	**37%**	**35%**	**34%**	**32%**	**31%**	**29%**	**28%**	**26%**	**25%**

Source: own estimates.

Considering cohort specific retirement histories

The retirement history provides another rationale for differences in pension behaviour. Generally, the higher the proportion of already retired persons in one cohort the lower is the sum of future retirement probabilities for this birth year. The retirement history is reflected in the age- and gender-specific retirement rate r.[15] Retirement rates r of male Hungarians are illustrated exemplarily in **Error!**

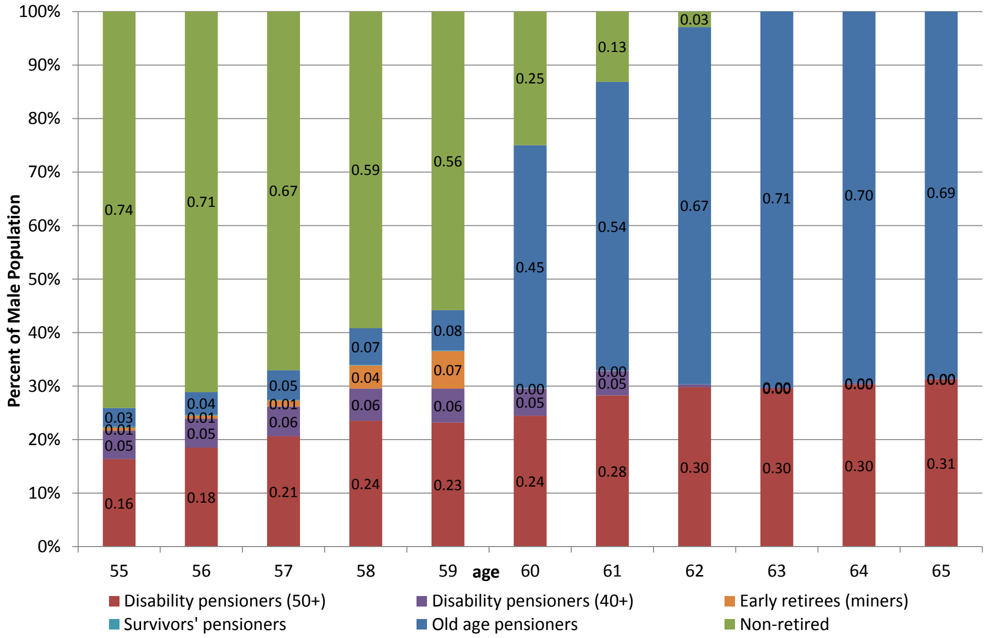
[13] For all age groups who retire above the (future) statutory retirement age, the base year retirement probabilities are kept constant in the model over time. The share of this group is marginal.

[14] They receive a full pension benefit despite early retirement, i.e. pension decrements are disregarded for this group of women.

[15] It is estimated by dividing the number of retirees R^{past} by the total number of scheme participants at age s and gender g in a year b, as outlined in the following equation: $r_{s,g,b} = R^{past}_{s,g,b} / (p^{part\ pen}_{s,g,b} * P_{s,g,b})$. Again we

Reference source not found. for the beginning of 2010. As expected, they increase with age: While at the age of 55 only about 26% of the population receives pension benefits, this ratio increases to 100% at the age of 63. Astonishing is the very high proportion of disability retirees in 2010: Almost every third male citizen received a disability benefit. The pension model takes into account that these numbers will decline over time due to lower inflows into disability (described below).

Figure 4: Male retirement and non-retirement rates (end of 2010)



Source: own estimation based on data provided by the ONYF.

In the model we consider that each cohort of age s and gender g is with a certain probability already retired in the base year or will retire in future years f at age z ($z > s$). We assume that the accumulated life-cycle retirement probabilities (LCRP) sums up to unity for each cohort (see **Error! Reference source not found.**). In other words, the sum of the retirement rate $r_{s,g,b}$, the accumulated future standard old age retirement probabilities ($\sum_{z=s+1}^{D} i_{z,g,f}^{old}$), the accumulated future net-disability retirement probabilities $\left(\sum_{i=s+1}^{D}\left(i_{z,g,f}^{disab} - l_{z,g,f}^{disab}\right)\right)$ and, in case of women, the probability to retire early with 40 service years ($\sum_{z=s+1}^{D} i_{z,g,f}^{40sy}$) amounts to one. Thus, cohorts which are already fully retired in 2010 (i.e. $r_{s,g,b} = 1$), will not retire in future years. This aspect is considered further below with the phi factor. But beforehand let us discuss the impact of disability retirement on old age retirement considered in the model.

$$LCRP_{s,g,b} = r_{s,g,b} + \sum_{z=s+1}^{D} i_{z,g,f}^{old} + \sum_{z=s+1}^{D}\left(i_{z,g,f}^{disab} - l_{z,g,f}^{disab}\right) + \sum_{z=s+1}^{D} i_{z,g,f}^{40sy} = 1. \qquad (10)$$

Considering changes in disability behavior and their impact on old age probabilities

---

assume that the entire population P is participating in the general public pension system (i.e. the participation probability $p_{s,g,b}^{part\ pen}$ equals to unity).

In line with the pension rules, disability beneficiaries stay in the disability system after the legal retirement. They do not become old age pensioners and hence their pensions are not recalculated at the legal retirement age. To reflect this feature of the Hungarian pension system, we calculate the probabilities to become a disability beneficiary and to remain in this group until the legal retirement. On this basis we are able to measure the impact of changes in disability incidence rates on the likeliness to retire into old age. We can, for instance, reflect that the relatively low disability probabilities observed in recent years will increase the overall number of old age retirees. This differentiation is valuable because the generosity between disability and old age pension benefits differs.[16] Additionally, this endogenization allows to reflect possible feedback loops which go along with recent pension reforms: In the coming decade, standard early retirement channels will be closed and legal retirement ages augment to 65. These legislative changes increase the inflows into disability, as the maximum age to be eligible to a disability pension rises from 59 to 64. Additionally, the probability to become disabled is increasing sharply in the age groups 59-64. As a consequence, less scheme participants can be expected to enter into old age after the cut in early retirement channels. The magnitude of this shift from old age to disability is outlined below.

More precisely, future disability patterns are reflected via the accumulated net-disability retirement probabilities ($\sum_{z=s+1}^{D}\left(i_{z,g,f}^{disab} - l_{z,g,f}^{disab}\right)$) which denote the annual inflow probabilities i^{disab} as well as the outflow probabilities l^{disab} to/from the disability pension scheme.[17] The stock of disability retirees in a future year f is estimated on the basis of the base year stock $r_{s,g,b}^{disab}$ and the sum of future net disability flows $\sum_{j=s+1}^{z}\left(\mathrm{i}_{z,g,f}^{disab} - \mathrm{l}_{z,g,f}^{disab}\right)$ – see **Error! Reference source not found.**. This calculation is used to derive probability to be a disability beneficiary at the legal retirement age, which determines the likeliness to retire into old age.

$$\mathrm{r}_{s+(f-b),g,f}^{disab} = \mathrm{r}_{s,g,b}^{disab} + \sum_{z=s+1}^{x+(f-b)}\left(\mathrm{i}_{z,g,f}^{disab} - \mathrm{l}_{z,g,f}^{disab}\right). \quad (11)$$

Correcting future retirement probabilities by the phi factor

As mentioned, **Error! Reference source not found.** does not necessarily have to be fulfilled. In some age groups, for instance, most cohort members are already retired in the base year. Therefore, they

[16] Individuals who enter into disability often expect a lower benefit than the standard old age pensioner (under current rules).

[17] Both flows are measured in per cent of the population. The exit flow variable $l_{s,g,f,z}^{disab}$ itself is calculated based on the stock of disability retirees in a future year f-1 at age z-1 and the probability of a disability beneficiary to leave the disability scheme $e_{s,g,f,z}^{disab}$ in year f at age z for any reason apart from death (with the following equation: $l_{s,g,f,z}^{disab} = \left(r_{s,g,f-1,z-1}\right) * e_{s,g,f,z}^{disab}$). The level of $e_{s,g,f,z}^{disab}$ is based on the age and gender specific distribution of the base year exit probabilities and hold constant over time.

cannot be expected to retire in the future. To take this into account, we correct old age retirement probabilities i^{old} with the factor φ over the remaining life-cycle (i.e. from a future age $z = s + 1$ until age $z = D$) – see **Error! Reference source not found.**.[18]

$$\varphi_{s,g,b} = \frac{1 - r_{s,g,b} - \sum_{z=s+1}^{D}\left(i_{z,g,f}^{disab} - l_{z,g,f}^{disab}\right) - \sum_{z=s+1}^{D} i_{z,g,f}^{40sy}}{\sum_{z=s+1}^{D} i_{z,g,f}^{old}}. \tag{12}$$

For the group of women, we consider in this step that a certain proportion of them can benefit from the 40 service year rules and will retire early with the exogenous retirement probabilities i^{40sy} outlined in **Error! Reference source not found.**.[19]

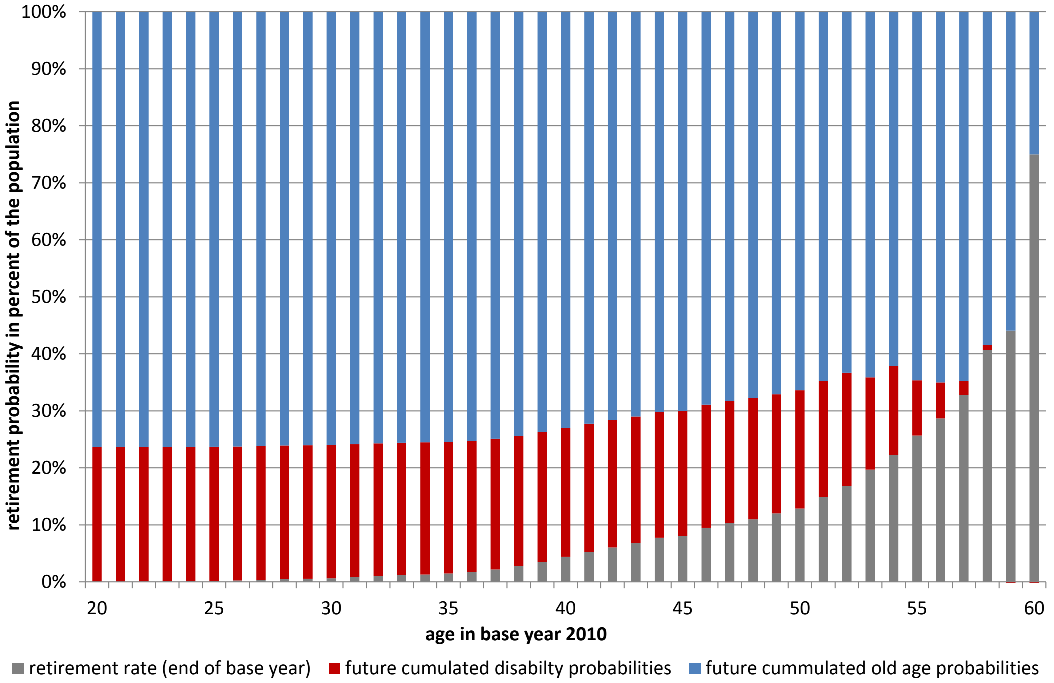
$$\mathrm{i}_{\mathrm{z,g,f}}^{\mathrm{old,corrected}} = \mathrm{i}_{\mathrm{z,g,f}}^{\mathrm{old}} * \varphi_{\mathrm{s,g,b}} \qquad for\; s < z < D. \tag{13}$$

After these steps, we know the probability of a Hungarian citizen of age s and gender g to retire into old age or disability in future years. It is illustrated exemplarily for males in **Error! Reference source not found.** considering a cut in early retirement channels from 2013 onwards. To better understand this figure, let us look at the example of a 40-year-old male Hungarian. As shown, his retirement rate amounts to about 4%, i.e. with a probability of 4% he is already retired (into old age, disability or another pension type). The probability that he will become a new disability pensioner in future years, stay in this scheme until the legal retirement age and, therefore, not enter into the old age pension scheme amounts to 23%. This figure is quite high and is to some extent driven by the cut in early retirement channels (see below). Nevertheless, it is still lower than current disability retirement rates (see **Error! Reference source not found.**). The residual of the retirement rate and the future net-disability probabilities (73%) reflects the probability that a current 40 year old male will become a new old age pensioner after the base year. Particularly interesting is the impact of the cut of early retirement channels on life cycle retirement probabilities. In a scenario without this recent reform, the probability of the 40 year old male to become a new disability pensioner in future years and not to enter into the old age pension scheme is much lower. It adds up to 17% compared to 23% in case of the early retirement cut scenario.

[18] This approach ensures that the life-cycle retirement probability (LCRP) of each birth year is equal to unity. Disability retirement probabilities of the base year are, generally, not corrected but held constant over time. Only for the early retirement cut scenario higher incidence rates are assumed for those aged above 59.

[19] Beneficiaries of the 40 service year rule are not covered in i^{old}. For males i^{40sy} always equals to zero.

Figure 5: Accumulated life-cycle retirement probabilities, males – With early retirement abolishment



Source: own estimation based on ONYF.

As a final outcome, we derive cohort and gender specific retirement probabilities. As discussed in the previous passages, they reflect possible changes in retirement behaviour due to 1) pension reforms 2) cohort specific retirement histories and 3) future changes in disability net-inflows. **Error! Reference source not found.** in the main text provides an example of these final retirement probabilities for male individuals. As can be seen, retirement probabilities gradually shift to higher ages. Furthermore, the probabilities to retire as an old age beneficiary increase. This is mainly explained by the fact that less individuals are expected to become disability pensioners.[20]

[20] Moreover, the rise in total future old age retirement probabilities is explained by the fact that in younger age groups less are already retired in the base year.

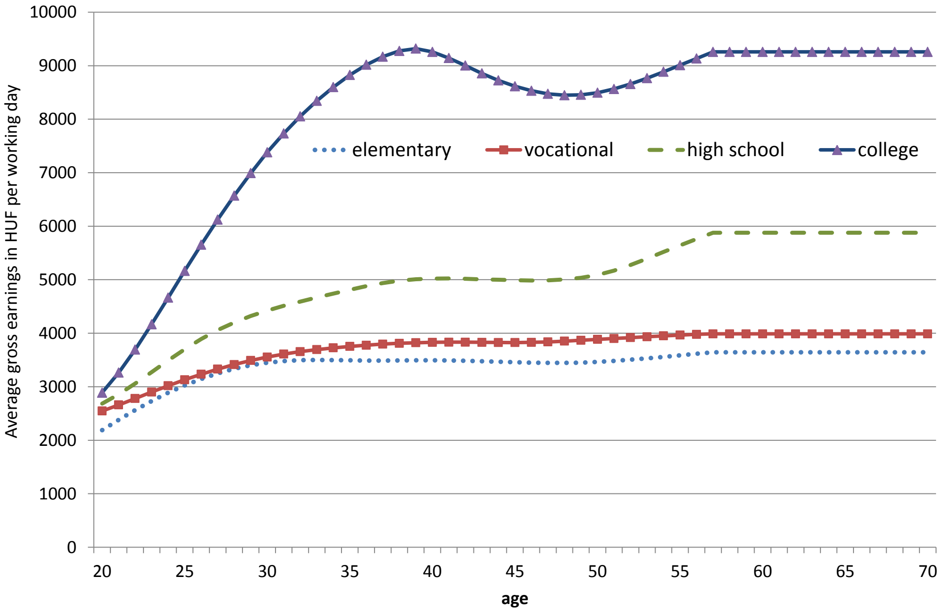
D. Extended description of the micro simulation model

In Appendix D we present some stylized facts in relation to the micro simulation model (see section 4.1 in the main paper) and provide extended description of the model with regards to estimation of gross earnings (pension contribution base) and the calculation of accrual rates.[21]

Deriving gross earnings from 1988 onwards

Estimated earnings profiles (age-specific profiles of average gross earnings per working day which are differentiated by four education groups and by gender) are corrected for the average wage growth observed in periods 1988-1996 and 2007 and after. **Error! Reference source not found.** illustrates such profiles exemplarily for males in the latter period.[22] They follow the usual age-specific patterns. Moreover, the profiles illustrate a clear difference between education groups.

Figure 6: Average (non-zero) gross earnings of males in 2006 by age and education (HUF per working day)



Source: own estimation based on CDB.

To draw conclusions about the future distribution of pension benefits, it is not only vital to reflect the age, education and gender specific dimension of earnings but also to consider the individual income position. What we know is how individuals performed in terms of earnings in the period 1997-2006 relative to their fellows in the same age, education and gender group. On this basis we estimate the average deviation from the mean age and gender specific salaries by educational attainment over the

[21] Content of Appendix D shows considerable overlap with the equivalent counterpart in the main paper.
[22] We apply these gross earnings profiles later only for the status of contributing more than zero working days. Consequently, all zero values of gross earnings in the CDB are disregarded for the calculation of these profiles.

given contribution history. Thus, an indvidual who was earning only 50 % of the average salary per working day (of his age, education and gender group) in the period 1997-2006 is assumed to remain in this relative income position also in the years for which no data has been provided.

Calculation of accrual rates (AR)

In line with the accrual schedule set in the pension law[23], the AR amounts to: 33% for the first 10 years of service time,

- \+ 2% for each of the service years between 11-25,
- \+ 1% for each of the service years between 26-36,
- \+ 1.5% for each of the service years between 37-40,
- \+ 2% for each of the service years after 40 years.[24]

The service time sums up contributory periods as well as non-contributory periods - such as national service, various types of childcare allowance, tertiary education (only prior to 1998 for old age but without limit for disability benefit), sick leave, etc.[25]

Data on service time is available for every single taxpayer who contributed at least one day in the period 1997-2006. For a small sample of these individuals, also annual service days are provided for the period 1958-1996. All other service day data needs to be estimated newly. For this calculation, the missing working career is split into two periods: 1) years before 1988 and 2) *years after 1987* (i.e. years 1988-1996 and 2007+). For the first period, full employment for age groups 18 to 70 is assumed. The missing contribution history in the years after 1987 is estimated with a two-step regression based approach.

In a first step the question is answered whether an individual is likely to work or not in the missing years after 1987. For the period 1988-1996, a backward estimation of the working status ($work_{it}$) is applied.[26] It is based on the smaller panel dataset and the linear probability model outlined in **Error! Reference source not found.**. The dependent variable $work_{it}$ reflects the probability of an individual

[23] See Section 12 of Act LXXXI of 1997 on Social Security Pension Benefits.
[24] It was envisaged to change this current non-linear accrual schedule (outlined above) with a linear accrual schedule for new retirees in 2013. More precisely it was planned to apply a constant accrual rate of 1.65% (1.22%) for each service year in case of single pillar participants (of mixed scheme participants). This legislated change of the benefit formula was, however, abolished before its introduction in 2013.
[25] A definition of service time is provided in section 37 to 43 of the Act LXXXI of 1997 on Social Security Pension Benefits.
[26] The observations of the smaller dataset are weighted to mimic the age, gender and accumulated service year profile of the large database.

i to accrue non-zero working days in a given year t.[27] A number of independent variables are employed to estimate $work_{it}$. First, it is considered that contribution activity in a given year t depends greatly on the contribution performance over the working life-cycle.[28] Three leads ($work_{it+1,} work_{it+2}, work_{it+3}$) of the dependent variable are used in the regression in order to capture status such dependency over time. The time lead of up to three periods is valuable as strong hysteresis effects can be observed on the Hungarian labour market. Furthermore, an educational attainment categorical variable ($educ_i$)[29], continuous age variables (age_{it}) with a non-linear age specification as well as the number of service years accumulated between 1997 and 2006 ($\mathrm{SUMWD9706_i}$) are included in the regression analysis. Separate regressions are used for both genders.

$$work_{it} = \propto + \sum_{j=1}^{3} \beta_j work_{i,t+j} + \beta_4 year_i + \sum_{j=5}^{8} \beta_j (educ_i = j-4) + \beta_9 SUMWD9706_i + \beta_{10} SUMWD9706_i^2 + \sum_{j=11}^{14} \beta_j age_{it}^{j-10} + \mu_{it}. \quad (14)$$

For the estimation of employment statuses after 2006, basically, the same technique is applied (see **Error! Reference source not found.**). In this forward-looking estimation, the accumulated service years variable is, however, left out. The continuous age specification is replaced with a discrete age variable which interacts with the categorical variable education.[30] Besides, leads of the working status are exchanged by respective lagged variables.

$$work_{it} = \propto + \sum_{j=1}^{3} \beta_j work_{i,t-j} + \beta_4 year_i + \sum_{e=1}^{4} \sum_{a=18}^{70} \gamma_{ea} (educ_i = e) \times (age_i = a) + \mu_{it}. \quad (15)$$

As stated, we observe strong hysteresis effects on the Hungarian labour markets in the data. In other words, the chance of an inactive individual to re-enter into the labour market is a decreasing function of the duration in unemployment/inactivity. As outlined exemplarily in Figure , male individuals who have been inactive for the last consecutive three years before year $t+1$ have a very high probability of around 70 % to remain inactive also in year $t+1$. If they look back on only one year of inactiveness in year t, this probability is significantly lower adding up to about 50 %. Individuals who have been active in period t, on the contrary, show a very small probability, generally, well below 5% to switch into the status of inactiveness in year $t+1$. These effects are captured in the regression model. Figure

[27] One may criticize the choice of only two statuses (zero and non-zero working days). The distribution of working days observed in the period 1997-2006, however, confirms this approach. In fact, the majority of contributors feature either zero contribution days (about 25 per cent) or a level of contribution days close to 365 (366) days per year (about 60 per cent).

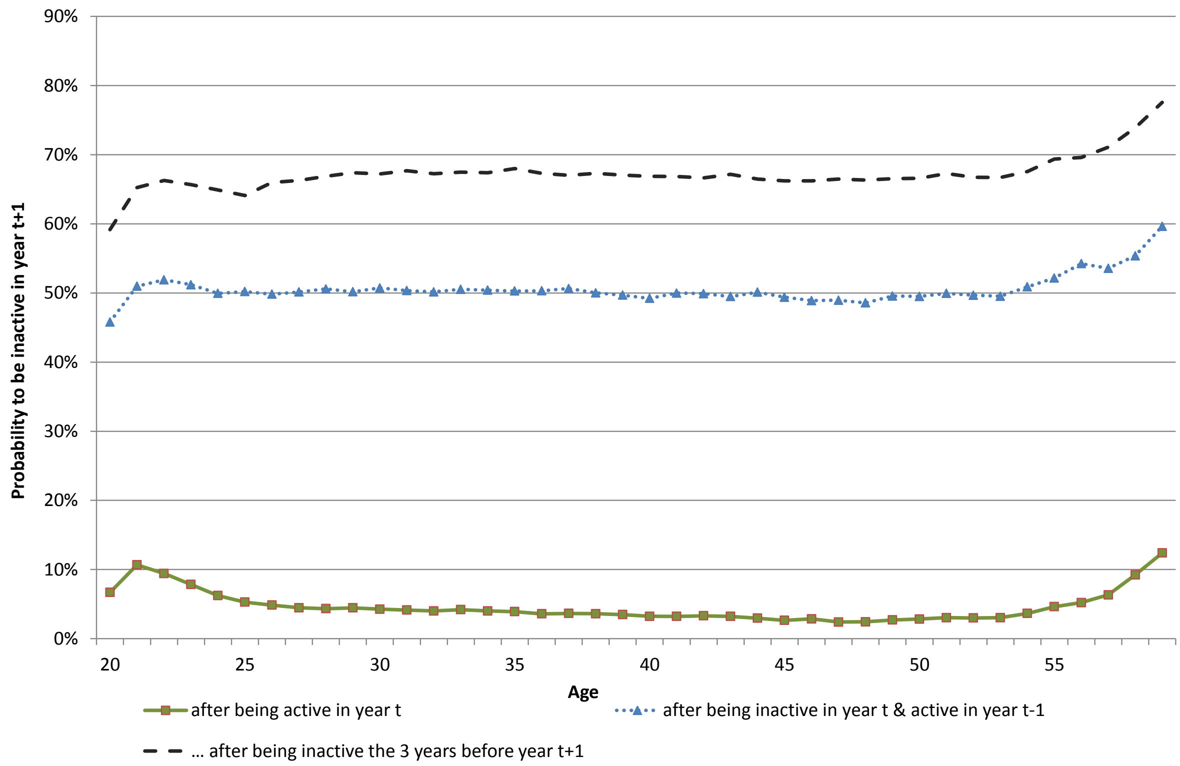
[28] Since the Breusch-Pagan Lagrange multiplier test rejects random effects (variances across entities is zero, i.e. no panel effect), a simple OLS regressions is used.

[29] Four educational groups are considered: individual with 1) elementary, 2) vocational, 3) high school and 4) college degree.

[30] The reason why we use a continuous age specification in the backward simulation is that we have very few observations for individuals above 50 between 1988 and 1996.

also outlines that the probability to switch into the status of inactivity is relatively stable over the working life-cycle and only increases at older ages.[31]

Figure 7: Probability to become inactive after different contribution histories – example of males



Source: own estimation based on CDB.

Based on the regression analysis, we can estimate the probability ($work_{it}$) for each individual i that he or she accrues at least 1 working day in a given year t. The final employment status in the period 1988-1996 and after 2006 is then estimated by a stochastic process. For this purpose, a pseudorandom number $rand_{it}$ is generated from a uniform distribution [0,1] for each of these years and for each individual. If $rand_{it}$ exceeds the value of $work_{it}$, then the respective individual in a given year is inactive. The backward-looking simulation starts in the year 1996. Employment statuses are estimated from the observed leads of 1997-1999 and are then going backward to the year of 1988. Year effects are responsible for capturing general labor market shocks. The same method is applied for the estimation of the individual work status after 2006. Since no data on year effects ahead is available, the year effect is fixed at its estimated value of 2006.

After these steps, each individual is attributed with a working status (active/inactive) in the years 1988-1996 and the years after 2006. In a next step, the amount of working days (wd_{it}) accrued in a given year is estimated. For inactive individuals this procedure is simple. They accrue zero working days in a

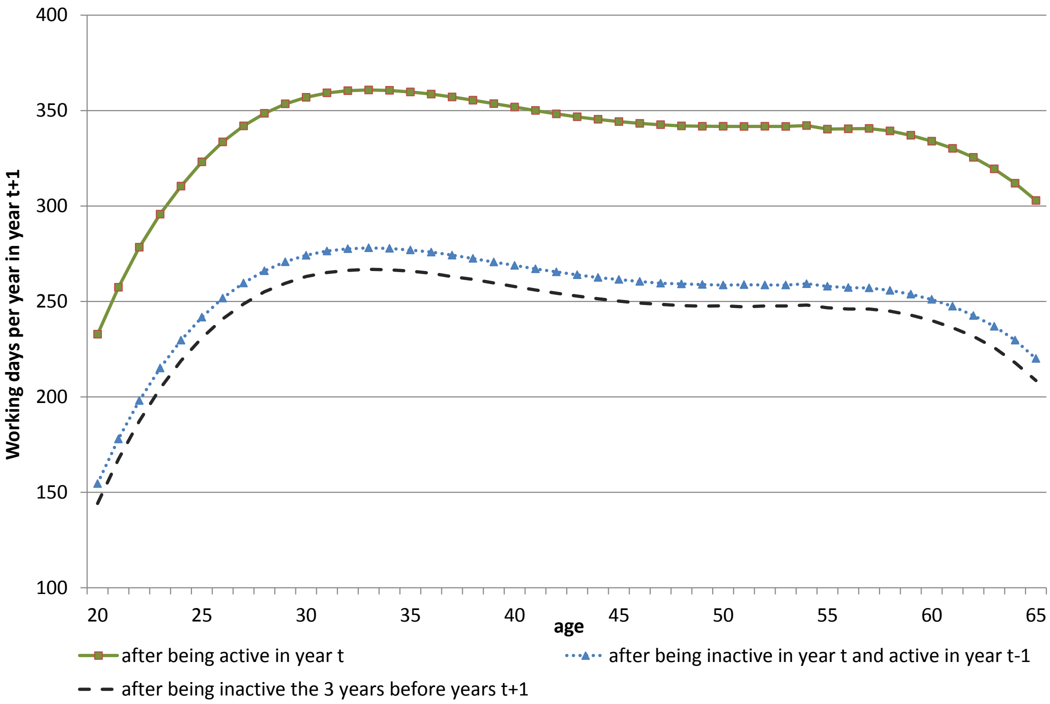
[31] The transition probabilities presented in Figure are held constant from age 57 onwards as they reflect mainly retirement exits at these ages which shall be treated endogenously in the model.

given year. For active individuals, on the contrary, the $\mathrm{wd_{it}}$ variable is approximated on the basis of a fixed effects model (see **Error! Reference source not found.**).[32] This model is applied separately for four education groups and two gender classes, i.e. eight separate regressions are used. It considers the age (age_{it})[33] of the individual, the working status reached in previous years $(work_{it-1}, work_{it-2}, work_{it-3})$ as well as individual fixed effects (μ_i). Moreover, timing effects are reflected in the fixed effects regression by the use of year dummies $(year_i)$.

$$wd_{it} = \alpha + \sum_{j=1}^{3} \beta_j wd_{i,t-j} + \beta_4 year_t + \sum_{j=5}^{8} \beta_j age_{it}^{j-4} + \mu_i + \varepsilon_{it}. \quad (16)$$

On this basis, we can derive gender and education specific age profiles of working days accrued by active scheme participants. As shown in Figure , working days increase steeply until the age of 30. For the age groups of 30-60 year olds, working days remain at a relatively high level. Interesting is the decline at the age of 60. This drop may be explained by a higher proportion of (still active) scheme participants working part-time.[34] The value of working days depends also significantly on the working status observed in the previous year. In fact, working days are considerably lower after periods of inactivity (see lower lines in Figure). The value of working days is, however, not so much affected by the duration of the working inactivity. As shown, working days do not significantly differ after a contribution pause of 1 or 3 years.

Figure 8: Working days profile after different contribution histories – example of males



---

[32] The Hausman test confirms the application of a fixed effects model.

[33] A non-linear relationship between the dependent variable of working days (wd) and the explanatory variable of age is assumed. More precisely, we apply a polynomial approximation of fourth order.

[34] Please note that we shift the working days profile to the right in line with legislated increases in retirement ages.

Source: own estimation based on CDB.

After the regression analysis, the total of service time (incl. contributory and non-contributory service time) at the point of retirement is estimated for each individual. For the calculation of average yearly earnings (AYI), only the sum of contributory service years (reflected in the variable DCI) is required. To derive this sum we, finally, subtract non-contributory service years (based on data provided by the ONYF) from the total service year variable.

To draw conclusions about the future distribution of pension benefits, it is not only vital to reflect the age, education and gender specific dimension of earnings but also to consider the individual income position. What we know is how individuals performed in terms of earnings in the period 1997-2006 relative to their fellows in the same age, education and gender group. On this basis we estimate the average deviation from the mean age and gender specific salaries by educational attainment over the given contribution history. Thus, an individual who was earning only 50% of the average salary per working day (of his age, education and gender group) in the period 1997-2006 is assumed to remain in this relative income position also in the years for which no data has been provided.

References

European Commission (2014): “The 2015 Ageing Report: Underlying Assumptions and Projection Methodologies”, European Economy, Report prepared by the European Commission.

Hirose, Kenichi (2011): “National Reports – Hungary”, in: Kenichi Hirose (ed.): “*Pension Reform in Central and Eastern Europe in times of crisis, austerity and beyond”*, International Labor Organization, pp. 171-198.

Murthi, M., Orszag, J. M., and Orszag, P. R. (2001): “Administrative costs under a decentralized approach to individual accounts: lessons from the United Kingdom“, in: Robert Holzmann and Joseph Stiglitz (eds), New Ideas about Old Age Security. Washington, DC: The World Bank.