Online Appendix from Bönke, Corneo and Lüthen "Lifetime Earnings Inequality in Germany"

Contents

Appendix I: Data

I.1 Earnings concept

The earnings information provided in the Insurance Account Sample is based on the employee's gross wage. In order to obtain the market value of earnings w_G, the social security contributions paid directly by the employer have to be added to the gross wage w according to equation (I.1):

$$w_G = w + \min(z_p, w)\, r_p + \min(z_h, w)\, (r_h + r_l) + \min(z_u, w)\, r_u \qquad (I.1)$$

In (I.1), z_s denotes the contribution ceiling and r_s the employer's contribution rate in the various branches $s = p(ension), h(ealth), u(nemployment), l(ongterm\ care)$ of the German social security scheme. The respective contribution ceiling and rate are provided in Table I.1.

Some categories of employees like miners, sailors and distinctive employees of the federal railways have special social security arrangements (*knappschaftlich Versicherte*). For these employees the contribution ceiling and contribution rate of the pension insurance, z_p and r_p, differ. Both are higher than in the regular scheme (see Table I.2) and mirror the historically higher risks of employees in these industries. Due to the higher health risks, this pension scheme includes an additional pension component used for earlier retirement entries, additional invalidity care and higher pensions (§§ 40, 45, 85, 238, 239 and 242; Social Code VI [*Sozialgesetzbuch VI*]). This pension scheme is especially relevant for male employees of older cohorts. For example an average of 10% to 15% born between 1935 and 1940 are subject to these special social security arrangements where it is negligible for the younger cohorts.

In order to provide the most accurate picture of the comparison between younger and older employees, we want to include employees with have special social security arrangements (*knappschaftlich Versicherte*), as their share is non-negligible in older cohorts. The differences in the contributions ceiling could be accounted for with by imputations methods. However, the differences in the contribution rates range from 6% to 10%, depending on the year (see Table I.1 below). This could potentially bias the analysis of earnings inequality and to account for that matter, we turn to the concept of market value of labor. As our robustness-section III shows, our results are nonetheless robust regardless the treatment of employer's social security contributions.

Table I.1: Key parameters of German social security (default case for regular insured)

Year	Average earnings[A]	Pension insurance ceiling	Pension insurance rate	Health insurance ceiling	Health insurance rate[B]	Unemployment insurance ceiling	Unemployment insurance rate	Long-term care insurance rate[C]
1952	3,852	7,800	5.000	6,000	3.000	6,000	2.000	
1953	4,061	9,000	5.000	6,000	3.000	6,000	2.000	
1954	4,234	9,000	5.000	6,000	3.100	6,000	2.000	
1955	4,548	9,000	5.375	6,000	3.100	6,000	1.630	
1956	4,844	9,000	5.500	6,000	3.100	6,000	1.500	
1957	5,043	9,000	6.750	6,480	3.900	9,000	1.083	
1958	5,330	9,000	7.000	7,920	4.200	9,150	1.000	
1959	5,602	9,600	7.000	7,920	4.200	9,150	1.000	
1960	6,101	10,200	7.000	7,920	4.200	9,150	1.000	
1961	6,723	10,800	7.000	7,920	4.700	9,150	1.000	
1962	7,328	11,400	7.000	7,920	4.800	9,150	0.775	
1963	7,775	12,000	7.000	7,920	4.800	9,150	0.700	
1964	8,467	13,200	7.000	7,920	4.850	9,150	0.650	
1965	9,229	14,400	7.000	8,880	4.950	9,150	0.650	
1966	9,893	15,600	7.000	10,800	5.000	9,150	0.650	
1967	10,219	16,800	7.000	10,800	5.050	10,650	0.650	
1968	10,842	19,200	7.500	10,800	5.100	15,600	0.650	
1969	11,839	20,400	8.000	11,250	5.250	18,000	0.650	
1970	13,343	21,600	8.500	14,400	4.100	21,600	0.650	
1971	14,931	22,800	8.500	17,100	4.100	22,800	0.650	
1972	16,335	25,200	8.500	18,900	4.200	25,200	0.850	
1973	18,295	27,600	9.000	20,700	4.600	27,600	0.850	
1974	20,381	30,000	9.000	22,500	4.700	30,000	0.850	
1975	21,808	33,600	9.000	25,200	5.200	33,600	1.000	
1976	23,335	37,200	9.000	27,900	5.600	37,200	1.500	
1977	24,945	40,800	9.000	30,600	5.700	40,800	1.500	
1978	26,242	44,400	9.000	33,300	5.700	44,400	1.500	
1979	27,685	48,000	9.000	36,000	5.600	48,000	1.500	
1980	29,485	50,400	9.000	37,800	5.700	50,400	1.500	
1981	30,900	52,800	9.250	39,600	5.900	52,800	1.500	
1982	32,198	56,400	9.000	42,300	6.000	56,400	2.000	
1983	33,293	60,000	9.083	45,000	5.900	60,000	2.300	
1983	34,292	62,400	9.250	46,800	5.700	62,400	2.300	
1985	35,286	64,800	9.454	48,600	5.900	64,800	2.150	
1986	36,627	67,200	9.600	50,400	6.100	67,200	2.000	
1987	37,726	68,400	9.350	51,300	6.300	68,400	2.150	
1988	38,896	72,000	9.350	54,000	6.500	72,000	2.150	
1989	40,063	73,200	9.350	54,900	6.500	73,200	2.150	
1990	41,946	75,600	9.350	56,700	6.300	75,600	2.150	
1991	44,421	78,000	8.980	58,500	6.100	78,000	3.090	
1992	46,820	81,600	8.850	61,200	6.400	81,600	3.150	
1993	48,178	86,400	8.750	64,800	6.700	86,400	3.250	
1994	49,142	91,200	9.600	68,400	6.600	91,200	3.250	
1995	50,665	93,600	9.300	70,200	6.600	93,600	3.250	0.500
1996	51,678	96,000	9.600	72,000	6.700	96,000	3.250	0.850
1997	52,143	98,400	10.150	73,800	6.800	98,400	3.250	0.850
1998	52,925	100,800	10.150	75,600	6.800	100,800	3.250	0.850
1999	53,507	102,000	9.850	76,500	6.800	102,000	3.250	0.850
2000	54,256	103,200	9.650	77,400	6.800	103,200	3.250	0.850
2001	55,216	104,400	9.550	78,300	6.800	104,400	3.250	0.850
2002	28,626	54,000	9.550	40,500	7.000	54,000	3.250	0.850
2003	28,938	61,200	9.750	41,400	7.200	61,200	3.250	0.850
2004	29,060	61,800	9.750	41,856	7.200	61,800	3.250	0.850
2005	29,202	62,400	9.750	42,300	7.100	62,400	3.250	0.850
2006	29,494	63,000	9.750	42,756	6.500	63,000	3.250	0.850
2007	29,951	63,000	9.950	42,756	6.800	63,000	2.100	0.850
2008	30,625	63,600	9.950	43,200	6.900	63,600	1.650	0.850
2009	30,879	64,800	9.950	44,100	7.000	64,800	1.400	0.975

Note: Average earnings and contribution ceilings denoted in current prices and currency (1952 - 2001 in DM, 2002 - 2009 in Euro), reported rates are employer's contribution rates. [A]Subject to social security contributions. [B]Average contribution rate. Employees with high earnings who are eligible to opt for private health insurance (*Versicherungsfreigrenze*) are considered to remain in the public health insurance. [C] The contribution ceilings of the long-term care and the health insurance coincide.
Source: Appendices 1 and 2 of Social Code VI (*Sozialgesetzbuch VI*), Federal Ministry of Labour and Social Affairs.

Table I.2: Regulations for the old age pension schemes of miners

Year	Average earnings[A]	Pension insurance ceiling	rate
1952	3,893	12,000	15.500
1953	4,104	12,000	15.500
1954	4,279	12,000	15.500
1955	4,596	12,000	15.500
1956	4,895	12,000	15.500
1957	5,096	12,000	15.200
1958	5,386	12,000	15.000
1959	5,661	12,000	15.000
1960	6,165	12,000	15.000
1961	6,794	13,200	15.000
1962	7,405	13,200	15.000
1963	7,857	14,400	15.000
1964	8,556	16,800	15.000
1965	9,326	18,000	15.000
1966	9,997	19,200	15.000
1967	10,327	20,400	15.000
1968	10,957	22,800	15.000
1969	11,965	24,000	15.000
1970	13,485	25,200	15.000
1971	15,090	27,600	15.000
1972	16,508	30,000	15.000
1973	18,489	33,600	15.000
1974	20,597	37,200	15.000
1975	22,039	40,800	15.000
1976	23,582	45,600	15.000
1977	25,209	50,400	15.000
1978	26,520	55,200	15.000
1979	27,979	57,600	15.000
1980	29,798	61,200	15.000
1981	31,228	64,800	15.000
1982	32,540	69,600	14.750
1983	33,646	73,200	15.170
1983	34,655	76,800	16.000
1985	35,660	80,400	15.300
1986	37,015	82,800	15.350
1987	38,125	85,200	15.100
1988	39,307	87,600	15.100
1989	40,486	90,000	15.100
1990	41,946	93,600	15.100
1991	44,421	96,000	14.645
1992	46,820	100,800	14.600
1993	48,178	106,800	14.500
1994	49,142	112,800	15.900
1995	50,665	115,200	15.400
1996	51,678	117,600	15.900
1997	52,143	121,200	16.750
1998	52,925	123,600	16.750
1999	53,507	124,800	16.380
2000	54,256	127,200	15.950
2001	55,216	128,400	15.850
2002	28,626	66,600	15.850
2003	28,938	75,000	16.150
2004	29,060	76,200	16.150
2005	29,202	76,800	16.150
2006	29,494	77,400	16.150
2007	29,951	77,400	16.450
2008	30,625	78,600	16.450
2009	30,506	79,800	16.450

Note: Average earnings and contribution ceilings denoted in current prices and currency (1952 - 2001 in DM, 2002 - 2009 in Euro), reported rates are employer's contribution rates. [A]Miners subject to social security contributions.
Source: Appendices 1 and 2 of Social Code VI (*Sozialgesetzbuch VI*)

I.2 Imputation of top-coded earnings

The imputation of incomes for top-coded observations assumes that top incomes are distributed according to the Pareto law. Several studies investigating income distributions in various countries indicate that this is a good assumption.

Assume that individual earnings w_i exceeding $\tilde{w}$ are Pareto-distributed. Then, the probability to observe an income greater or equal to $w_i > \tilde{w}$ is given by

$$1 - F(w_i) = \left(\frac{w_i}{\tilde{w}}\right)^{-\alpha} \tag{I.2}$$

where $F(w_i)$ denotes the cumulative probability density function. Consider n to be the number of earners with $w_i > \tilde{w}$ and $i = 1, \dots, n$. Furthermore, earners i are ranked in ascending order according to their income. From equation (I.2) each individual's rank r_i in the income distribution is determined as

$$r_i = nF(w_i) = n\left(1 - \left(\frac{w_i}{\tilde{w}}\right)^{-\alpha}\right) \tag{I.3}$$

In top-coded data, individual earnings are available up to a contribution ceiling, z. If an individual earns more, reported earning is $w_i = z$. Consider m out of the n earners to receive an income above the contribution ceiling $z > \tilde{w}$. Since for m earners neither r_i nor w_i is observable, we estimate the parameters of the Pareto-distribution by exploiting earnings data from the interval $[\tilde{w}, z]$. Rearranging equation (I.3) yields

$$\ln\left(1 - \frac{r_i}{n}\right) = -\alpha \ln\left(\frac{w_i}{\tilde{w}}\right) \tag{I.4}$$

We employ equation (I.4) to estimate the Pareto-coefficient α. Suppose at least the top 10% of individual earnings w_i in the interval $[0, z)$ to be Pareto-distributed. Accordingly, $\tilde{w}$ is assigned the value of the 90th percentile in the respective distribution of earnings below z. The Pareto-coefficient is estimated by means of an OLS regression without constant. The regression is conducted separately for all years t and birth cohorts c. Hence, the cohort and year specific Pareto-coefficient $\hat{\alpha}_{c,t}$ is derived for $c = 1935, \dots, 1969$ and $t = 1952, \dots, 2009$ distributions. With the estimated Pareto-coefficient at hand, unobserved earnings above the contribution ceiling z can be estimated by rearranging (I.3):

$$\hat{w}_i = \tilde{w}\left(1 - \frac{\hat{r}_i}{n}\right)^{-\frac{1}{\hat{\alpha}}} \tag{I.5}$$

where $\widehat{w}_i$ denotes the estimated earned income and $\hat{r}_i$ the assumed rank. The conjectures regarding $\hat{r}_i$ have an immediate effect on measures of income mobility and, therefore, are crucial when investigating earnings dynamics. In our preferred imputation, we choose $\hat{r}_i$ under the minimal mobility assumption. Thereby, the rank $\hat{r}_i$ is based on the last observable rank in relation to all individuals at or above the contribution ceiling in the cohort-specific earnings distribution. [1] This imputation procedure leads to plausible annual earnings distributions. Comparing the obtained annual earnings distributions to (almost) uncapped survey-based micro data reveals a good fit (see Figure A1).

---

[1] For illustration consider two earnings distributions in subsequent periods $t-1$ and t made out of three individuals a, b and c. Suppose the following ordering of earnings in $t-1$: $w_{a,t-1} < w_{b,t-1} < z_{t-1} < w_{c,t-1}$ and resulting ranks $r_{a,t-1} = 1, r_{b,t-1} = 2$ and the estimated rank $\hat{r}_{c,t-1} = 3$ since c's earnings exceed z_{t-1}. In t individual a has earnings above the contribution ceiling such that $w_{b,t} < z_t$ and $w_{a,t}, w_{c,t} > z_t$ where it is not observable whether a or c earns more. Then, the ranking order in t is $r_{b,t} = 1, \hat{r}_{a,t} = 2$ and $\hat{r}_{c,t} = 3$ because of $r_{c,t-1} > r_{a,t-1}$. Thus, the relative ordering of a and c remains unchanged for future years unless either a's or c's earnings fall below the contribution ceiling. To establish whether mobility results are robust, two alternative mobility scenarios are calculated: an equal ranking with imputation of estimated average earnings above the contribution ceiling and a maximum mobility scenario. In the maximum mobility scenario, the ranking order is reversed between years t and $t+1$. All alternative results and a scenario without imputation are provided in Online Appendix III.3.

Table I.3: Cohort-specific means of shares of censored spells

Cohort	Men	Women
1935	8.00	0.37
1936	7.47	0.37
1937	7.95	0.30
1938	7.98	0.38
1939	9.35	0.47
1940	9.09	0.62
1941	9.31	0.42
1942	10.63	0.54
1943	10.45	0.79
1944	9.68	0.65
1945	9.03	0.60
1946	7.84	0.48
1947	7.98	0.75
1948	7.58	0.55
1949	7.88	0.65
1950	7.64	0.45
1951	7.94	0.48
1952	7.53	0.64
1953	6.97	0.49
1954	7.59	0.63
1955	6.42	0.31
1956	5.96	0.56
1957	6.60	0.69
1958	6.47	0.41
1959	6.41	0.58
1960	6.14	0.39
1961	5.47	0.60
1962	5.43	0.65
1963	5.11	0.58
1964	5.37	0.46
1965	4.88	0.88
1966	4.37	0.49
1967	3.99	0.60
1968	3.77	0.65
1969	3.24	0.49
Mean	7.07	0.54

Note: Means are calculated as share of censored spells on all annual spells in all years for each cohort. Differences in shares of censored spells across cohorts are due to changes in the contribution ceiling (see Table I.1).
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke.

I.3 Correction of structural break 1983/1984

Starting with 1984, one-time payments (e.g. holiday and Christmas allowances or bonuses) are subject to social security contributions and included in the basis of assessment and hence the earnings measure. This leads potentially to an artificial increase in inequality in the annual earnings distributions after 1983. Facing the same problem for comparable but cross sectional data, Fitzenberger (1999) suggests fixing this structural break by estimating quantile specific deviations from the median growth rate between 1983 and 1984.[2] A similar strategy is adopted by Dustman et al. (2009) and Card et al. (2013). In order to meet our data requirements we adjust Fitzenberger's (1999) strategy to panel data.

The imputation of one-time payments for observations before 1984 is accomplished as follows. First, we generate a variable containing the average individual rank in the cohort

---

[2] Fitzenbergers (1999) study is based on the *IAB Beschäftigtenstichprobe*, also obtained from social security administration data.

specific earnings distribution between age 35 and 40. This variable serves as an approximation for the individual's permanent position in the earnings distribution and reflects the finding by Fitzenberger (1999) that spurious growths due to one-time payments is more pronounced for higher earnings. The earnings position is coded as dummy d_q for $q = 1, \dots, 20$ quantiles and $d_q = 1$ if the individuals average rank in the annual earnings distributions between 35 and 40 falls into the respective quantile. Furthermore, we define earnings growth between t and $t + 1$ as $\Delta w_{i,t} = ln(w_{i,t}) - ln(w_{i,t-1})$. Earnings growth is estimated with a generalized least squares random effects regression in an unbalanced panel restricted to prime age individuals from 26 to 59 according to equation I.6:

$$
\begin{aligned}
\Delta w_t = \alpha_0 + \; &\alpha_1 d_{1984} + \alpha_2 age_t + \alpha_3 age_t^2 + \alpha_4 age_t^3 \\
&+ \alpha_5 d_{1984} age_t + \alpha_5 d_{1984} age_t^2 + \alpha_5 d_{1984} age_t^3 + \boldsymbol{\beta} \mathbf{d}_q' \\
&+ \boldsymbol{\gamma} d_{1984} \mathbf{d}_q' + \boldsymbol{\delta} age_t \mathbf{d}_q' + \varepsilon
\end{aligned} \qquad (\text{I.6})
$$

To identify spurious growth between 1983 and 1984, we include a dummy variable marking the structural break with $d_{1984} = 1$ in 1984 and zero else. Furthermore, we model age-earnings profiles by including age as a third order polynomial function and the vector of average earnings rank dummies $\mathbf{d}_q$ as well as interactions for all rank dummies with age respectively the structural break dummy d_{1984}.

Regression results confirm the cross sectional pattern reported in Fitzenberger (1999) with higher spurious growth rates for above median annual earnings. Figure I.1 displays the spurious growths pattern for selected quantiles. Depending on their positions in the cohort specific permanent earnings distribution, individual earnings are corrected by the quantile and age specific excessive growth factor for years predating 1984. Due to top coding, we assume the 17th quantile's for earnings in the 18th, 19th and 20th quantile (see Fitzenberger 1999).

Figure I.1 Spurious growth rates for selected quantiles, men
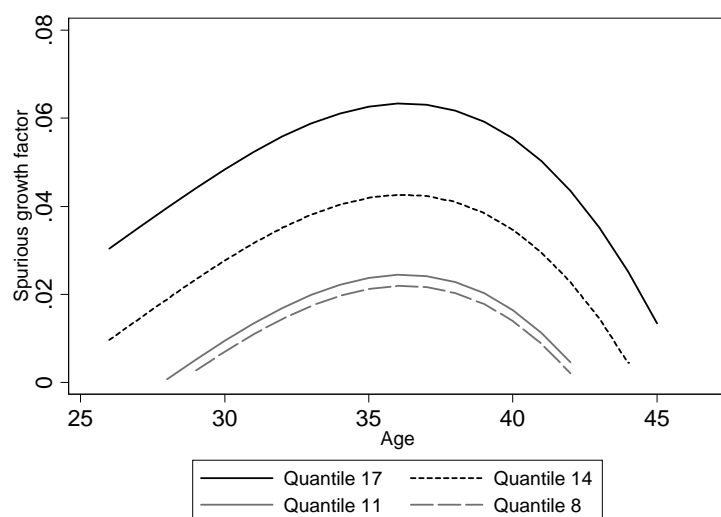
Note: The estimation is based on prime age males. Displayed is the relevant age range only. Spurious growth is identified as excessive growth between 1983 and 1984.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke.

I.4 Sample selection: General

To ensure that our lifetime earnings capture all relevant labor market activities, we restrict our sample to individuals with "valid biographies". To construct valid biographies, we measure the time where we cannot eliminate the possibility that an individual earned income besides what is recorded in our data. In order to achieve this, we exploit the fact that our data provides information on times apart from regular employment. Depending on the information on labor market activities we distinguish three cases:

(1) The individual is a regular employee subject to social security contributions (times of regular employment). Here we observe all relevant labor market activities and do not alter the earnings recorded in our data.

(2) The individual is not an employee subject to social security contributions but accumulates times relevant for old age pension. In this sense, the information provided in our data excludes the possibility of income earned on the regular labor market (e.g. times of educational training, care, sickness, unemployment, community or military service, disability or retirement). In this case we treat the respective monthly earnings as zeros.

(3) The individual is not an employee subject to social security contributions and we cannot exclude the possibility of income earned apart from what is recorded in our data. This is where times of self-employment, working as civil servant or labor market withdrawal cannot be distinguished and we do not have the necessary information to conclude that we observe all relevant labor market activities. Therefore we recode these monthly earnings into times of missing information.

In order to select individuals with complete occupational biographies, we now exclude all individuals who display more than one month of missing information per year after the age of

30. Hence, for an individual to be included in the analysis, the up-to-age 40 concept allows for up to 10 month of missing information before an individual is excluded, the up-to-age 41 concepts for up to 11 month and so forth.

I.5 Sample selection: Men

Table I.4: Number of observed men

		Sample size after exclusion of:					
Cohort	Original number of observations	Foreigners	East Germans	Repatriates (*Aussiedler*)	Self-employed craftsmen	Rejoiners (marriage-law)[A]	Completed biographies (up to age 40)
1935	3,236	2,195	1,746	1,661	1,606	1,606	1,114
1936	3,214	2,210	1,746	1,671	1,623	1,623	1,067
1937	3,201	2,253	1,751	1,680	1,626	1,626	1,081
1938	3,198	2,269	1,768	1,683	1,653	1,653	1,104
1939	3,228	2,295	1,791	1,733	1,678	1,678	1,207
1940	3,230	2,259	1,735	1,666	1,623	1,623	1,095
1941	3,328	2,330	1,802	1,716	1,677	1,677	1,121
1942	3,289	2,335	1,799	1,723	1,686	1,686	1,109
1943	3,320	2,382	1,772	1,697	1,653	1,653	1,107
1944	3,324	2,376	1,724	1,673	1,623	1,623	1,087
1945	3,334	2,406	1,818	1,761	1,725	1,725	1,154
1946	3,280	2,311	1,852	1,803	1,767	1,767	1,172
1947	3,380	2,427	1,837	1,771	1,741	1,741	1,175
1948	3,472	2,461	1,913	1,840	1,806	1,806	1,189
1949	3,514	2,517	1,903	1,816	1,775	1,775	1,163
1950	3,706	2,624	1,934	1,839	1,802	1,802	1,202
1951	3,988	2,787	2,034	1,913	1,869	1,869	1,228
1952	4,087	2,806	1,976	1,850	1,813	1,813	1,212
1953	4,122	2,832	2,100	1,958	1,927	1,927	1,223
1954	4,215	2,789	2,103	1,979	1,949	1,949	1,271
1955	4,497	2,911	2,135	2,029	2,003	2,003	1,293
1956	4,505	2,877	2,178	2,056	2,029	2,029	1,311
1957	4,806	2,914	2,210	2,094	2,052	2,052	1,295
1958	5,130	2,948	2,250	2,133	2,094	2,094	1,322
1959	5,510	3,027	2,305	2,194	2,156	2,156	1,345
1960	6,174	3,117	2,339	2,242	2,211	2,211	1,377
1961	7,013	3,259	2,465	2,360	2,318	2,318	1,417
1962	7,338	3,380	2,527	2,423	2,394	2,394	1,481
1963	7,488	3,436	2,590	2,506	2,464	2,464	1,494
1964	7,595	3,325	2,522	2,463	2,428	2,428	1,437
1965	7,646	3,305	2,555	2,496	2,460	2,460	1,493
1966	7,750	3,347	2,606	2,558	2,516	2,516	1,507
1967	7,699	3,307	2,567	2,519	2,476	2,476	1,511
1968	7,830	3,254	2,528	2,488	2,453	2,453	1,531
1969	8,044	3,355	2,641	2,608	2,577	2,577	1,622
	168,691	96,626	73,522	70,602	69,253	69,253	44,517

Note: Second column: Original sample size. Columns three to eight: All numbers denote observation after the stepwise exclusion of the respective groups. [A]Insured who left and rejoined the statutory pension system due to the law of marriage refunds (valid until 1967), which was possible until 1995. The earnings reported of these insured do not correspond to those actually earned.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using not weighted data.

Table I.5: Weighted shares of groups from the initial dataset, men

Cohort	Foreigners	East Germans	Repatriates (*Aussiedler*)	Self-employed craftsmen	Rejoiners (marriage-law) [A]	"Incomplete" Up to 40	"Complete" Up to 40
1935	0.1540	0.1761	0.0317	0.0235	-	0.1959	0.4187
1936	0.1593	0.1703	0.0281	0.0197	-	0.2166	0.4061
1937	0.1696	0.2016	0.0259	0.0211	-	0.1997	0.3822
1938	0.1704	0.1994	0.0255	0.0124	-	0.2083	0.3839
1939	0.1697	0.2090	0.0178	0.0212	-	0.1792	0.4031
1940	0.1786	0.2093	0.0220	0.0168	-	0.1976	0.3758
1941	0.1719	0.2138	0.0266	0.0144	-	0.1989	0.3743
1942	0.2090	0.2013	0.0222	0.0128	-	0.2006	0.3541
1943	0.1967	0.2169	0.0228	0.0148	-	0.1975	0.3514
1944	0.1964	0.2218	0.0160	0.0172	-	0.1940	0.3547
1945	0.2566	0.1899	0.0160	0.0125	-	0.1927	0.3323
1946	0.2492	0.1527	0.0140	0.0130	-	0.2021	0.3689
1947	0.2357	0.1787	0.0199	0.0107	-	0.1981	0.3570
1948	0.2233	0.1775	0.0230	0.0109	-	0.2095	0.3557
1949	0.2000	0.1982	0.0280	0.0139	-	0.2050	0.3549
1950	0.1920	0.2086	0.0282	0.0130	-	0.1970	0.3611
1951	0.1752	0.2255	0.0350	0.0143	-	0.2015	0.3486
1952	0.1745	0.2273	0.0375	0.0128	-	0.1991	0.3488
1953	0.1697	0.2220	0.0395	0.0102	-	0.2244	0.3342
1954	0.1724	0.2184	0.0348	0.0099	-	0.2121	0.3525
1955	0.1738	0.2305	0.0310	0.0085	-	0.2120	0.3443
1956	0.1743	0.2098	0.0367	0.0087	-	0.2160	0.3545
1957	0.1648	0.2067	0.0364	0.0134	-	0.2258	0.3529
1958	0.1625	0.2112	0.0353	0.0131	-	0.2237	0.3542
1959	0.1545	0.2202	0.0326	0.0133	-	0.2209	0.3585
1960	0.1606	0.2186	0.0279	0.0103	-	0.2230	0.3595
1961	0.1467	0.2337	0.0278	0.0133	-	0.2271	0.3515
1962	0.1529	0.2221	0.0275	0.0095	-	0.2282	0.3598
1963	0.1562	0.2302	0.0206	0.0120	-	0.2364	0.3445
1964	0.1607	0.2230	0.0160	0.0103	-	0.2433	0.3467
1965	0.1658	0.2151	0.0146	0.0115	-	0.2385	0.3545
1966	0.1649	0.2051	0.0124	0.0131	-	0.2458	0.3587
1967	0.1657	0.2031	0.0130	0.0131	-	0.2304	0.3747
1968	0.1788	0.1990	0.0105	0.0111	-	0.2322	0.3685
1969	0.1890	0.1984	0.0090	0.0096	-	0.2295	0.3644

Note: The numbers denote observation share of the respective groups before excluding any observations. [A]Insured who left and rejoined the statutory pension system due to the law of marriage refunds (valid until 1967), which was possible until 1995. The earnings reported of these insured do not correspond to those actually earned.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Table II.8: Weighted number of observations with valid UAX-biographies, men

Birth cohort	Up to 40	Up to 45	Up to 50	Up to 55	Up to 60
1935	214,783	210,073	206,947	197,408	193,415
1936	217,551	212,263	207,186	196,663	191,881
1937	207,309	206,856	203,374	195,114	186,527
1938	221,022	218,897	217,169	211,985	209,564
1939	245,519	236,111	231,068	223,601	219,909
1940	233,767	230,358	228,571	227,105	224,172
1941	216,453	214,801	213,377	210,465	209,591
1942	172,882	169,064	168,109	164,225	164,017
1943	175,621	174,271	171,203	168,750	166,712
1944	173,017	168,663	166,300	161,142	159,641
1945	126,931	125,355	124,422	123,337	122,304
1946	162,292	157,618	155,222	152,719	149,049
1947	178,106	174,483	171,523	167,621	164,812
1948	188,304	183,558	180,554	177,935	173,946
1949	201,483	194,494	189,937	186,256	182,587
1950	210,781	205,003	200,438	197,303	
1951	202,075	198,195	192,300	188,904	
1952	207,547	198,705	194,186	191,816	
1953	198,846	193,264	188,495	185,667	
1954	218,223	210,309	204,987	199,045	
1955	218,160	212,973	207,454	205,115	
1956	232,274	223,581	217,471		
1957	237,176	229,484	225,704		
1958	242,756	236,871	228,939		
1959	258,979	252,939	245,655		
1960	267,044	258,361			
1961	267,736	261,133			
1962	279,379	270,243			
1963	276,530	267,379			
1964	280,680	275,448			
1965	282,497				
1966	283,604				
1967	288,091				
1968	277,011				
1969	261,663				
	7,926,092	6,370,753	4,940,591	3,932,176	2,718,127

Note: Number of observations for a cohort changes because of the selection criterion for valid biographies.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

I.6: Sample selection: Women

Table I.7: Female labor force in West Germany for selected years, SOEP

Year	1988		1994		2000		2006	
Age range	20-53		25-59		31-59		37-59	
Labor force status	Observations	%	Observations	%	Observations	%	Observations	%
Employed[A]	7,688,808	76.48	8,571,636	78.89	7,935,596	80.88	7,112,982	77.5
Unemployed[A]	733,902	7.3	859,543	7.91	550,099	5.61	887,568	9.67
Apprentice[A]	448,453	4.46	71,369	0.66	9,951	0.1	16,838	0.18
Miner[A,B]	0	0	0	0	30,991	0.32	796	0.01
Sum of items above[A]	8,871,163	88.24	9,502,548	87.45	7,980,085	81.33	8,018,184	87.36
Civil servant	484,884	4.82	621,621	5.72	520,966	5.31	465,570	5.07
Self-employed	697,280	6.94	741,731	6.83	760,714	7.75	694,739	7.57
Total	10,053,327	100.00	10,865,900	100.00	9,811,864	100.01	9,178,493	100.00

Note: Sample selection mirrors the respective birth cohorts in our deployed FDZ-RV − VSKT2002, 2004-2009_Bönke data.
[A]Labor force covered in the VSKT, [B]not weighted cell size < 5.
Source: SOEP v28, own calculations using weighted data.

Table I.8: Number of observed women with valid UAX-biographies

| Birth cohort	Up to 40	Up to 45	Up to 50	Up to 55	Up to 60
1935	344	332	318	311	313
1936	354	349	349	329	336
1937	381	366	348	336	346
1938	407	403	387	360	362
1939	373	360	361	356	346
1940	425	420	419	432	440
1941	433	420	422	440	438
1942	468	463	472	479	476
1943	519	506	496	500	505
1944	479	481	481	478	472
1945	531	544	537	528	519
1946	525	502	492	496	500
1947	539	521	517	527	517
1948	508	501	512	514	500
1949	521	526	531	523	514
1950	504	515	507	512	
1951	522	535	538	539	
1952	551	539	533	528	
1953	528	523	526	522	
1954	499	503	514	511	
1955	558	569	565		
1956	503	527	550		
1957	567	567	581		
1958	501	526	541		
1959	552	567	578		
1960	545	553			
1961	565	561			
1962	654	654			
1963	647	646			
1964	657	661			
1965	692				
1966	657				
1967	700				
1968	698				
1969	888				
	18,795	15,140	12,075	9,221	6,584

Note: Number of observations for a cohort changes because of the selection criterion for valid biographies.
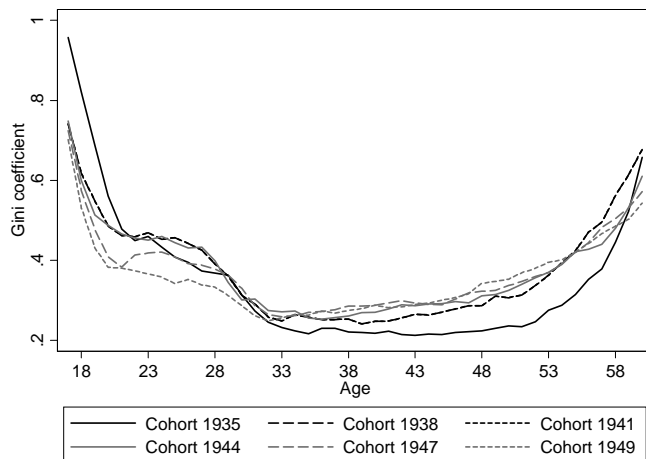Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using not weighted data.

Table I.9: Weighted number of observations with valid UA-40 biographies, women

Birth cohort	Observations with valid UA40-biographies (weighted)	Actual cohort size at age 40			
		without foreigners		including foreigners	
		size	coverage[A]	size	coverage[A]
1935	214,783	422058	0.1247	442171	0.1190
1936	217,551	429765	0.1242	450445	0.1185
1937	207,309	432061	0.1166	453370	0.1111
1938	221,022	456876	0.1355	480458	0.1288
1939	245,519	483157	0.1244	508572	0.1182
1940	233,767	486274	0.1330	515466	0.1254
1941	216,453	449906	0.1592	477378	0.1500
1942	172,882	362910	0.1667	394090	0.1535
1943	175,621	371086	0.1970	402179	0.1818
1944	173,017	361942	0.1928	394009	0.1771
1945	126,931	265921	0.2093	299934	0.1856
1946	162,292	309634	0.2393	346937	0.2136
1947	178,106	341762	0.2418	381039	0.2169
1948	188,304	360960	0.2321	404417	0.2072
1949	201,483	385314	0.2593	432152	0.2312
1950	210,781	.	.	440548	0.2118
1951	202,075	.	.	441156	0.2061
1952	207,547	.	.	452801	0.2179
1953	198,846	.	.	450526	0.2135
1954	218,223	.	.	464567	0.1989
1955	218,160	.	.	469187	0.2146
1956	232,274	.	.	486236	0.2007
1957	237,176	.	.	499854	0.2105
1958	242,756	.	.	508413	0.1921
1959	258,979	.	.	533458	0.2056
1960	267,044	.	.	553228	0.1863
1961	267,736	.	.	568443	0.1906
1962	279,379	.	.	577470	0.2007
1963	276,530	.	.	596598	0.1959
1964	280,680	.	.	602147	0.1963
1965	282,497	.	.	593553	0.2189
1966	283,604	.	.	595330	0.2034
1967	288,091	.	.	583619	0.2185
1968	277,011	.	.	569535	0.2093
1969	261,663	.	.	542073	0.2639

Note: Cohorts 1935 – 1949: West Germany including West Berlin, cohorts 1950 – 1969: West Germany including Berlin.
[A]Coverage equals the number of observations with valid UA-40 biographies (weighted) divided by actual cohort size at 40.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke. Actual cohort size at age 40 according to Federal Statistical Office, own calculations using weighted data.

Table I.10: Number of observed women

		Sample size after exclusion of:					
Cohort	Original number of observations	Foreigners	East Germans	Repatriates (*Aussiedler*)	Self-employed craftsmen	Rejoiners (marriage-law)[A]	Ccomplete biographies (up to age 40)
1935	4,678	3,359	2,546	2,390	2,387	2,072	344
1936	4,753	3,589	2,596	2,433	2,428	2,049	354
1937	5,004	3,766	2,777	2,563	2,553	2,144	381
1938	4,784	3,634	2,674	2,498	2,485	2,078	407
1939	4,900	3,705	2,635	2,458	2,446	2,016	373
1940	5,004	3,783	2,711	2,570	2,564	2,157	425
1941	4,923	3,608	2,575	2,440	2,432	2,051	433
1942	4,879	3,668	2,637	2,491	2,482	2,155	468
1943	4,777	3,572	2,593	2,485	2,472	2,228	519
1944	4,614	3,499	2,481	2,375	2,364	2,181	479
1945	4,423	3,249	2,442	2,333	2,316	2,196	531
1946	4,090	3,005	2,378	2,300	2,286	2,209	525
1947	4,071	3,049	2,275	2,175	2,157	2,110	539
1948	4,076	2,987	2,263	2,162	2,136	2,124	508
1949	3,969	2,892	2,165	2,054	2,040	2,033	521
1950	4,104	3,010	2,133	1,982	1,972	1,972	504
1951	4,259	3,138	2,211	2,047	2,035	2,035	522
1952	4,283	3,115	2,198	2,049	2,030	2,030	551
1953	4,381	3,161	2,196	2,011	1,995	1,995	528
1954	4,495	3,182	2,185	1,983	1,964	1,964	499
1955	4,591	3,196	2,239	2,066	2,055	2,055	558
1956	4,684	3,158	2,242	2,053	2,044	2,044	503
1957	4,812	3,209	2,330	2,167	2,150	2,150	567
1958	4,973	3,172	2,321	2,144	2,129	2,129	501
1959	5,218	3,261	2,379	2,222	2,207	2,207	552
1960	5,579	3,381	2,446	2,300	2,288	2,288	545
1961	5,886	3,533	2,555	2,412	2,399	2,399	565
1962	6,538	3,804	2,675	2,558	2,542	2,542	654
1963	6,550	3,723	2,672	2,537	2,519	2,519	647
1964	6,769	3,673	2,685	2,577	2,564	2,564	657
1965	6,995	3,799	2,711	2,621	2,605	2,605	692
1966	6,949	3,706	2,732	2,649	2,624	2,624	657
1967	7,002	3,689	2,695	2,624	2,609	2,609	700
1968	7,197	3,683	2,749	2,674	2,663	2,663	698
1969	7,390	3,725	2,785	2,718	2,703	2,703	888
	181,600	119,683	86,887	82,121	81,645	77,900	18,795

Note: Second column: Original sample size. Columns three to eight: All numbers denote observation after the stepwise exclusion of the respective groups. [A]Insured who left and rejoined the statutory pension system due to the law of marriage refunds (valid until 1967), which was possible until 1995. The earnings reported of these insured do not correspond to those actually earned.

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using not weighted data.

Table I.11: Weighted shares of groups from the initial dataset, women

Cohort	Foreigners	East Germans	Repatriates (*Aussiedler*)	Self-employed craftsmen	Rejoiners (marriage-law) [A]	"Incomplete" Up to 40	"Complete" Up to 40
1935	0.0618	0.2086	0.0362	0.0008	0.0897	0.5902	0.1024
1936	0.0635	0.2055	0.0356	0.0014	0.1059	0.5927	0.1013
1937	0.0697	0.2568	0.0446	0.0027	0.1004	0.5318	0.0944
1938	0.0724	0.2481	0.0370	0.0031	0.1034	0.5299	0.1094
1939	0.0736	0.2610	0.0355	0.0025	0.1066	0.5277	0.0997
1940	0.0822	0.2819	0.0271	0.0010	0.0920	0.5020	0.1058
1941	0.0842	0.2710	0.0279	0.0016	0.0851	0.4905	0.1249
1942	0.1082	0.2605	0.0270	0.0017	0.0758	0.4749	0.1276
1943	0.1082	0.2611	0.0238	0.0020	0.0563	0.4547	0.1503
1944	0.1130	0.2521	0.0221	0.0020	0.0417	0.4650	0.1457
1945	0.1557	0.2197	0.0221	0.0028	0.0292	0.4483	0.1513
1946	0.1580	0.1719	0.0177	0.0028	0.0175	0.4725	0.1772
1947	0.1526	0.2077	0.0253	0.0030	0.0124	0.4386	0.1727
1948	0.1543	0.1853	0.0244	0.0045	0.0027	0.4665	0.1650
1949	0.1456	0.1976	0.0311	0.0029	0.0012	0.4413	0.1815
1950	0.1492	0.2207	0.0420	0.0017	0.0000	0.4232	0.1632
1951	0.1368	0.2327	0.0404	0.0026	0.0000	0.4273	0.1602
1952	0.1397	0.2390	0.0341	0.0030	0.0000	0.4142	0.1700
1953	0.1315	0.2352	0.0445	0.0028	0.0000	0.4201	0.1659
1954	0.1289	0.2470	0.0468	0.0029	0.0000	0.4192	0.1552
1955	0.1267	0.2356	0.0395	0.0016	0.0000	0.4292	0.1673
1956	0.1198	0.2250	0.0469	0.0019	0.0000	0.4482	0.1582
1957	0.1116	0.2265	0.0387	0.0033	0.0000	0.4535	0.1664
1958	0.1126	0.2328	0.0422	0.0023	0.0000	0.4584	0.1517
1959	0.1079	0.2338	0.0389	0.0030	0.0000	0.4551	0.1614
1960	0.1113	0.2381	0.0355	0.0023	0.0000	0.4649	0.1479
1961	0.1032	0.2379	0.0370	0.0024	0.0000	0.4685	0.1510
1962	0.1087	0.2530	0.0254	0.0028	0.0000	0.4499	0.1603
1963	0.1094	0.2432	0.0322	0.0029	0.0000	0.4548	0.1575
1964	0.1148	0.2413	0.0256	0.0024	0.0000	0.4578	0.1581
1965	0.1196	0.2258	0.0193	0.0025	0.0000	0.4552	0.1776
1966	0.1212	0.2137	0.0193	0.0037	0.0000	0.4760	0.1660
1967	0.1241	0.2138	0.0178	0.0027	0.0000	0.4621	0.1796
1968	0.1362	0.2032	0.0161	0.0019	0.0000	0.4699	0.1727
1969	0.1500	0.1897	0.0131	0.0024	0.0000	0.4283	0.2166

Note: The numbers denote observation share of the respective groups before excluding any observations. [A]Insured who left and rejoined the statutory pension system due to the law of marriage refunds (valid until 1967), which was possible until 1995. The earnings reported of these insured do not correspond to those actually earned.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Table I.12: Weighted number of observations with valid UAX-biographies, women

Birth cohort	Up to 40	Up to 45	Up to 50	Up to 55	Up to 60
1935	52,622	50,353	48,384	47,316	48,175
1936	53,356	52,572	53,091	49,557	51,276
1937	50,362	47,803	45,417	43,887	45,512
1938	61,887	61,498	58,893	55,125	55,975
1939	60,105	58,912	58,727	57,776	57,050
1940	64,653	64,182	64,474	66,696	69,552
1941	71,620	68,661	69,313	72,244	70,475
1942	60,512	60,087	60,867	61,635	62,068
1943	73,109	71,529	69,801	70,128	71,327
1944	69,795	70,204	70,192	69,223	68,046
1945	55,662	56,701	54,976	54,022	53,185
1946	74,109	70,441	68,875	68,793	70,248
1947	82,645	79,152	79,550	80,563	79,427
1948	83,779	83,093	83,629	83,924	83,742
1949	99,915	99,882	100,330	97,999	96,264
1950	93,323	95,009	95,676	95,400	
1951	90,938	93,035	93,411	93,860	
1952	98,684	97,226	96,700	96,101	
1953	96,190	95,506	95,110	95,126	
1954	92,391	93,927	96,216	94,149	
1955	100,683	103,887	103,634		
1956	97,569	101,439	105,649		
1957	105,207	107,349	109,690		
1958	97,649	99,741	101,691		
1959	109,672	111,108	113,736		
1960	103,066	103,457			
1961	108,367	107,331			
1962	115,887	116,899			
1963	116,896	116,696			
1964	118,184	117,276			
1965	129,950				
1966	121,063				
1967	127,507				
1968	119,226				
1969	143,057				
	3,199,640	2,554,956	1,998,032	1,453,524	982,322

Note: Number of observations for a cohort changes because of the selection criterion for valid biographies.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Appendix II: Results for women

This section replicates graphs 1-10 from the paper for women and provides information on their educational attainment.

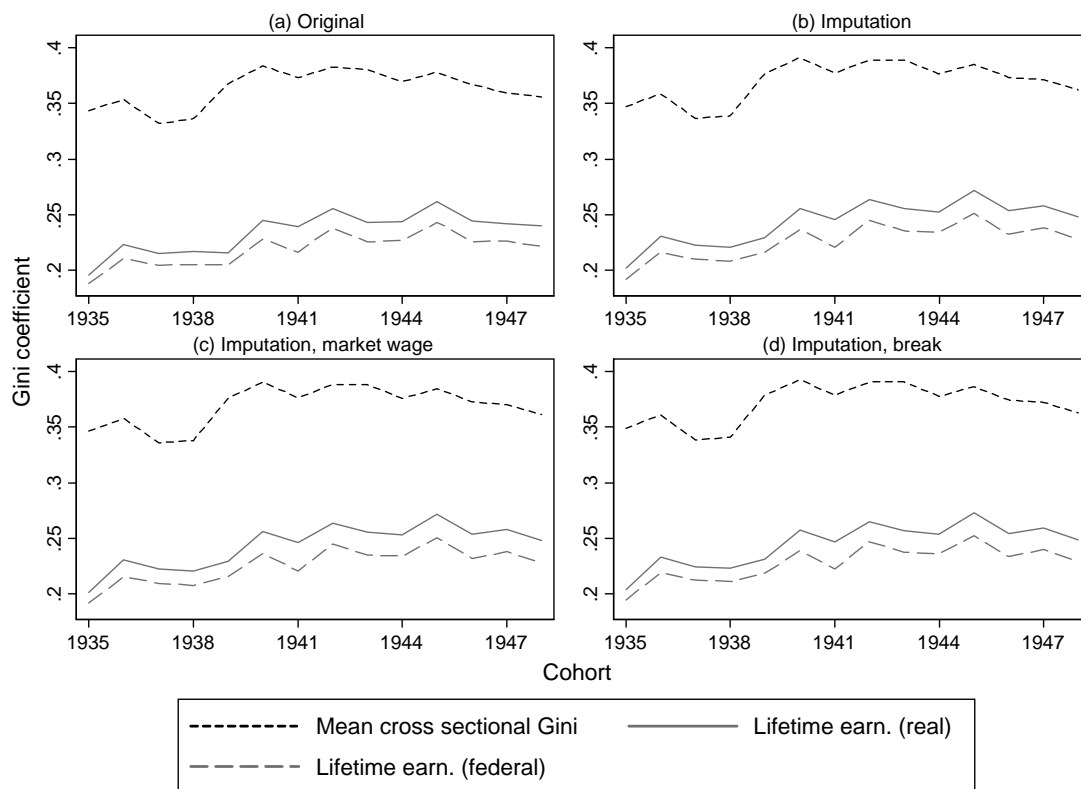Figure II.1: Means of annual Gini coefficients and Gini coefficients of lifetime earnings for cohorts 1935 - 1949
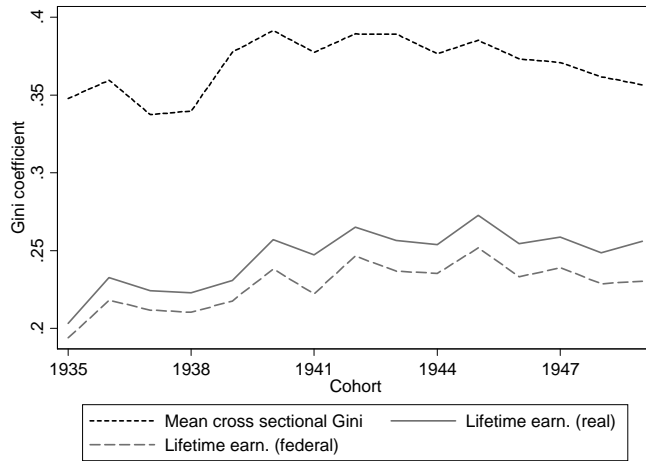
Note: real denotes CPI discounting, federal denotes federal bond discounting.
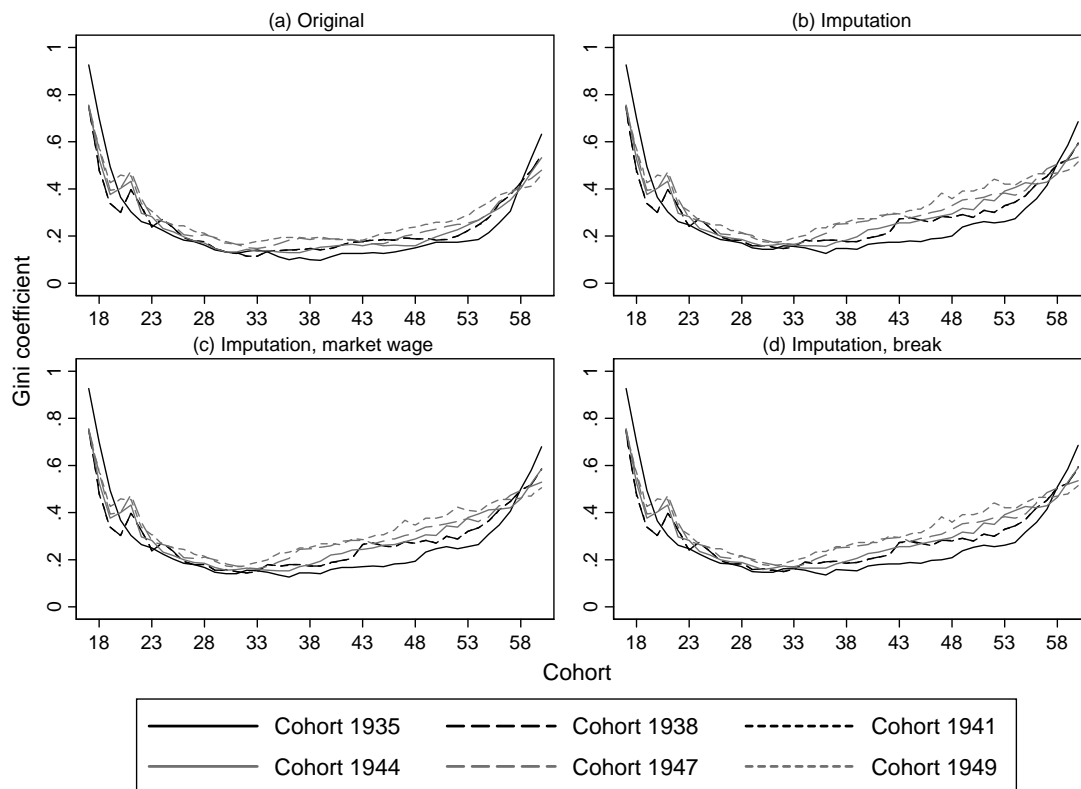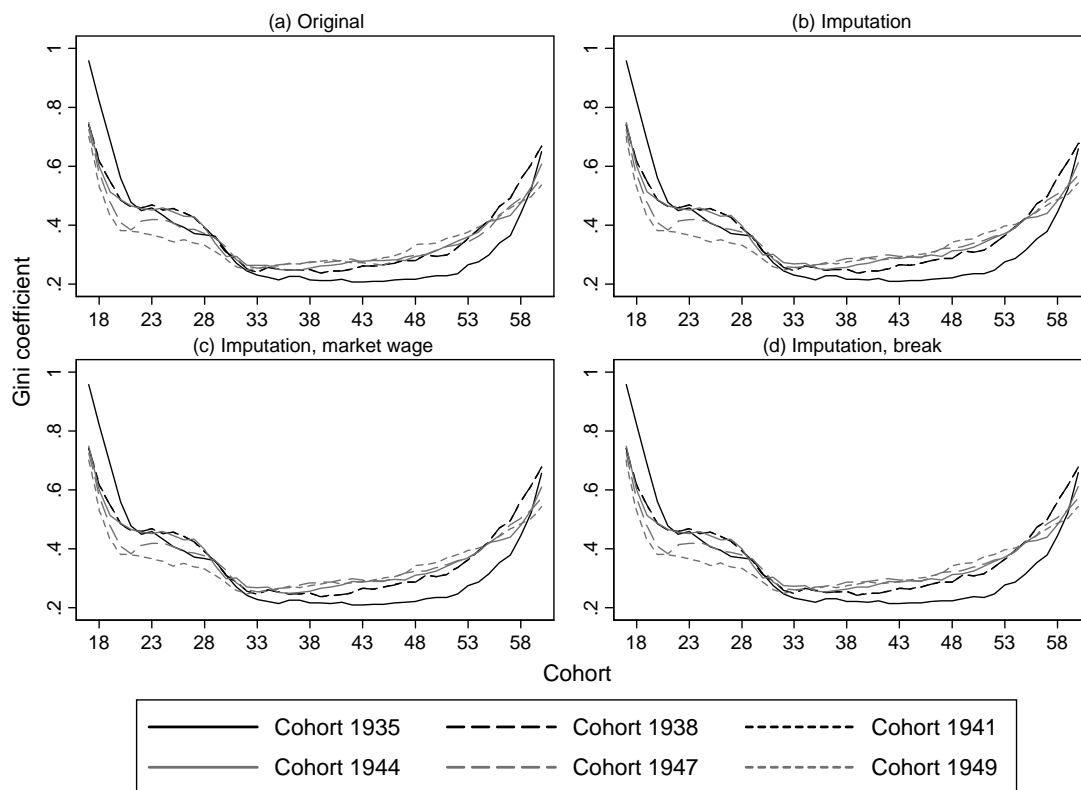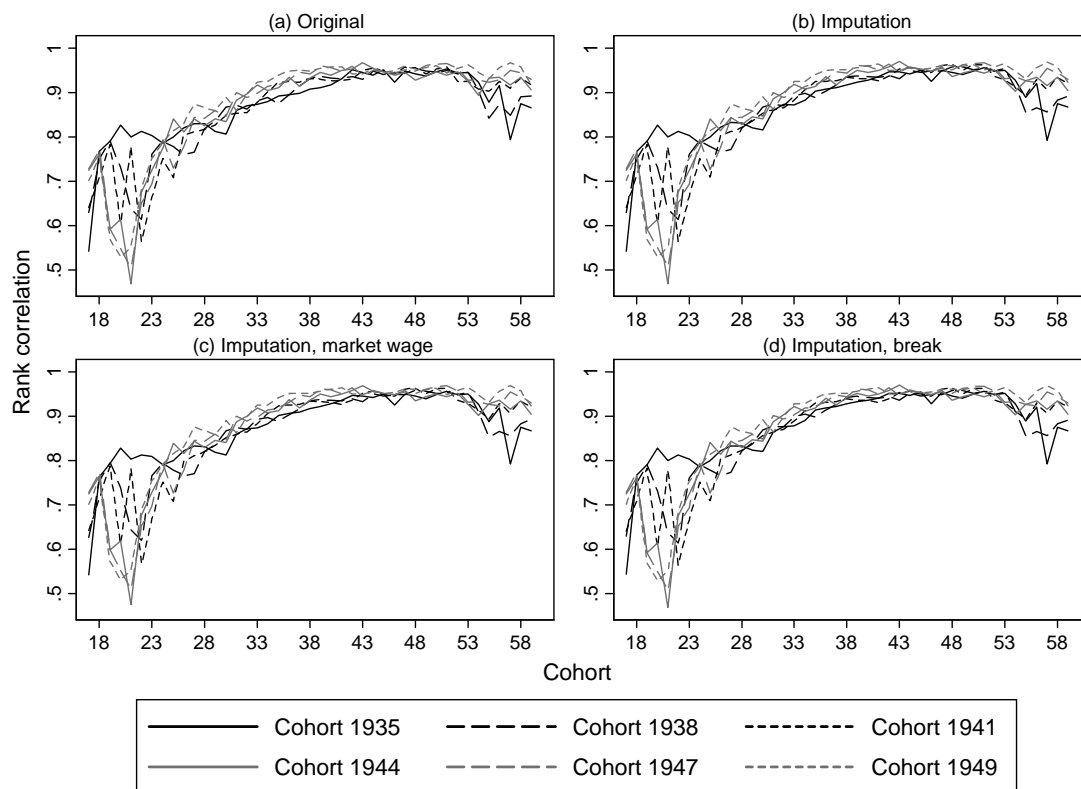Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

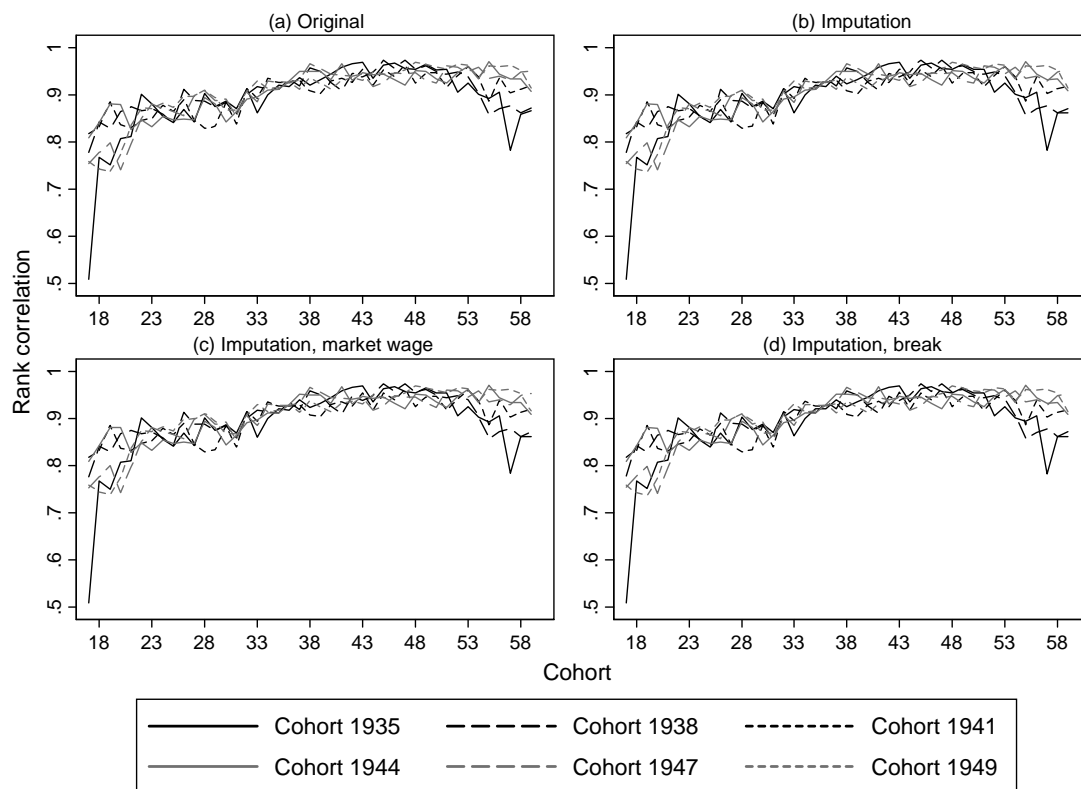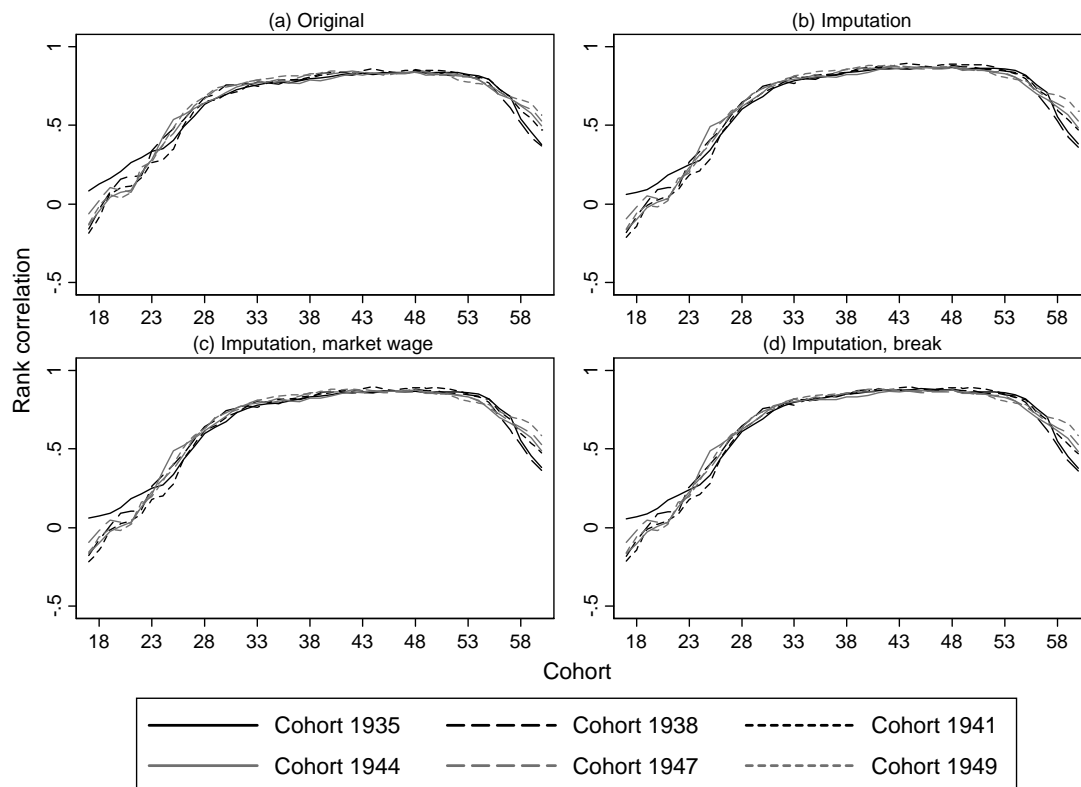Figure II.2: Annual Gini coefficients from age 17 to age 60 for cohorts 1935 - 1949

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

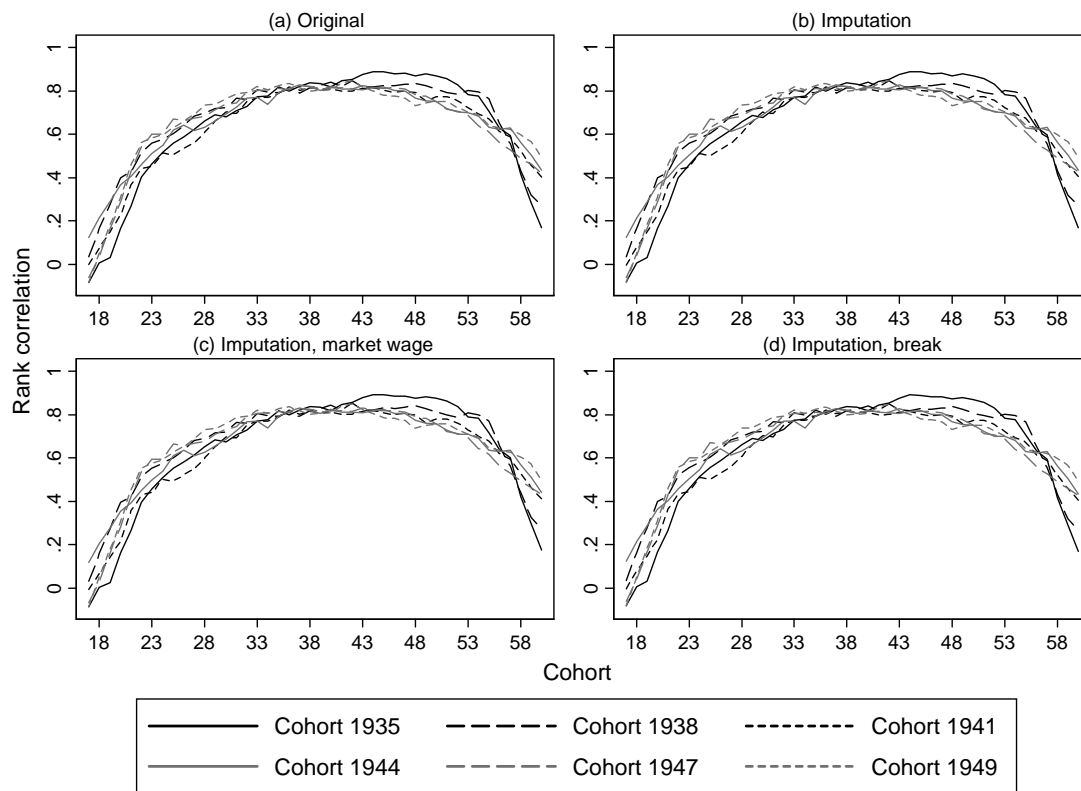Figure II.3: Earnings rank correlations between consecutive years for cohorts 1935-1949
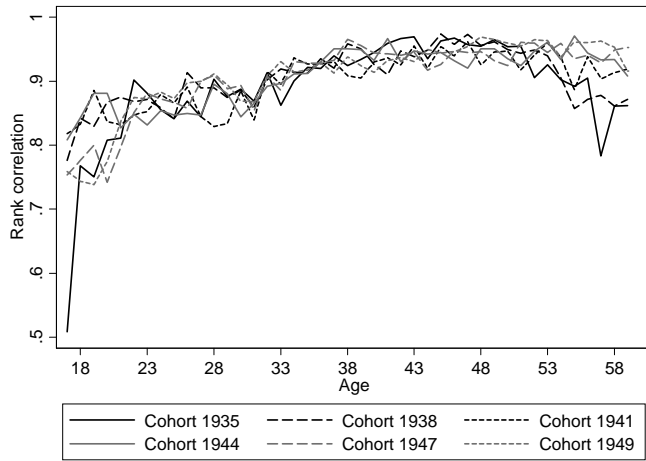
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure II.4: Rank correlation of annual and lifetime earnings for cohorts 1935-1949
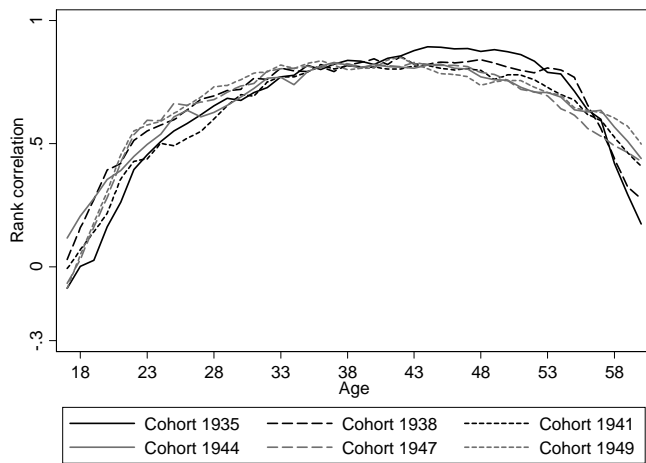
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure II.5: Decomposition of changes in inequality as of Eq. (3) for cohort 1944
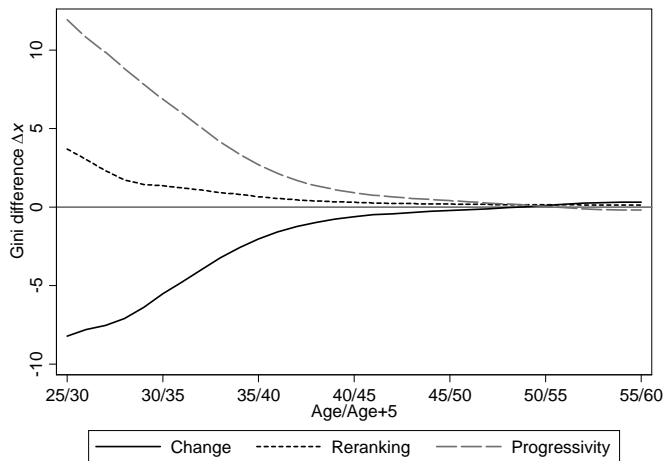
Note: Accumulated discounted earnings refer to the age in the abscissa as compared to accumulated earnings five years later, as in Eq. (3) in the paper. Coefficients are multiplied by 100.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure II.6: Age-earning-profiles by highest educational attainment for pooled cohorts 1935-1949
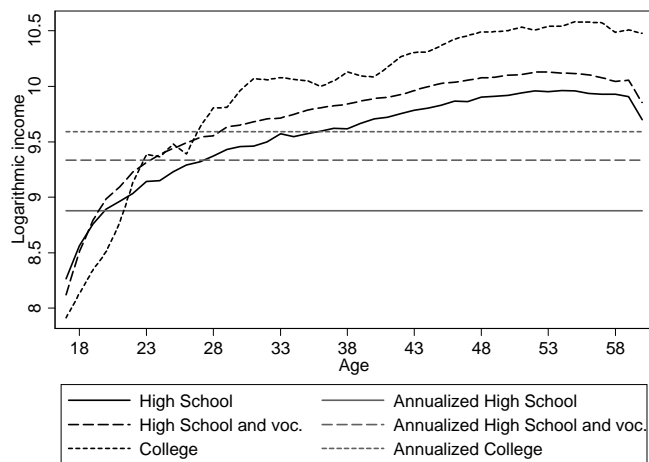
Note: voc. abbreviates vocational training.

Figure II.7: Gini coefficients of UAX for cohorts 1935-1969
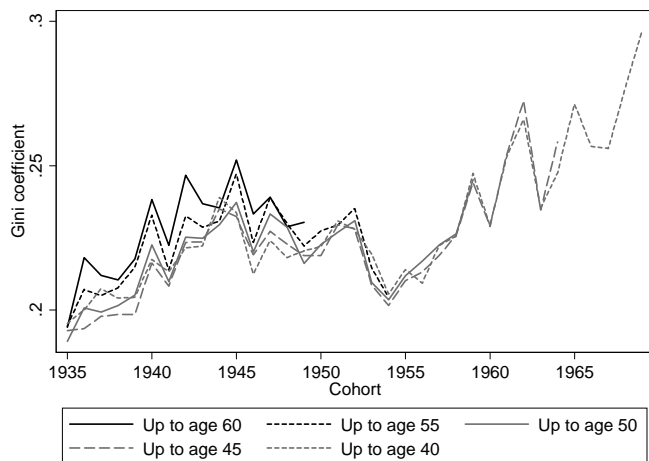
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure II.8: 85th / 15th ratio of UAX- earnings for cohorts 1935-1969
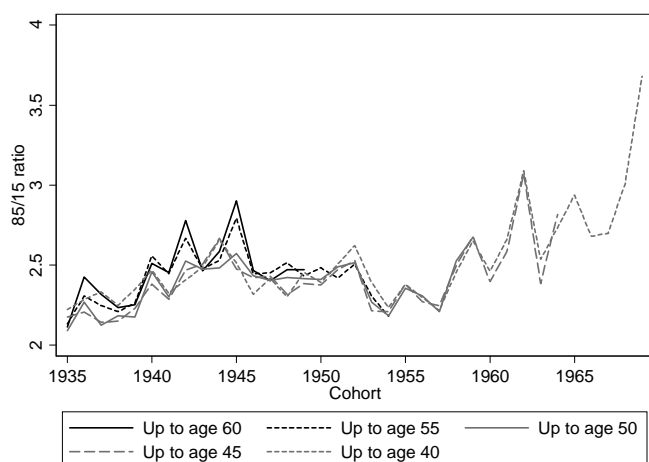
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure II.9: Gini coefficients of annual earnings at various ages for cohorts 1935-1969
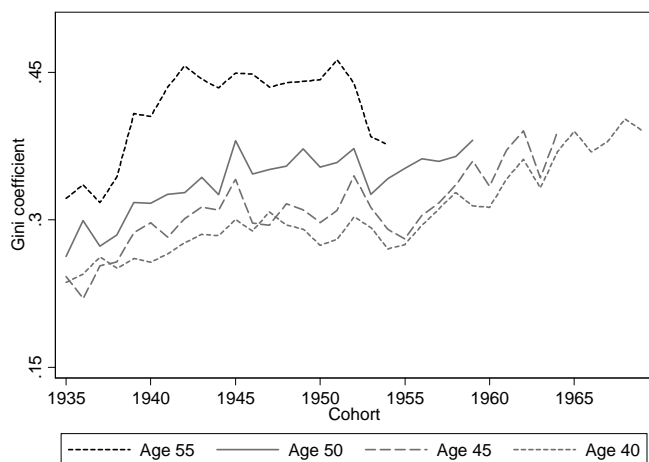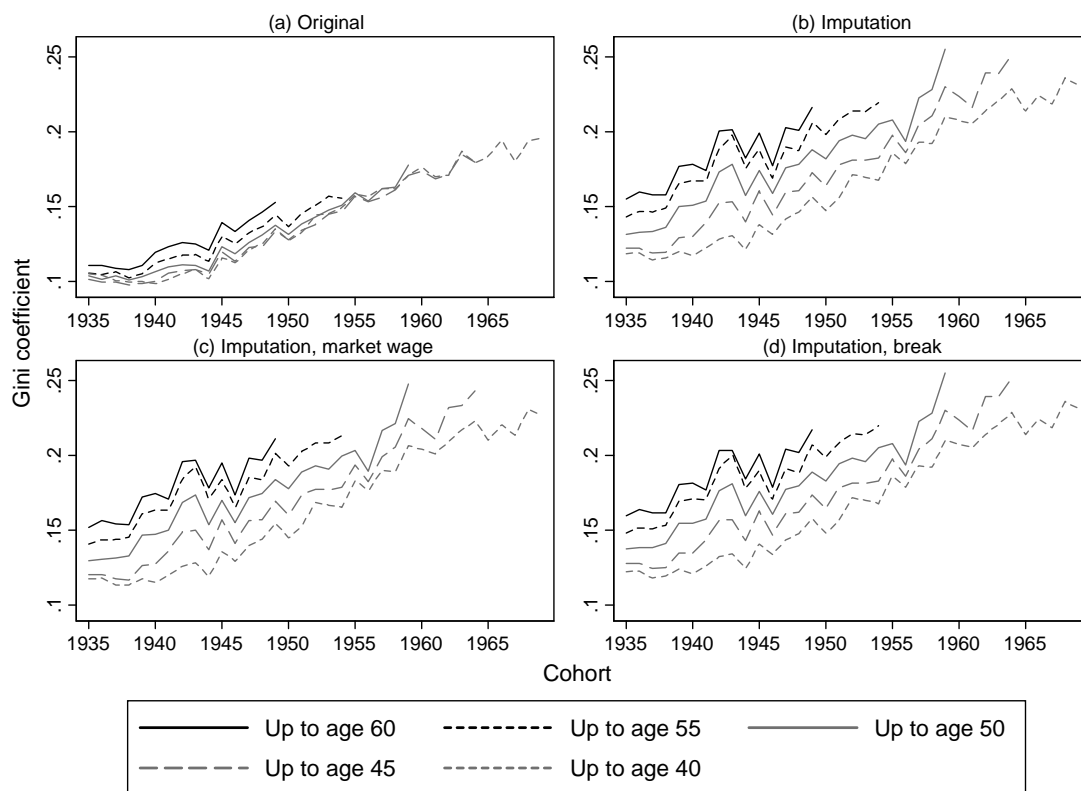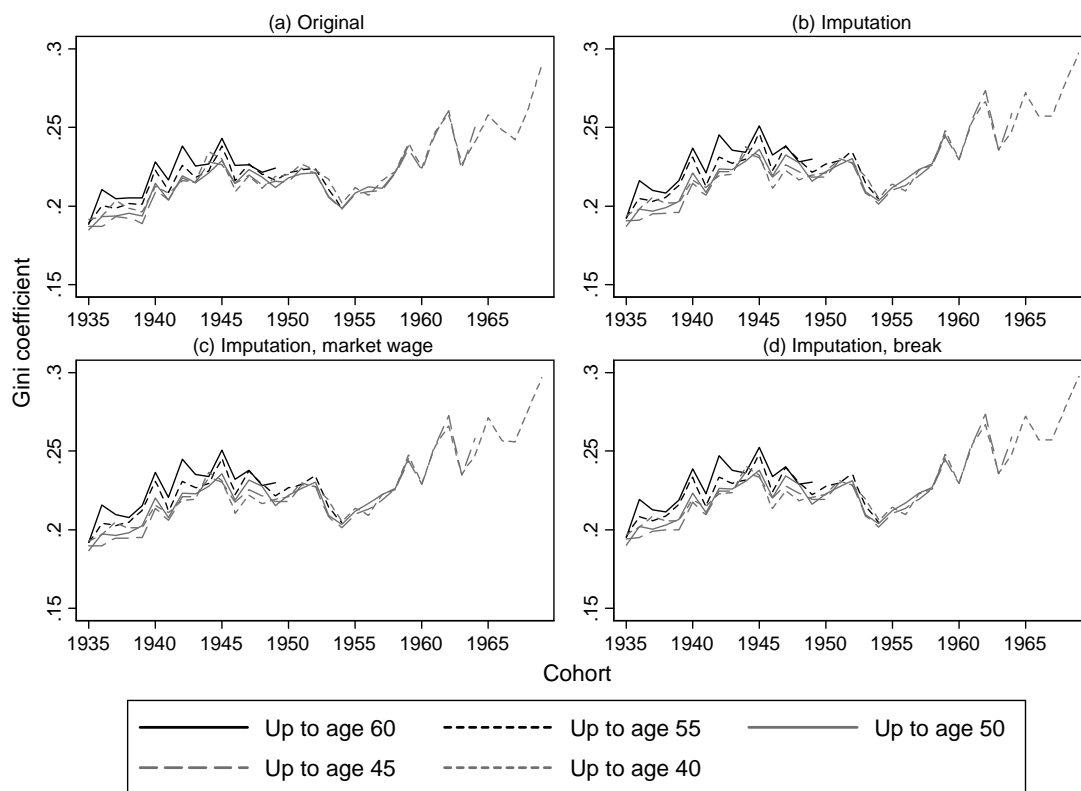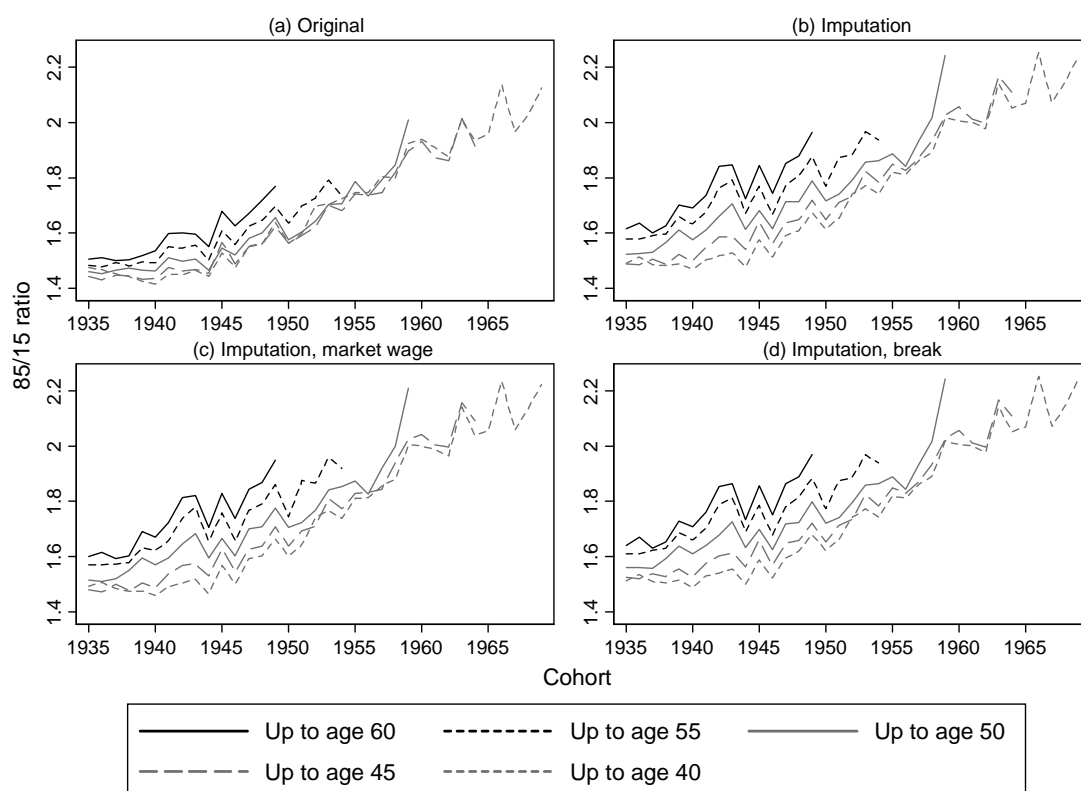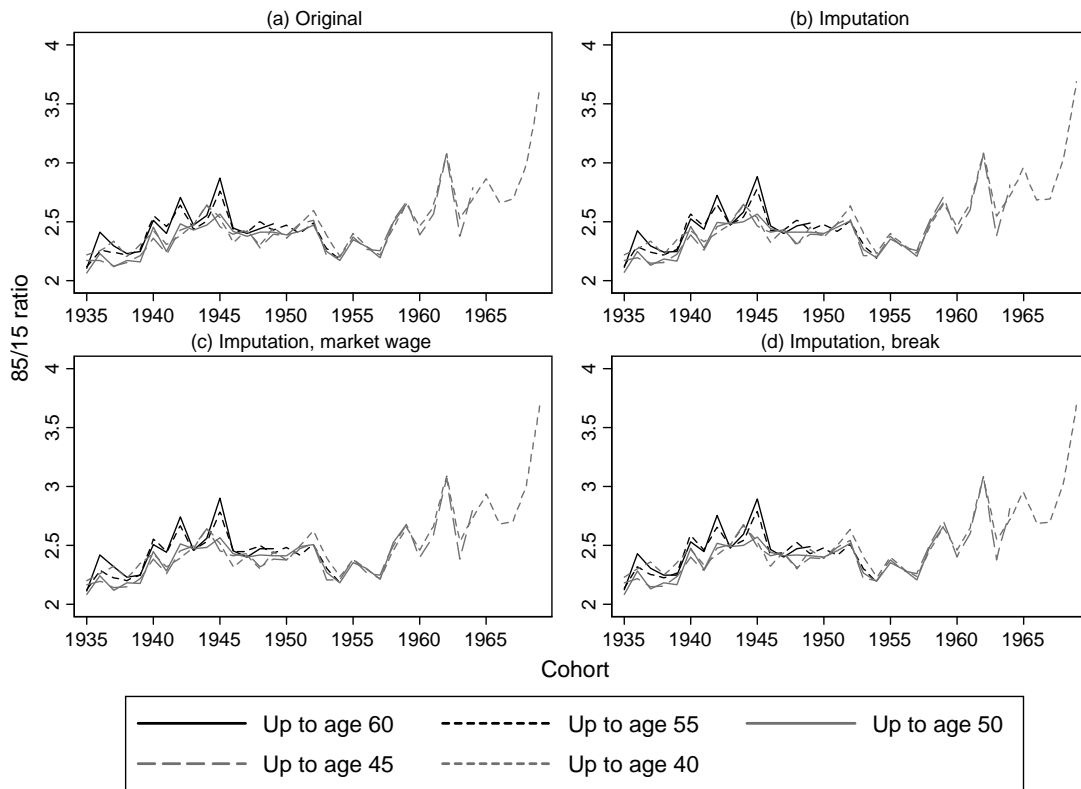

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure II.10: Rank correlation of UA-40 with selected UAX for cohorts 1935-1969
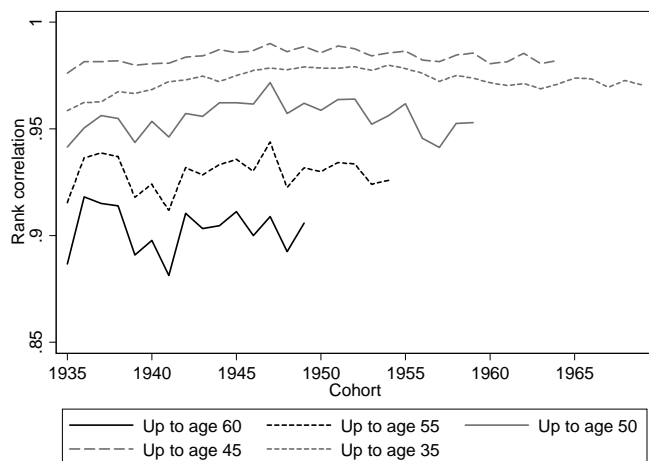
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure II.11: Educational attainment and inequality in our sample, women
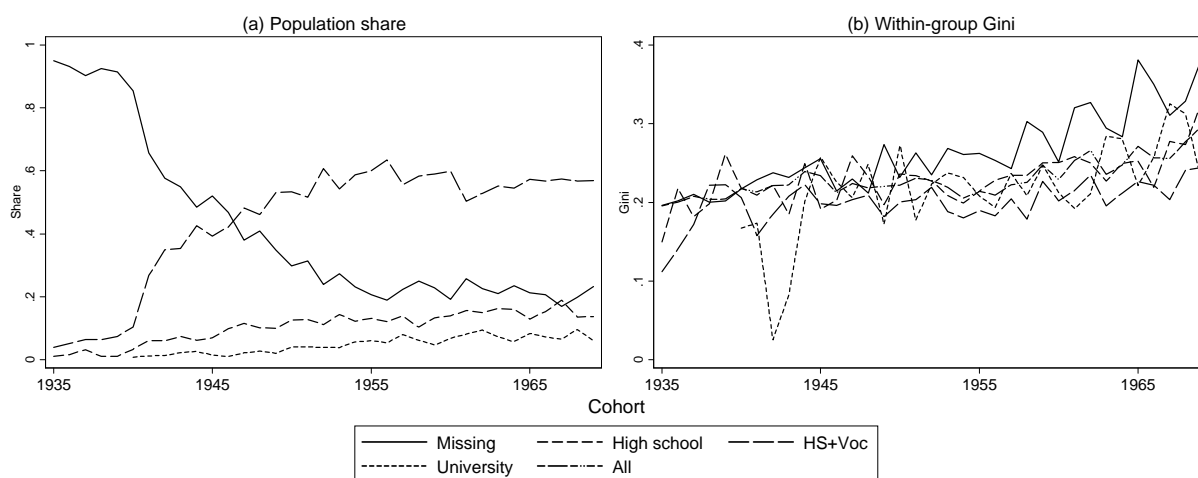
Note: Within-group Gini coefficients refer to the distributions of UA-40 with federal bond discounting.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure II.12: Educational attainment of cohorts of West German women according to the SOEP
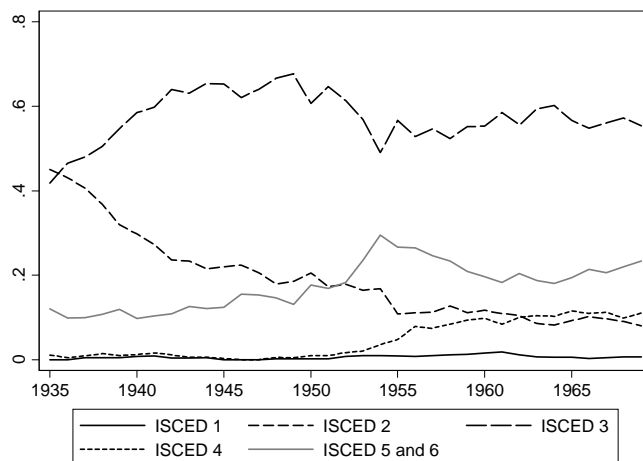
Note: The education groups are defined according to the International Standard Classification of Education 1997 (ISCED-97): ISCED 1: Primary education; ISCED 2: Lower secondary education; ISCED 3: Upper secondary education; ISCED 4: Post-secondary non-tertiary education; ISCED 5/6: Tertiary education. All cohorts born after 1944 are analyzed at age 40. Since the SOEP starts in 1984, older cohorts are analyzed at the closest distance to age 40, e.g. age 45 for those born in 1939. Source: SOEP v28, own calculations using weighted data.

Appendix III: Robustness and supplementary graphics

III.1 Confidence intervals for UAX-earnings

Table III.1: UAX Ginis for selected cohorts, men

Cohort	Up to 40	Up to 45	Up to 50	Up to 55	Up to 60
1935	0.121	0.125	0.135	0.145	0.156
	(0.114; 0.128)	(0.119; 0.134)	(0.127; 0.144)	(0.137; 0.155)	(0.145; 0.167)
1940	0.118	0.131	0.151	0.166	0.177
	(0.111; 0.126)	(0.123; 0.141)	(0.141; 0.164)	(0.155; 0.182)	(0.165; 0.192)
1945	0.138	0.159	0.172	0.185	0.196
	(0.130; 0.147)	(0.149; 0.172)	(0.160; 0.186)	(0.173; 0.203)	(0.183; 0.214)
1950	0.146	0.161	0.178	0.193	
	(0.138; 0.156)	(0.151; 0.173)	(0.167; 0.196)	(0.179; 0.210)	
1955	0.183	0.194	0.204		
	(0.173; 0.195)	(0.182; 0.208)	(0.191; 0.220)		
1960	0.204	0.218			
	(0.192; 0.218)	(0.205; 0.234)			
1965	0.210				
	(0.200; 0.223)				
1969	0.227				
	(0.215; 0.239)				

Note: The UAX are based on federal bond discounting. Bias corrected and accelerated bootstrap confidence intervals at the 95%-level in brackets.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Table III.2: UAX Ginis for selected cohorts, women

Cohort	Up to 40	Up to 45	Up to 50	Up to 55	Up to 60
1935	0.195	0.193	0.189	0.194	0.194
	(0.183; 0.210)	(0.180; 0.211)	(0.176; 0.203)	(0.180; 0.211)	(0.181; 0.212)
1940	0.218	0.216	0.223	0.233	0.238
	(0.203; 0.233)	(0.200; 0.233)	(0.207; 0.241)	(0.217; 0.248)	(0.221; 0.257)
1945	0.234	0.232	0.237	0.247	0.252
	(0.219; 0.249)	(0.218; 0.246)	(0.223; 0.254)	(0.233; 0.266)	(0.237; 0.270)
1950	0.222	0.219	0.223	0.227	
	(0.208; 0.240)	(0.204; 0.235)	(0.210; 0.238)	(0.214; 0.244)	
1955	0.214	0.210	0.212		
	(0.200; 0.229)	(0.196; 0.225)	(0.199; 0.228)		
1960	0.229	0.229			
	(0.215; 0.246)	(0.214; 0.248)			
1965	0.271				
	(0.254; 0.293)				
1969	0.297				
	(0.282; 0.313)				

Note: The UAX are based on federal bond discounting. Bias corrected and accelerated bootstrap confidence intervals at the 95%-level in brackets.
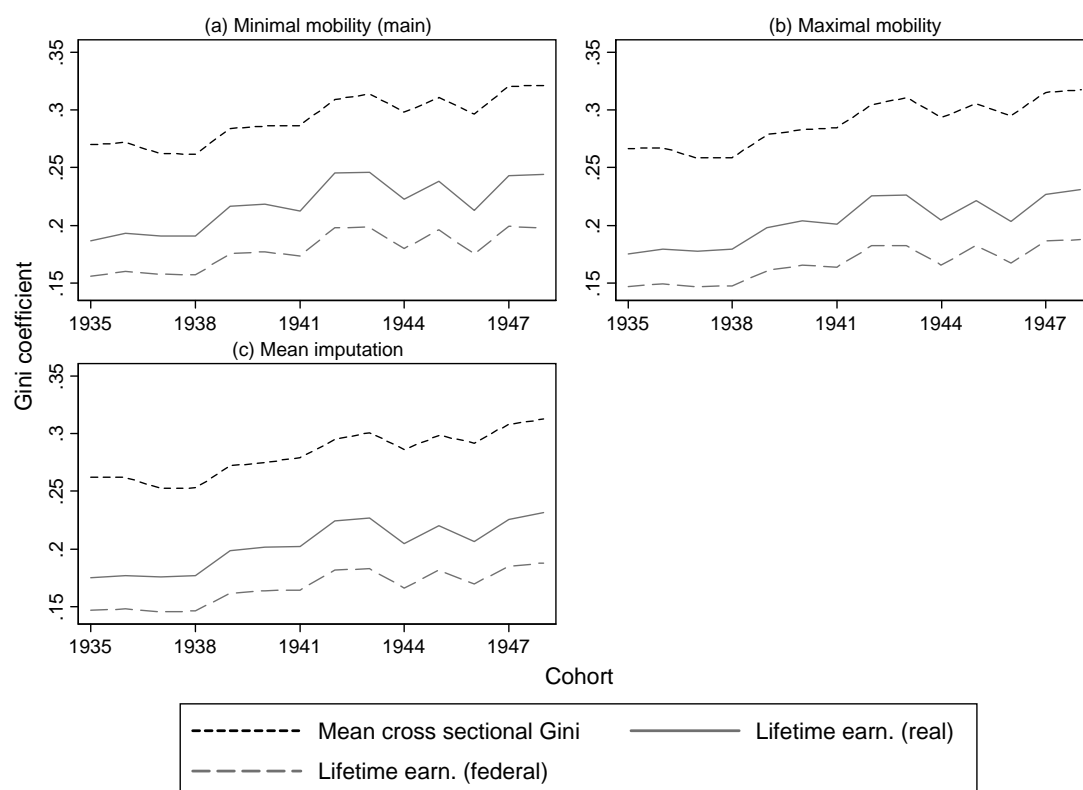Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

III.2: Alternative earnings concepts

Annotation: The calculations in this section are based on federal bond discounting unless stated otherwise. In order to obtain the earnings concept reported in the paper we apply three changes to the original earnings: the imputation of top coded earnings, the inclusion of the employers' social security contributions and the correction of the structural break. In this section we provide results for four alternative earnings concepts (see Table II.3):

(a) Original: Earnings as recorded in the dataset with no changes applied.

(b) Imputation: Original earnings with imputation of top coded earnings.

(c) Imputation, market wage: Original earnings with imputation of top coded earnings plus employers' social security contributions.

(d) Imputation, break: Original earnings with imputation of top coded earnings and correction of structural break.

Table III.3: Alternative earnings concepts

Earnings concept	Imputation of top-coded earnings	Including employers' social security contributions	Correction of the structural break
(a) Original			
(b) Imputation	X		
(c) Imputation, market wage	X	X	
(d) Imputation, break	X		X
Main concept in the paper	X	X	X

Note: X marks if the change is included in the respective earnings concept.

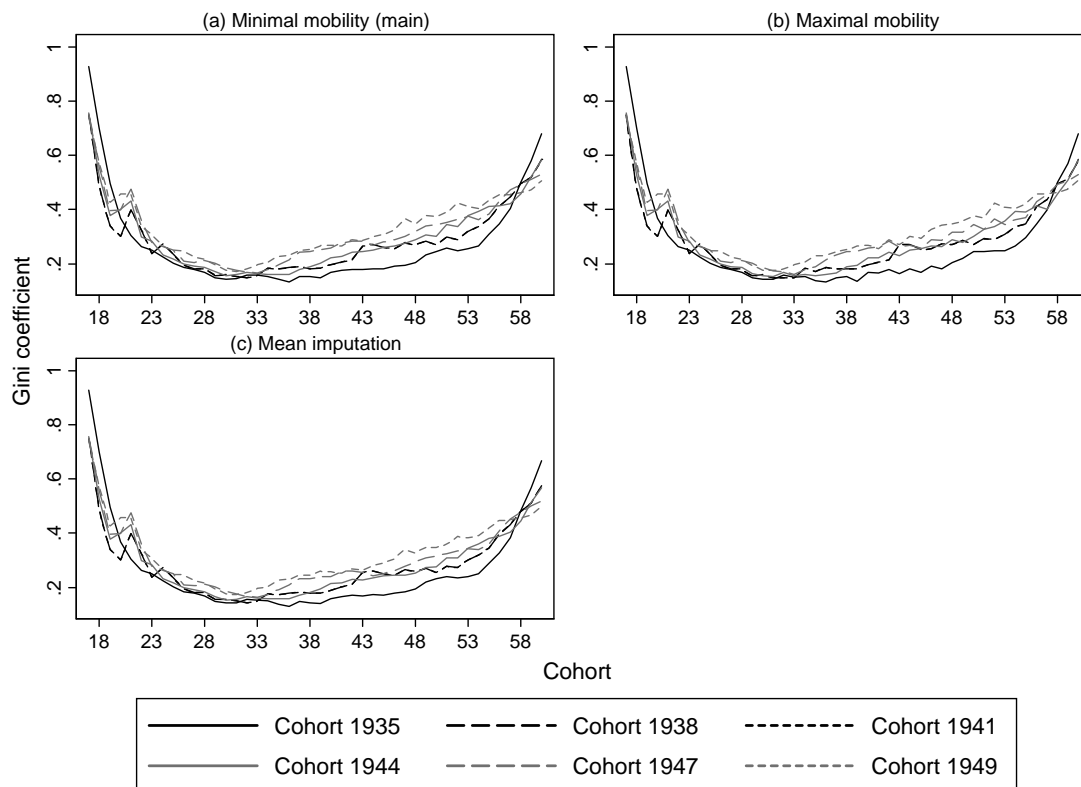Figure III.1: Means of annual Gini coefficients and Gini coefficients of lifetime earnings, men
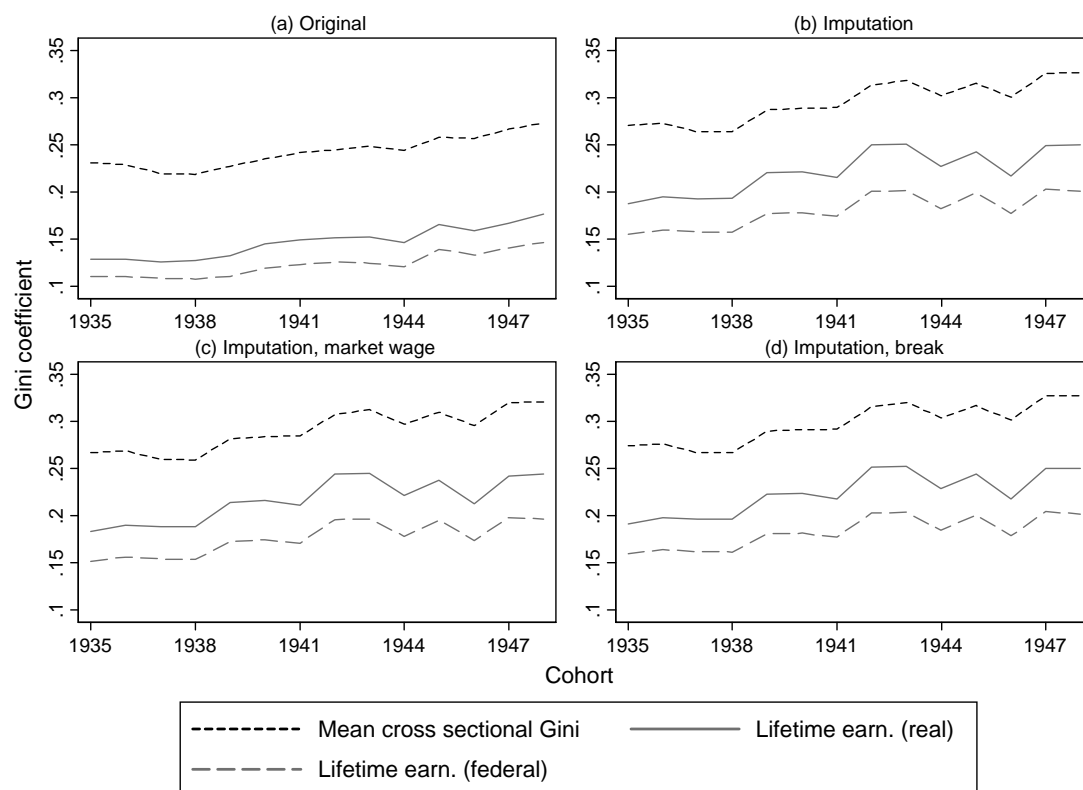
Note: "real" denotes CPI discounting, "federal" denotes federal bond discounting.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.2: Means of annual Gini coefficients and Gini coefficients of lifetime earnings, women

Note: "real" denotes CPI discounting, "federal" denotes federal bond discounting.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

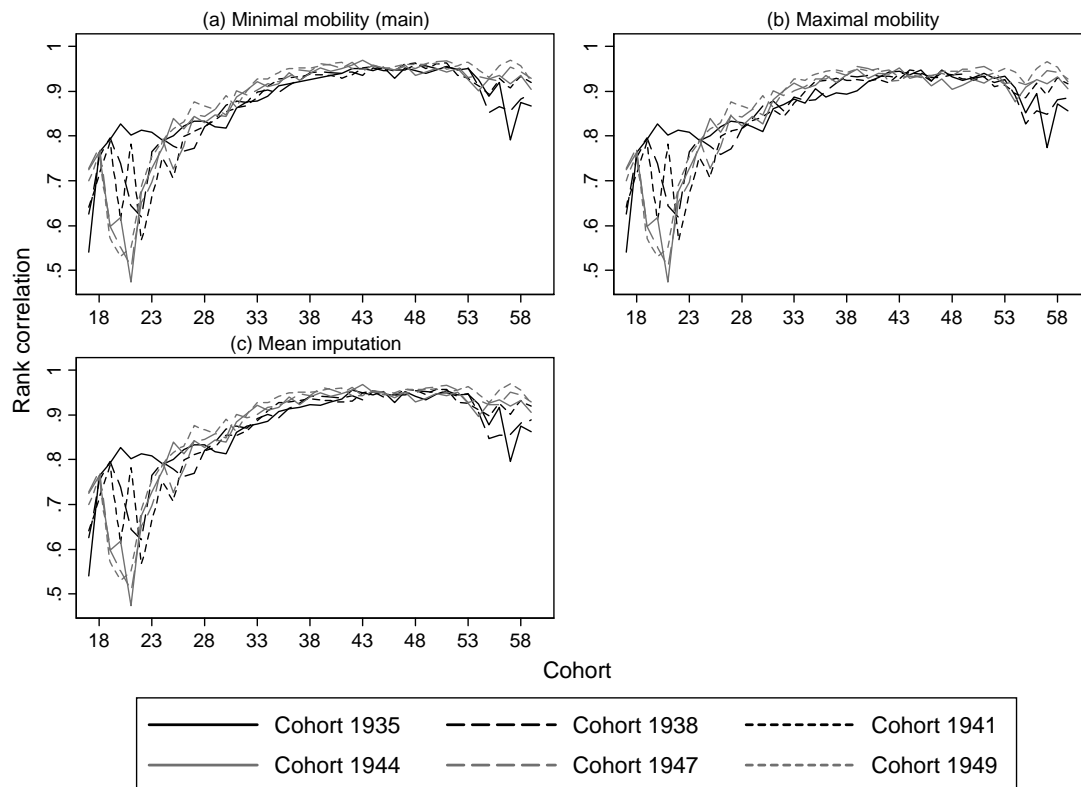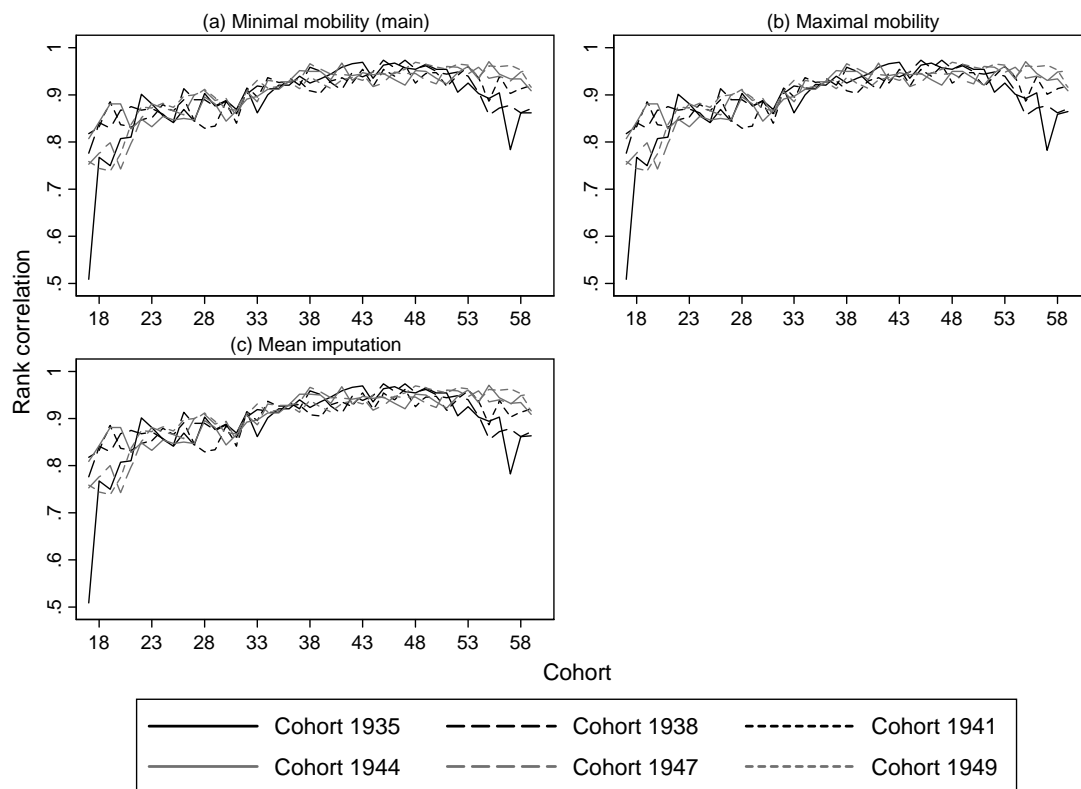Figure III.3: Annual Gini coefficients from age 17 to 60 for cohorts 1935-1949, men

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.4: Annual Gini coefficients from age 17 to 60 for cohorts 1935-1949, women

(a) Original

(b) Imputation

(c) Imputation, market wage

(d) Imputation, break

Cohort

| Cohort 1935 | Cohort 1938 | Cohort 1941 |
| Cohort 1944 | Cohort 1947 | Cohort 1949 |

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.5: Earnings rank correlations between consecutive years for cohorts 1935-1949, men

(a) Original

(b) Imputation

(c) Imputation, market wage

(d) Imputation, break

Cohort

| Cohort 1935 | Cohort 1938 | Cohort 1941 |
| Cohort 1944 | Cohort 1947 | Cohort 1949 |

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

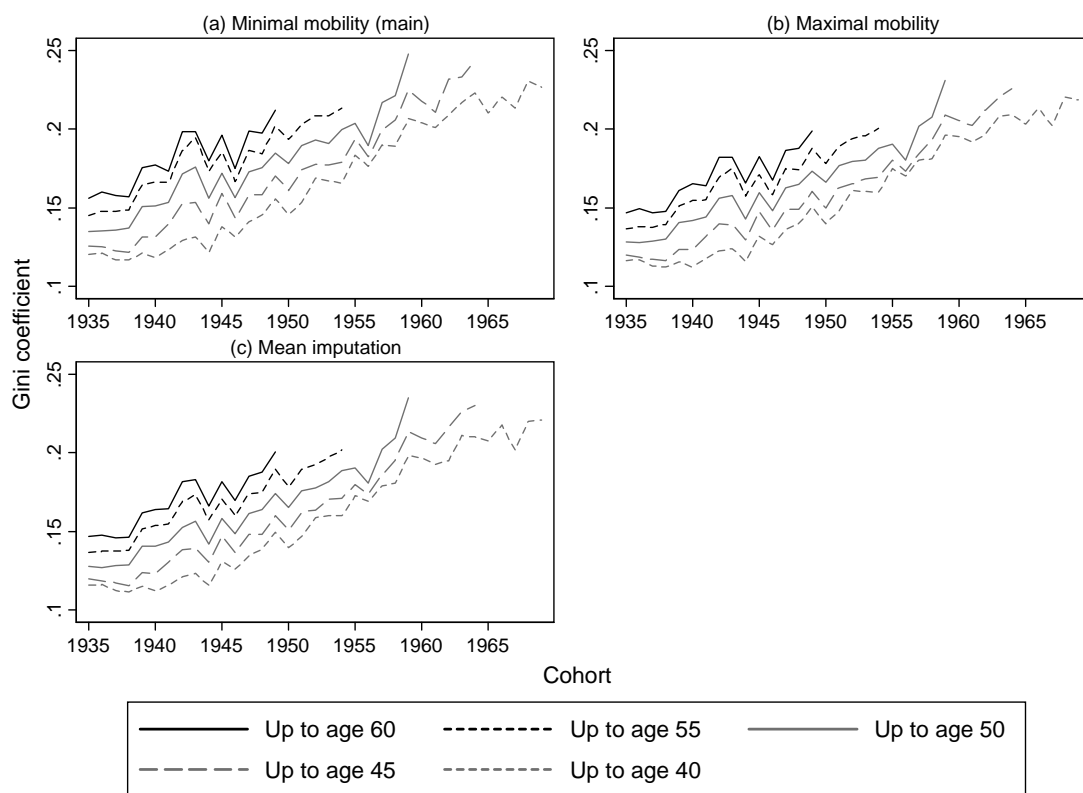Figure III.6: Earnings rank correlations between consecutive years for cohorts 1935-1949, women

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.7: Rank correlation of annual and lifetime earnings for cohorts 1935-1949, men

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

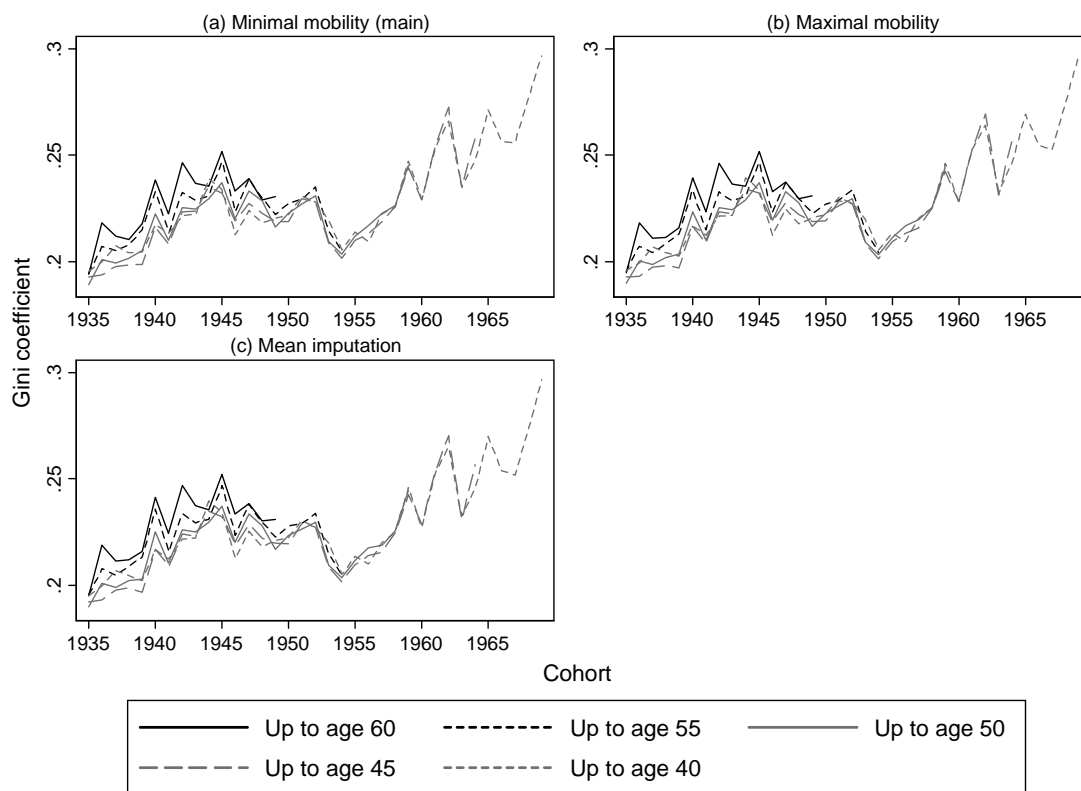Figure III.8: Rank correlation of annual and lifetime earnings for cohorts 1935-1949, women

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

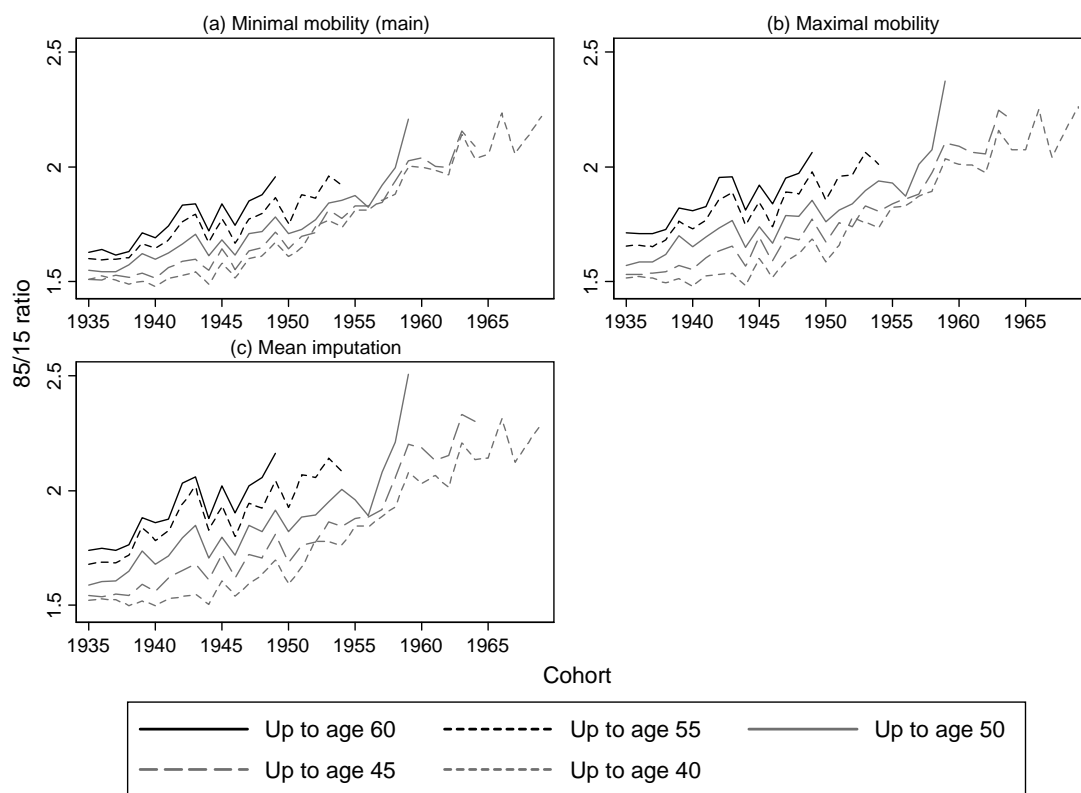Figure III.9: Gini coefficients of UAX for cohorts 1935-1969, men

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.10: Gini coefficients of UAX for cohorts 1935-1969, women

Figure III.11: 85th / 15th ratio of UAX- earnings for cohorts 1935-1969, men

Figure III.12: 85th / 15th ratio of UAX- earnings for cohorts 1935-1969, women

Figure III.13: Rank correlation of UA-40 with selected UAX, men
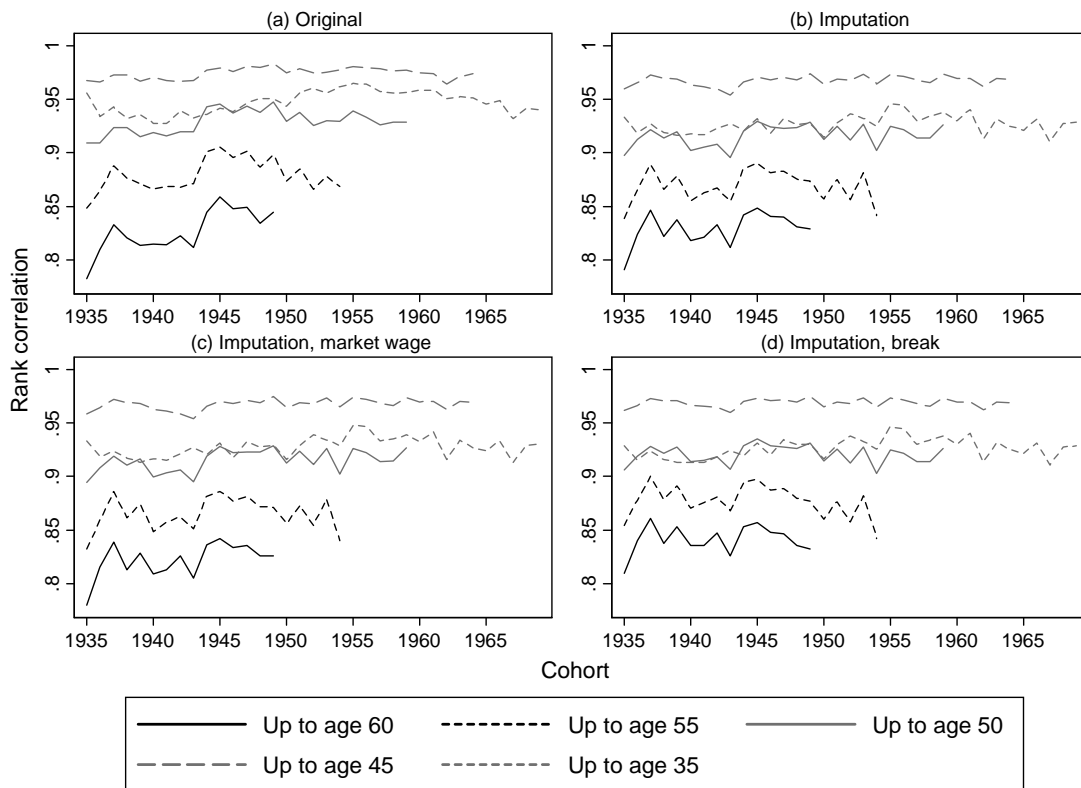
Figure III.14: Rank correlation of UA-40 with selected UAX, women
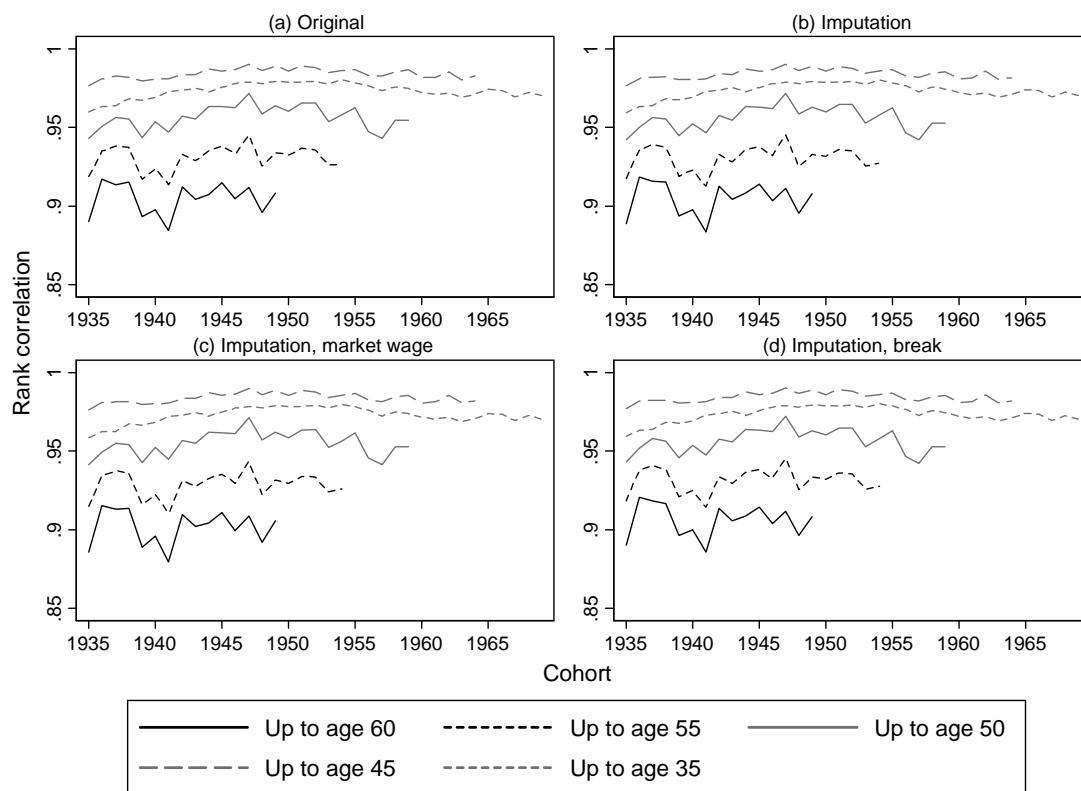
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.15: 50th / 15th ratio of selected UAX for cohorts 1935-1969, men
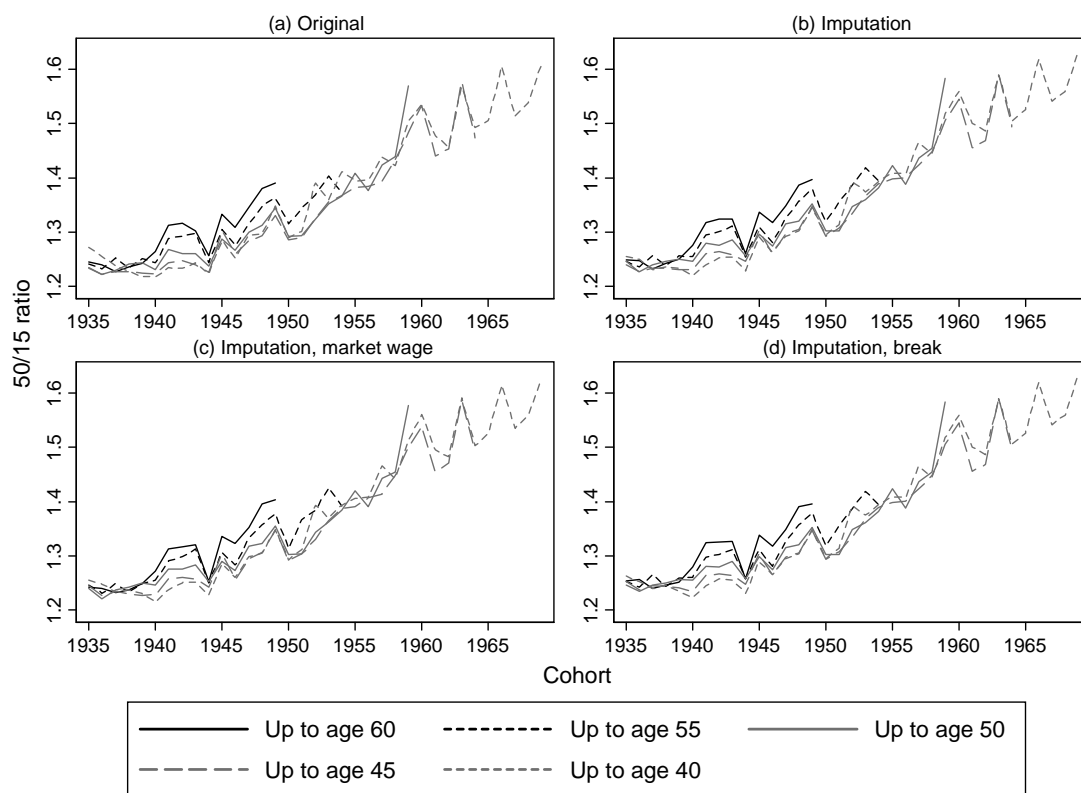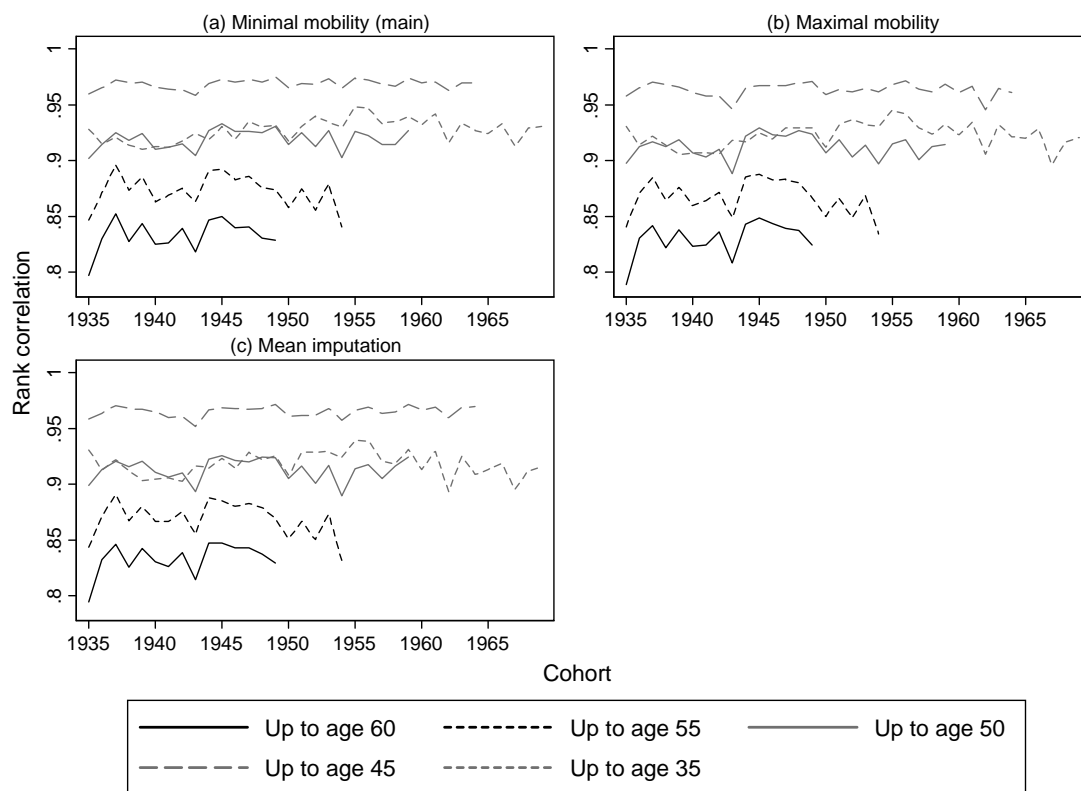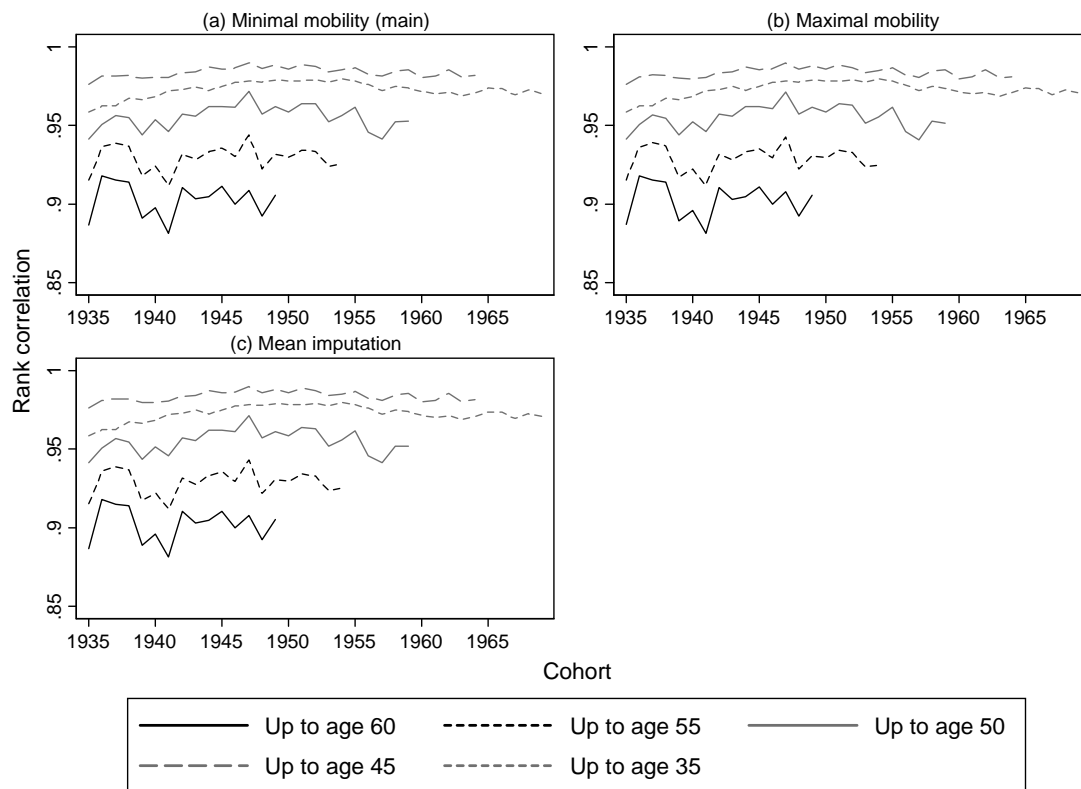

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.16: 50th / 15th ratio of selected UAX for cohorts 1935-1969, women
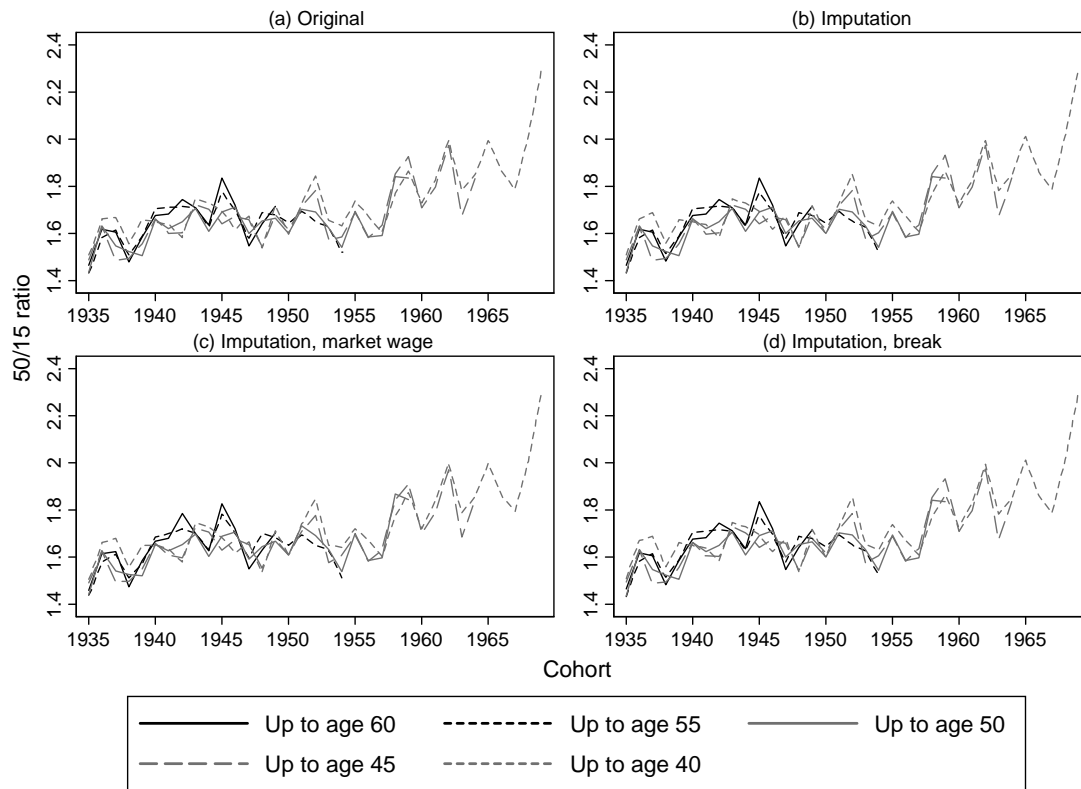
Figure III.17: 85th / 50th ratio of selected UAX for cohorts 1935-1969, men
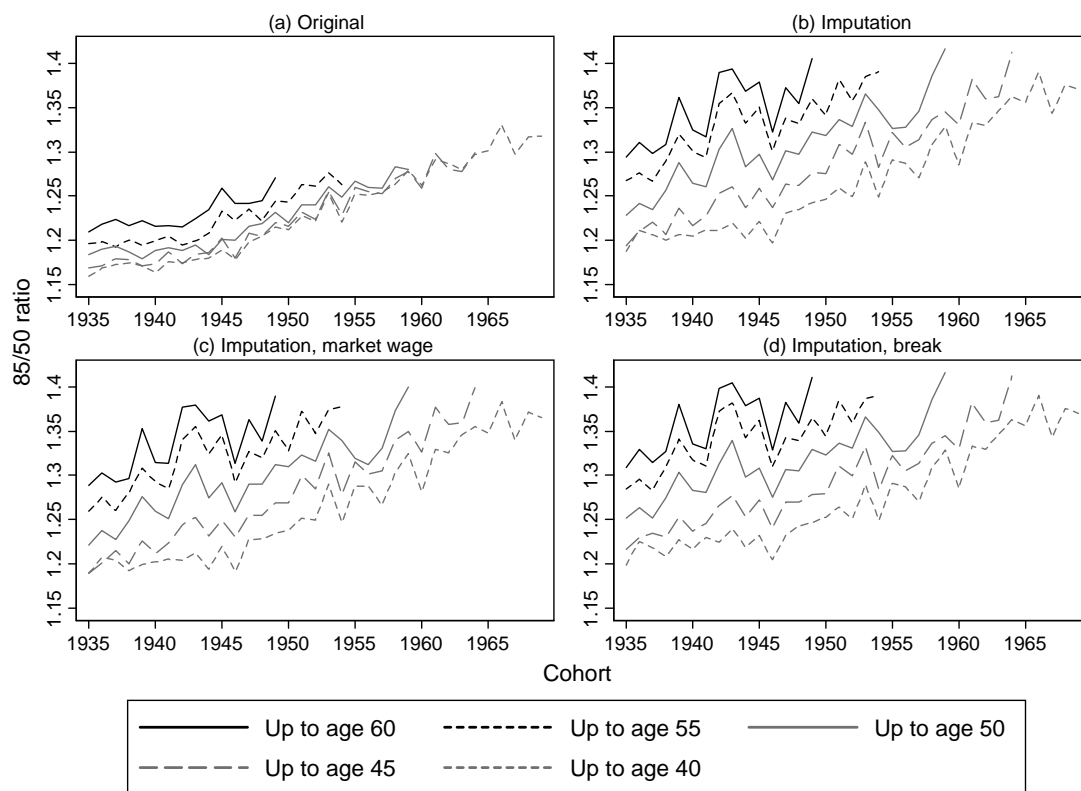
Figure III.18: 85th / 50th ratio of selected UAX for cohorts 1935-1969, women
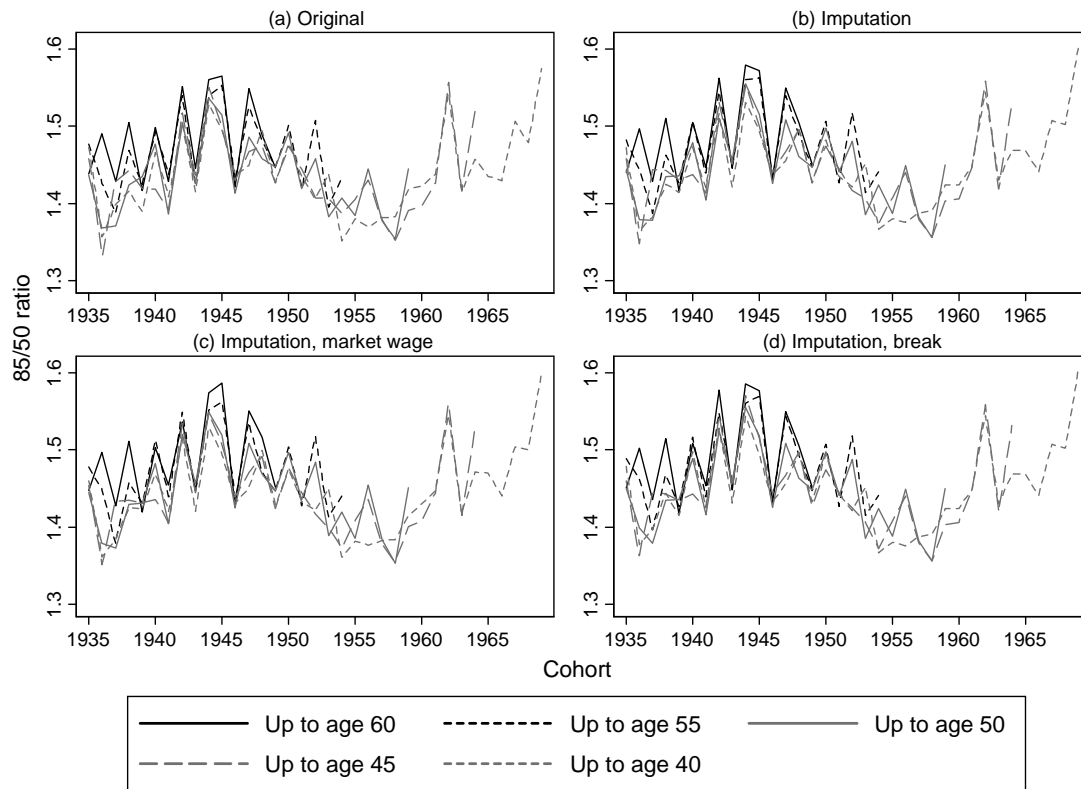
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

III.3 Alternative imputation assumptions

Annotation: The calculations in this section are based on federal bond discounting unless stated otherwise. Each graph shows three different imputation assumptions as described in Online Appendix I:

(a) Minimal mobility depicts our main concept of minimal mobility of the imputed earnings.

(b) Maximal mobility depicts perfect mobility of the imputed earnings.

(c) Mean imputation wage assigns the average imputed wage to everyone above the contribution ceiling.

Figure III.19: Means of annual Gini coefficients and Gini coefficients of lifetime earnings, men

Note: "real" denotes CPI discounting, "federal" denotes federal bond discounting.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.20: Means of annual Gini coefficients and Gini coefficients of lifetime earnings, women
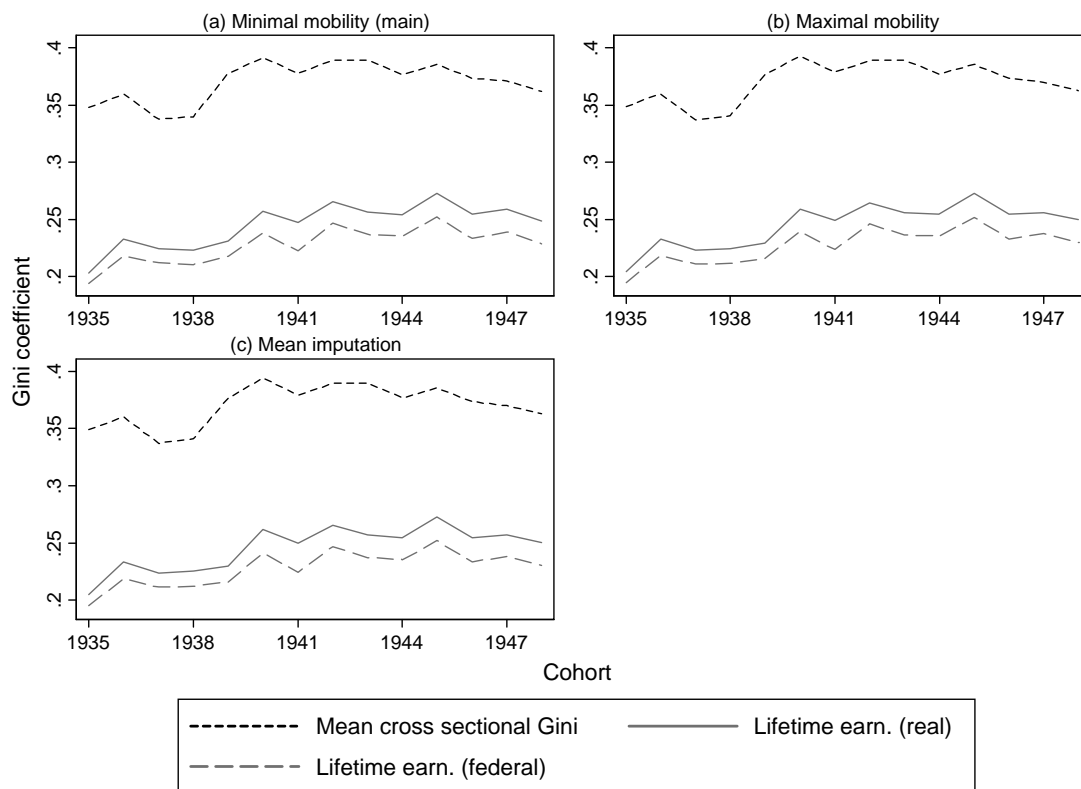
Note: "real" denotes CPI discounting, "federal" denotes federal bond discounting.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.22: Annual Gini coefficients from age 17 to 60 for cohorts 1935-1949, men

Note: "real" denotes CPI discounting, "federal" denotes federal bond discounting.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.23: Annual Gini coefficients from age 17 to 60 for cohorts 1935-1949, women
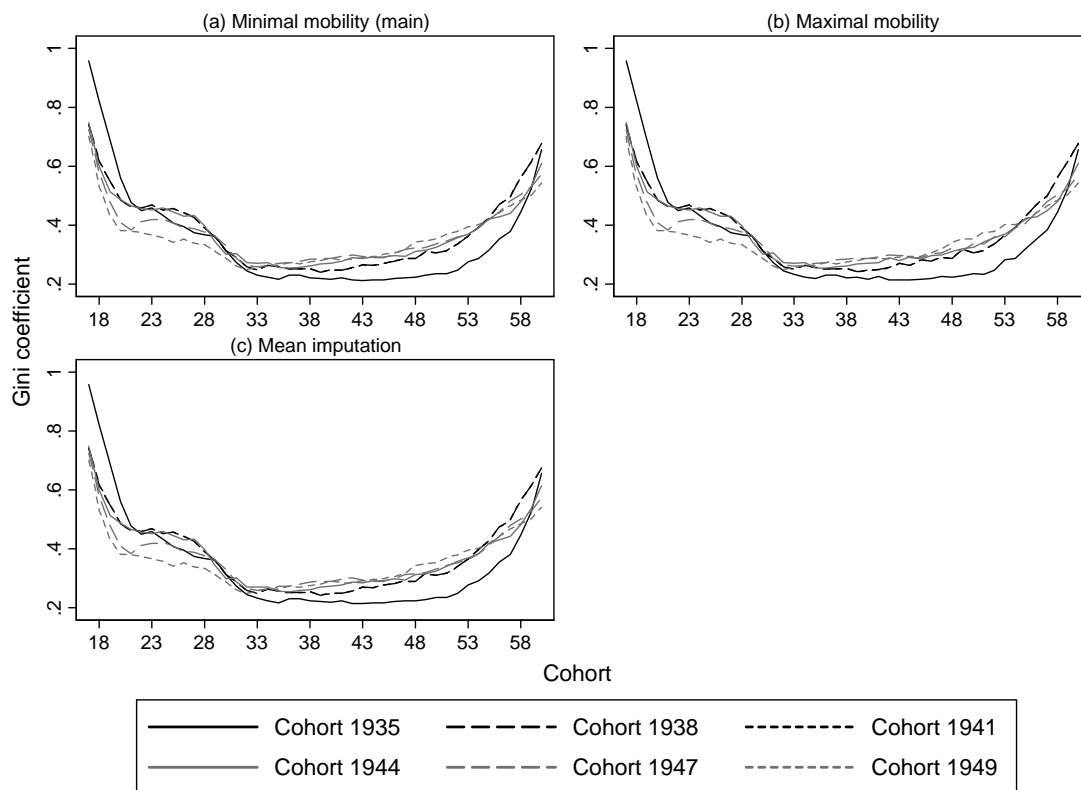
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.24: Earnings rank correlations between consecutive years for cohorts 1935-1949, men

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.25: Earnings rank correlations between consecutive years for cohorts 1935-1949, women

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.26: Rank correlation of annual and lifetime earnings for cohorts 1935-1949, men
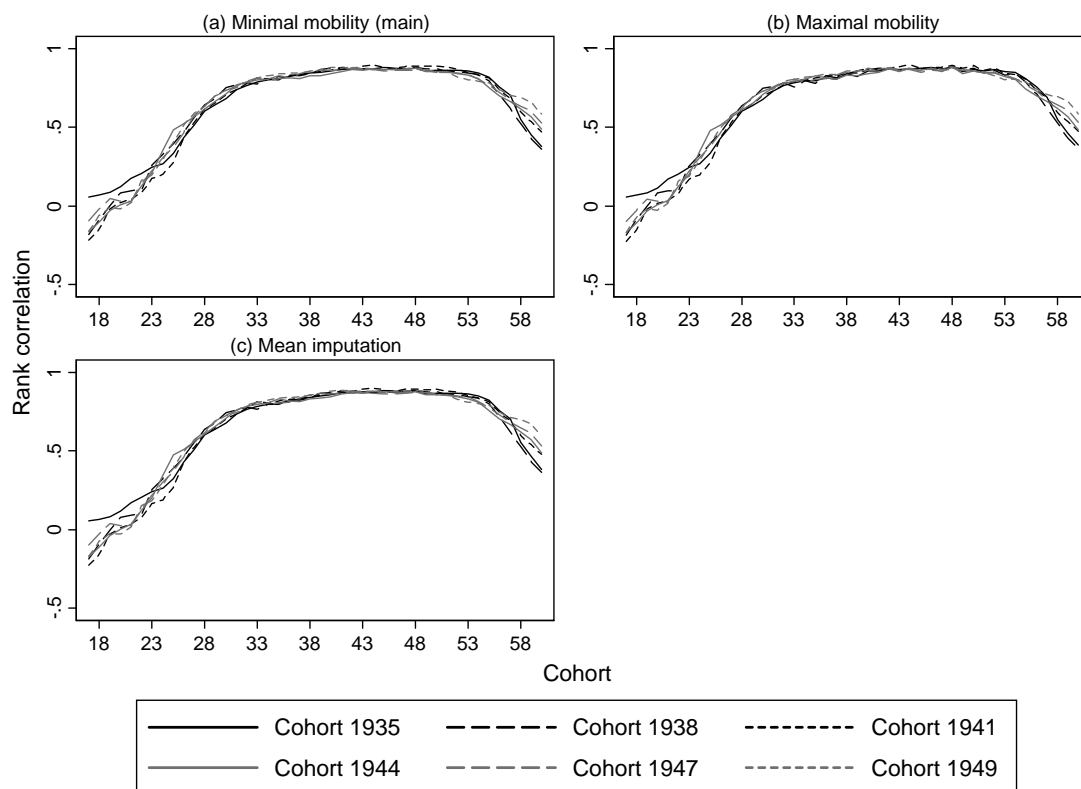
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

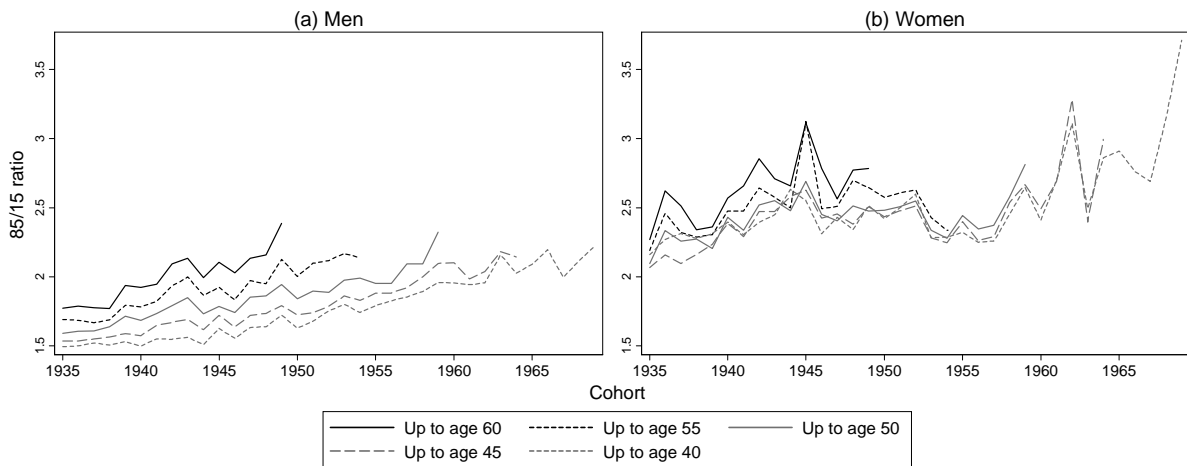Figure III.27: Rank correlation of annual and lifetime earnings for cohorts 1935-1949, women
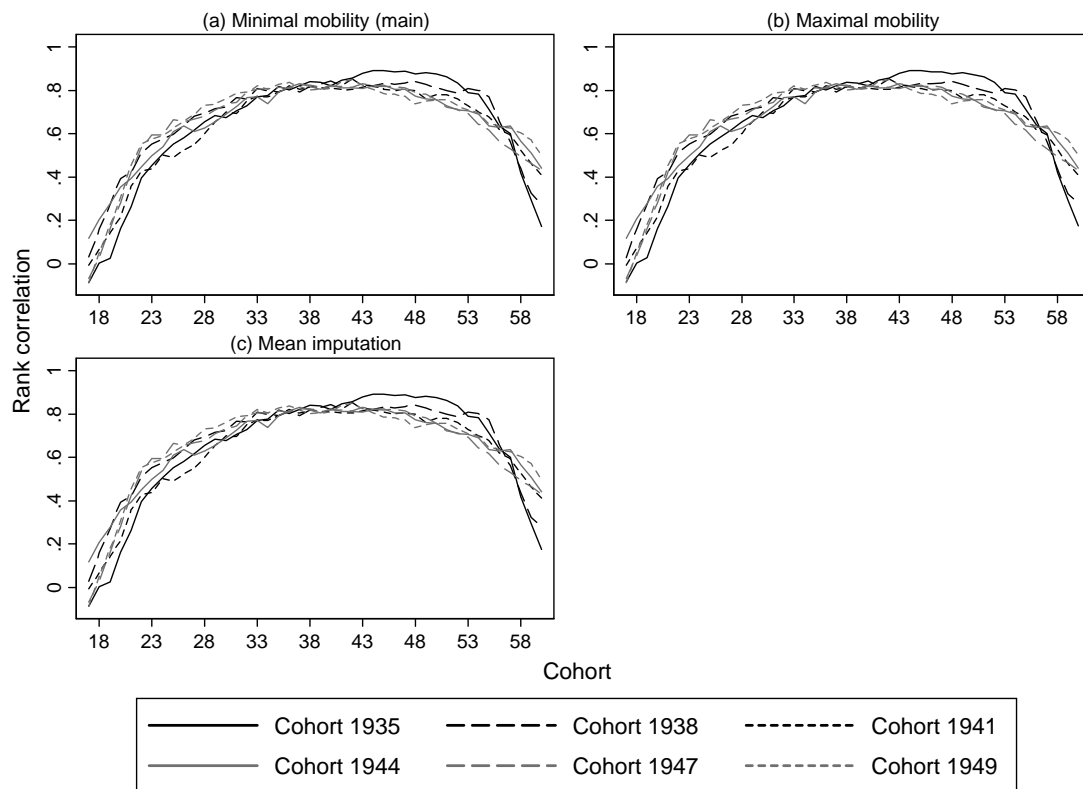
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.28: Gini coefficients of UAX for cohorts 1935-1969, men

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.29: Gini coefficients of UAX for cohorts 1935-1969, women

Figure III.30: 85th / 15th ratio of UAX- earnings for cohorts 1935-1969, men

Figure III.31: 85th / 15th ratio of UAX- earnings for cohorts 1935-1969, women.
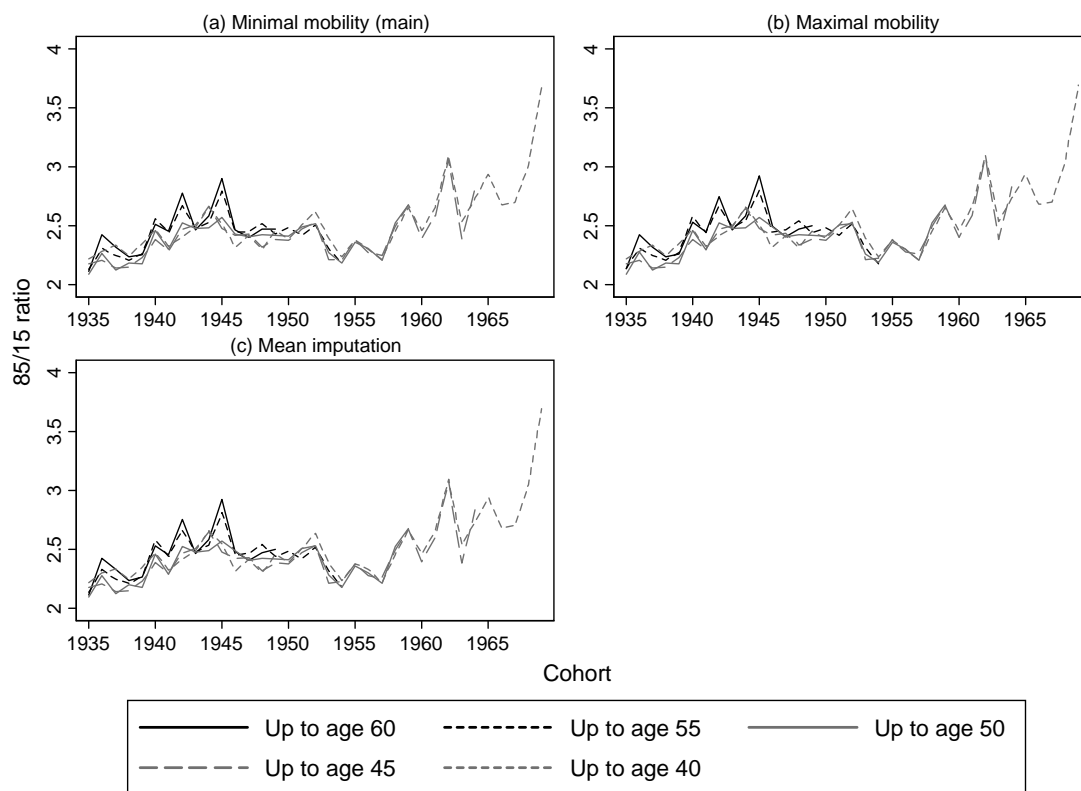
Figure III.32: Rank correlation of UA-40 with selected UAX, men

Figure III.33: Rank correlation of UA-40 with selected UAX, women

Figure III.34: 50th / 15th ratio of selected UAX for cohorts 1935-1969, men
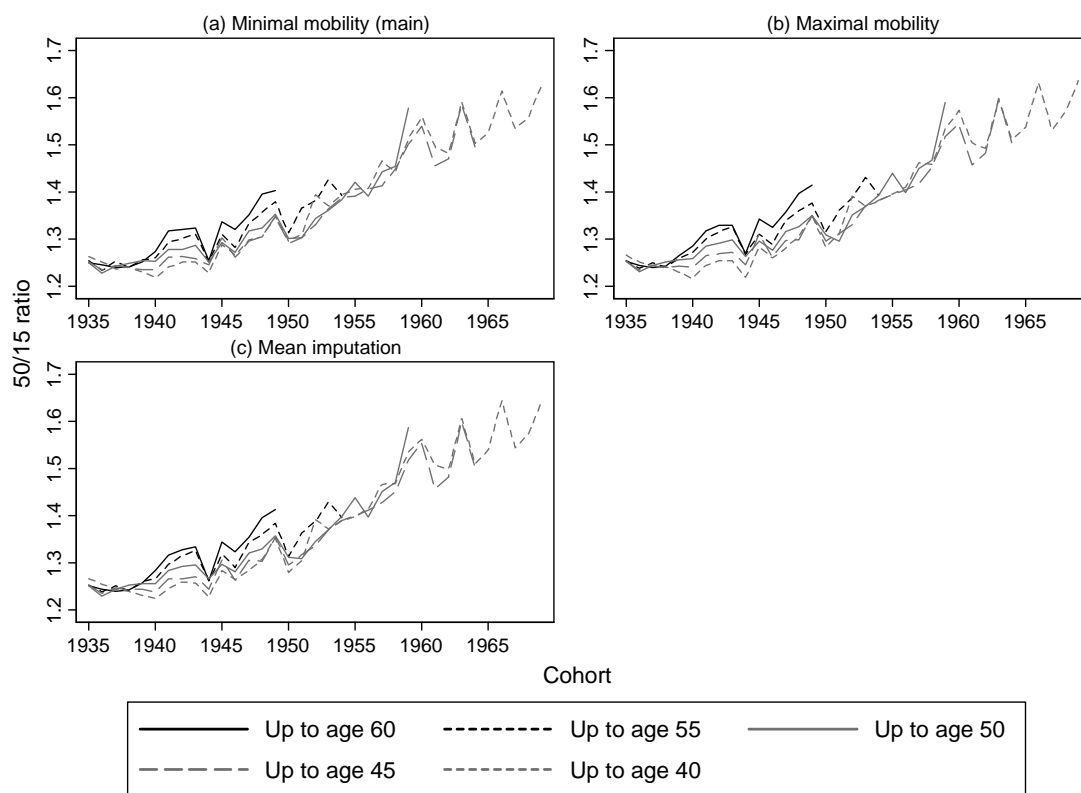
Figure III.35: 50th / 15th ratio of selected UAX for cohorts 1935-1969, women
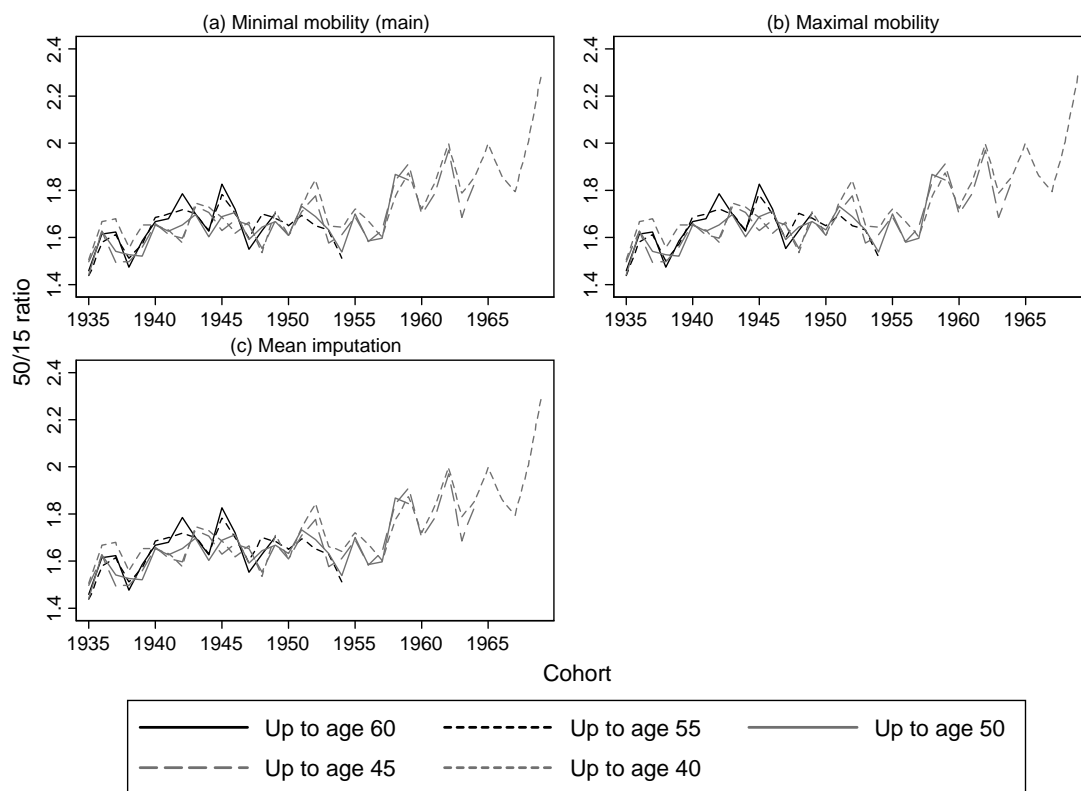
Figure III.36: 85th / 50th ratio of selected UAX for cohorts 1935-1969, men
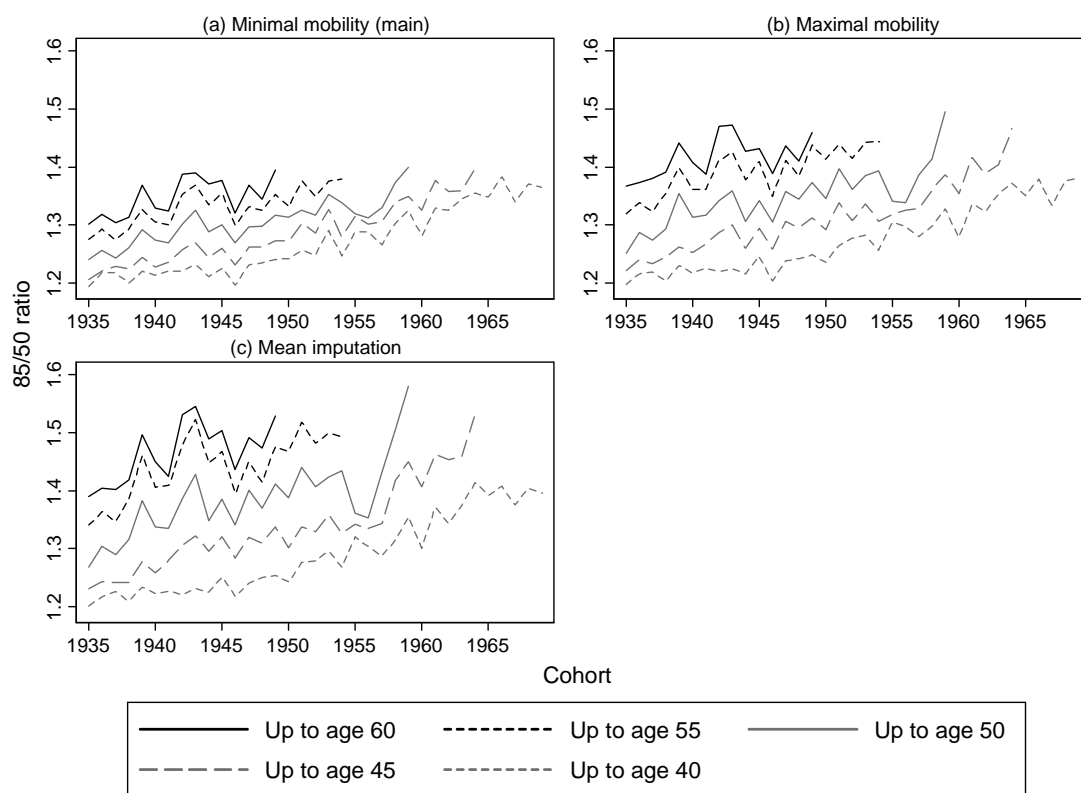
Figure III.37: 85th / 50th ratio of selected UAX for cohorts 1935-1969, women
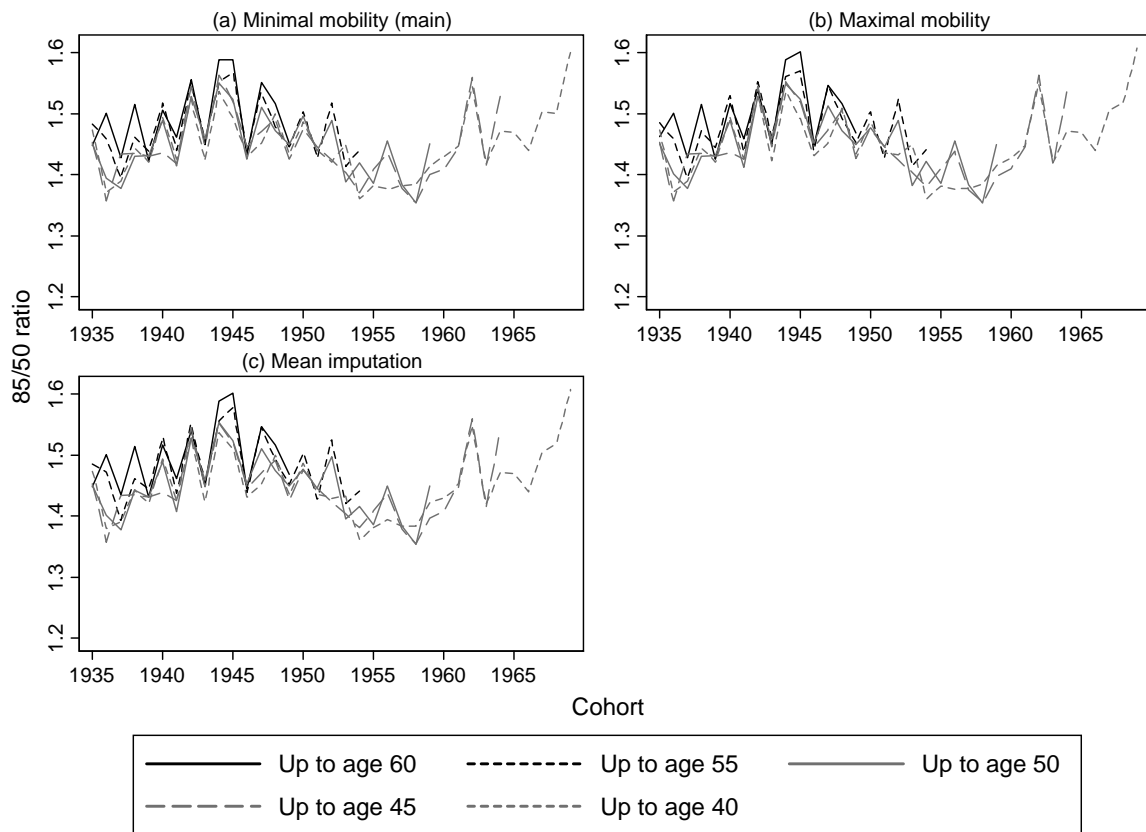
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

III.4 Alternative discounting method − real earnings

Annotation: Each graph in this section uses real earnings instead of federal bond discounted earnings. The section mimics the relevant graphs in the paper.

Figure III.38: Rank correlation of annual and lifetime earnings with for cohorts 1935-1949, men and women
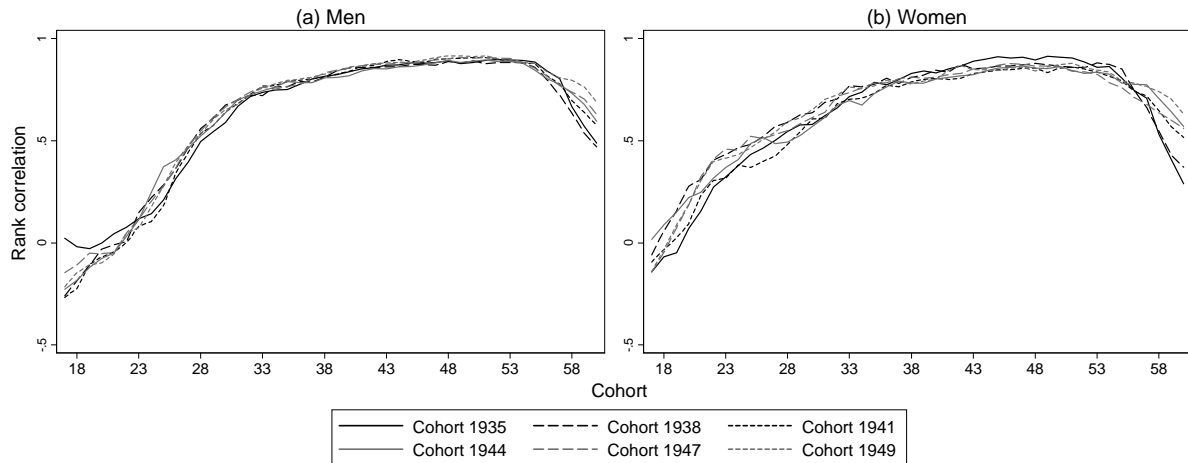
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.39: Gini coefficients of UAX for cohorts 1935-1969, men and women
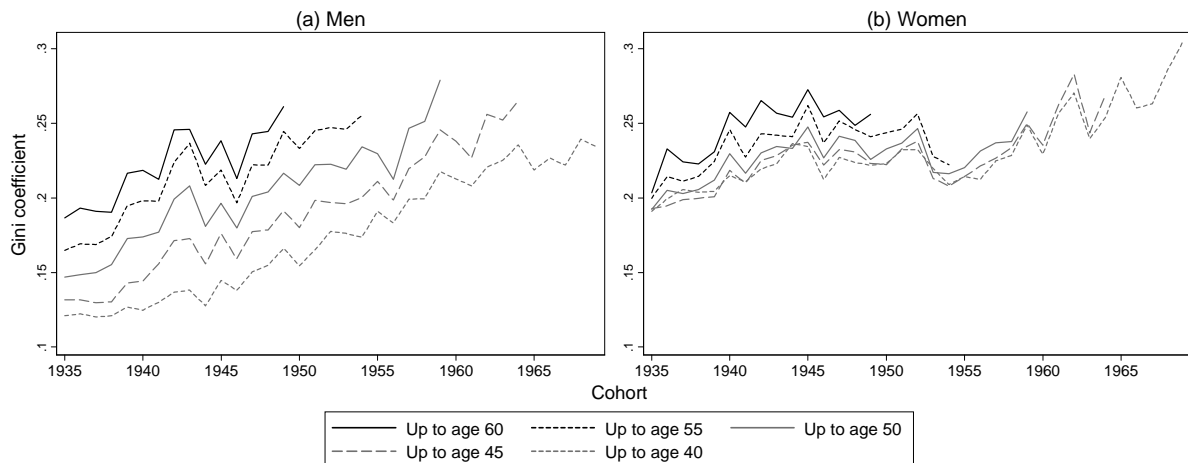
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.40: 85th / 15th ratio of UAX- earnings for cohorts 1935-1969, men and women

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.41: Rank correlation of UA-40 with selected UAX, men and women
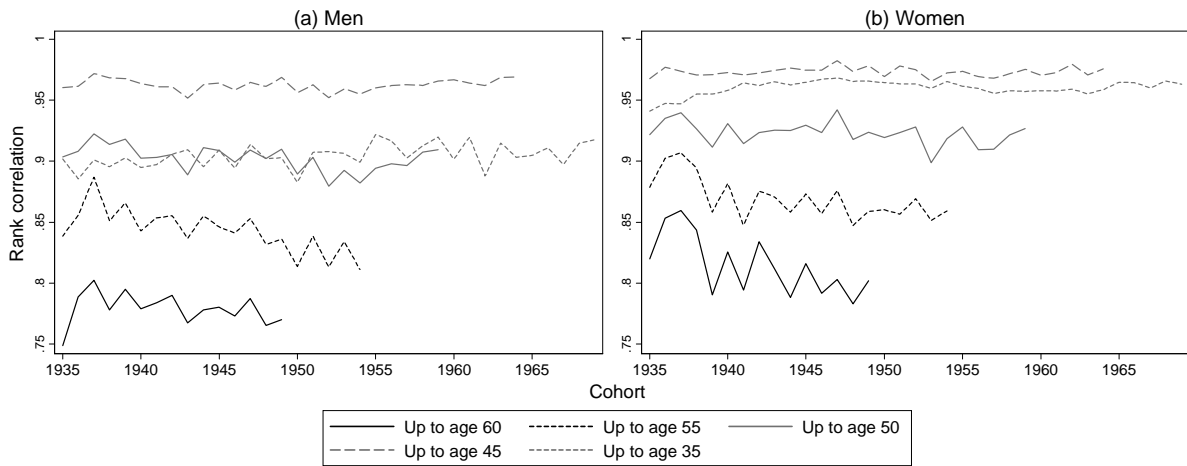
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.42: 50th / 15th and 85th / 50th ratio of selected UAX, federal bond discounting, men cohorts 1935-1969.
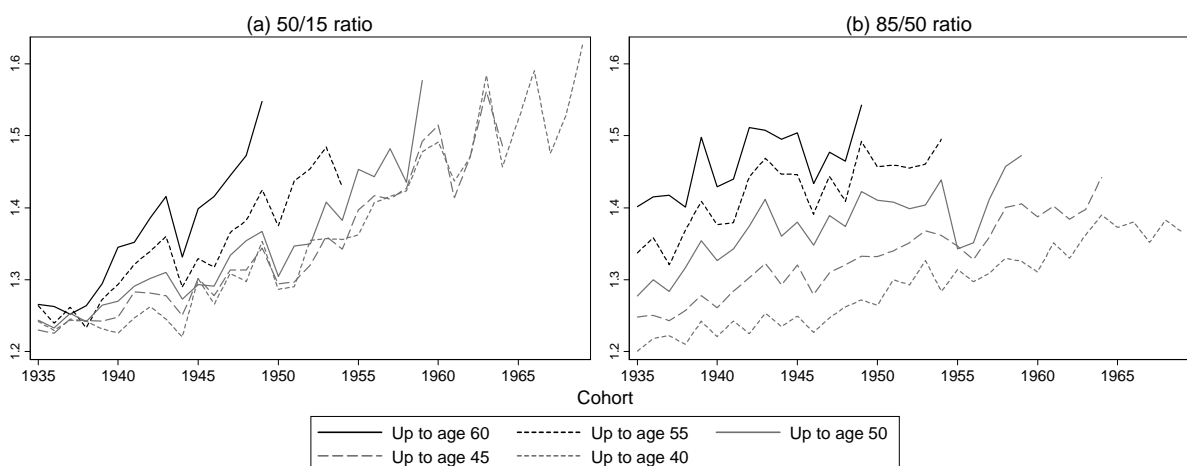
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.43: Gini coefficients of UAX for cohorts 1935-1969 with earnings imputation if individual is not employed, men
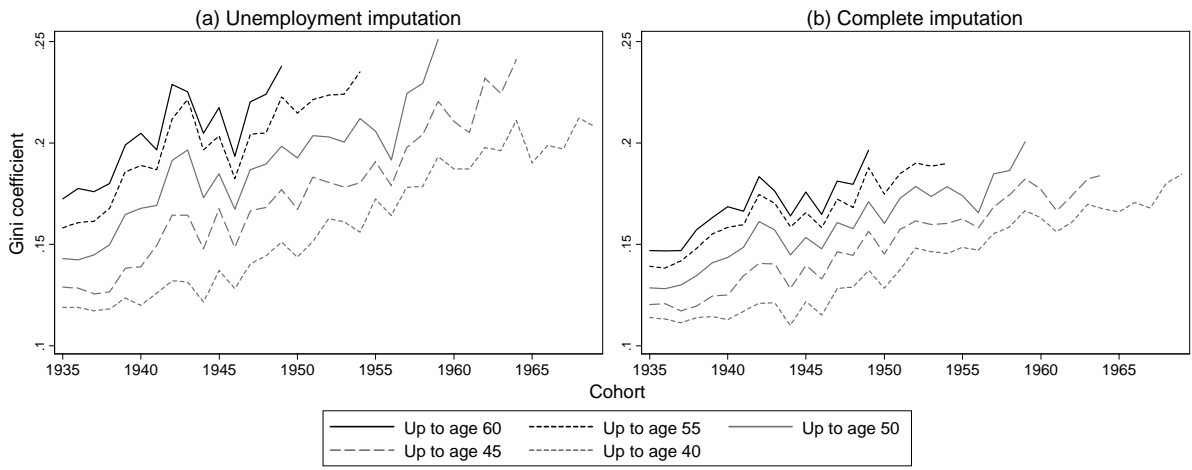
(a) Unemployment imputation

(b) Complete imputation

Gini coefficient

Cohort

| Up to age 60 | Up to age 55 | Up to age 50 |
| Up to age 45 | Up to age 40 | |

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

46

III.5 Generalized entropy measures and further results from the Gini-decomposition

Annotation: This sections first depicts the UAX-earnings results for the GE[0], the GE[1] and the GE[2]. Then it shows further results from Gini-decompositions. All graphs in this section are based on federal bond discounting.

Figure III.44: Generalized entropy measures of UAX for cohorts 1935-1969, men
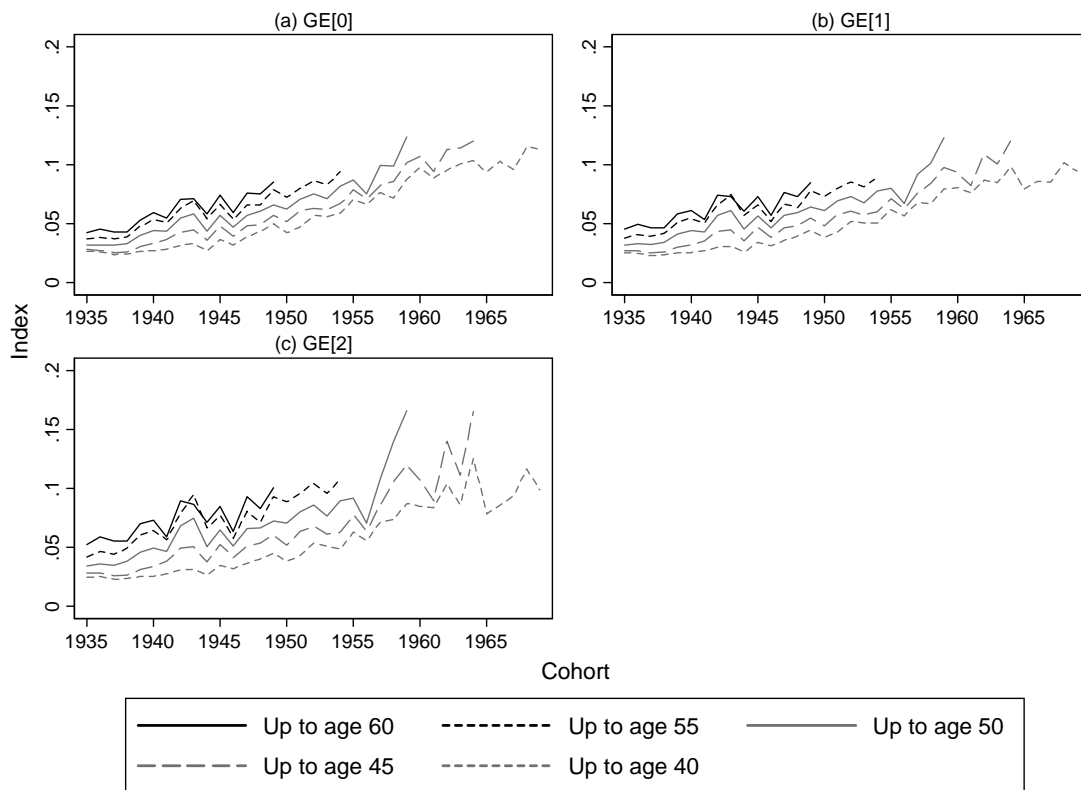
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.45: Generalized entropy measures of UAX for cohorts 1935-1969, women
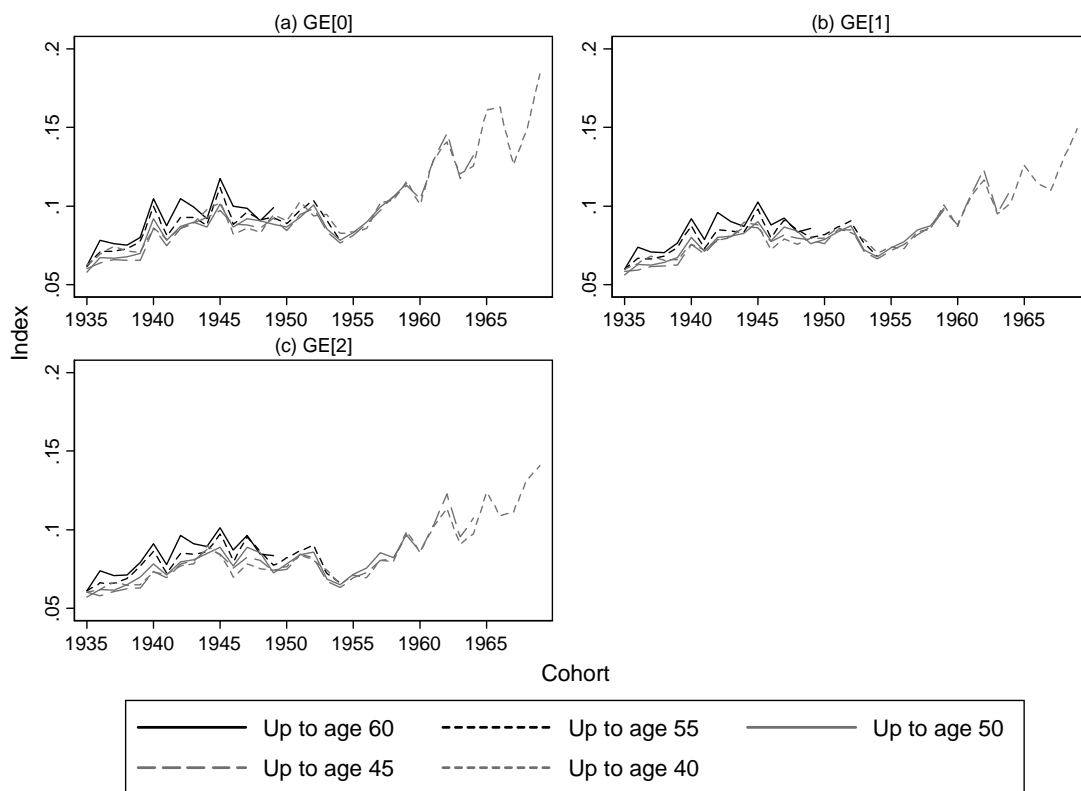
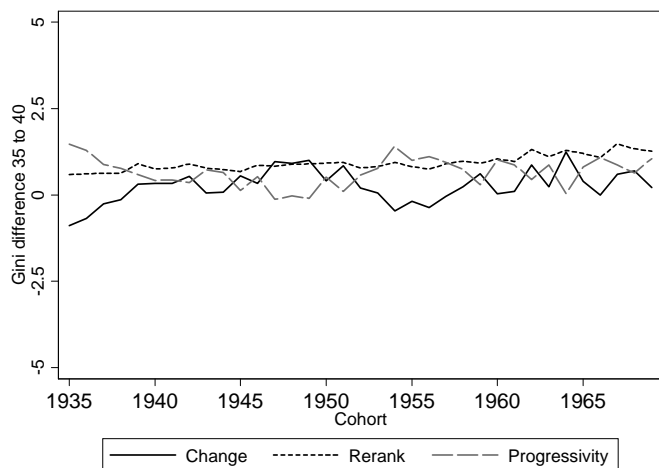Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.46: Decomposition of changes in inequality as of Eq. (3) in the paper for selected cohorts, men
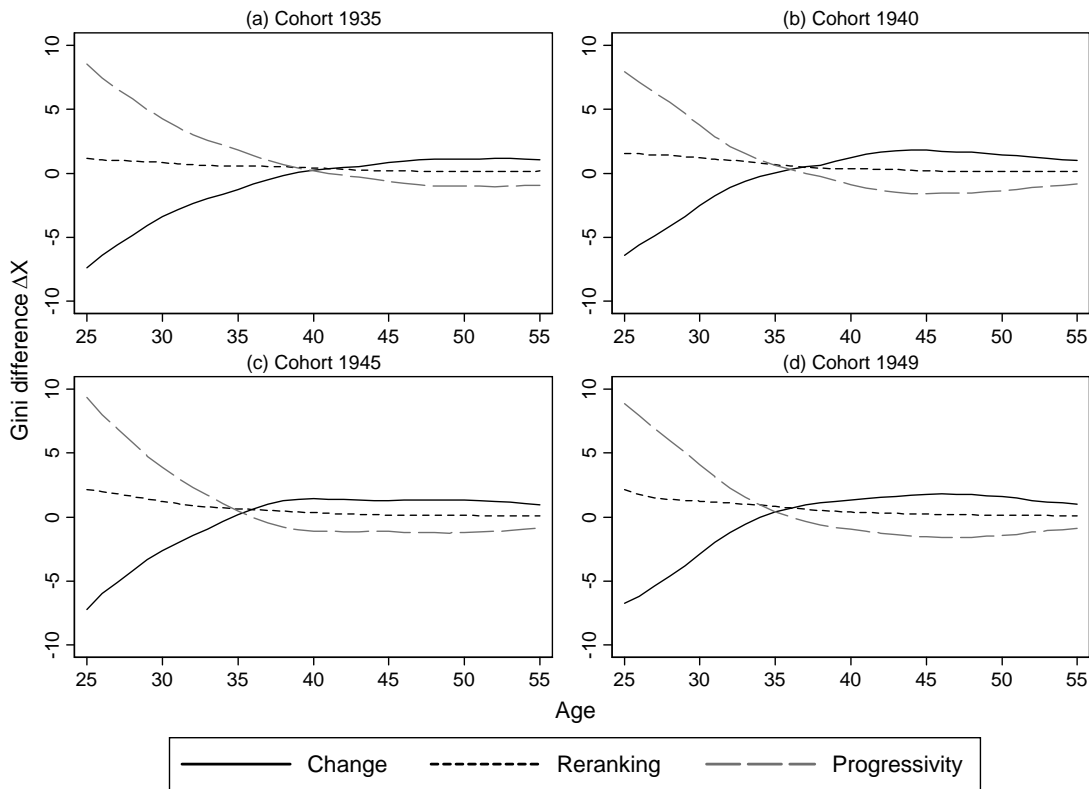
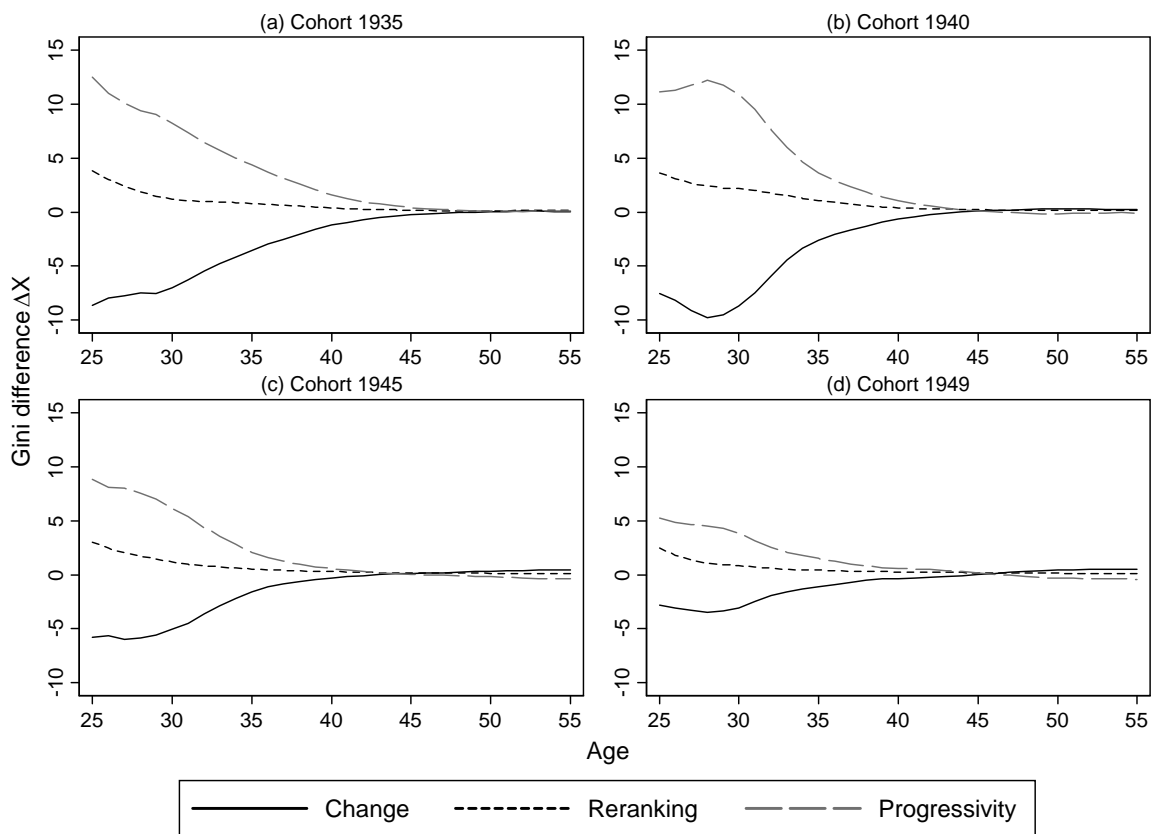Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.47: Decomposition of changes in inequality as of Eq. (3) in the paper for changes of lifetime earnings from age 35 to age 40 for cohorts 1935-1969, men

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.48: Decomposition of changes in inequality as of Eq. (3) in the paper for selected cohorts, women
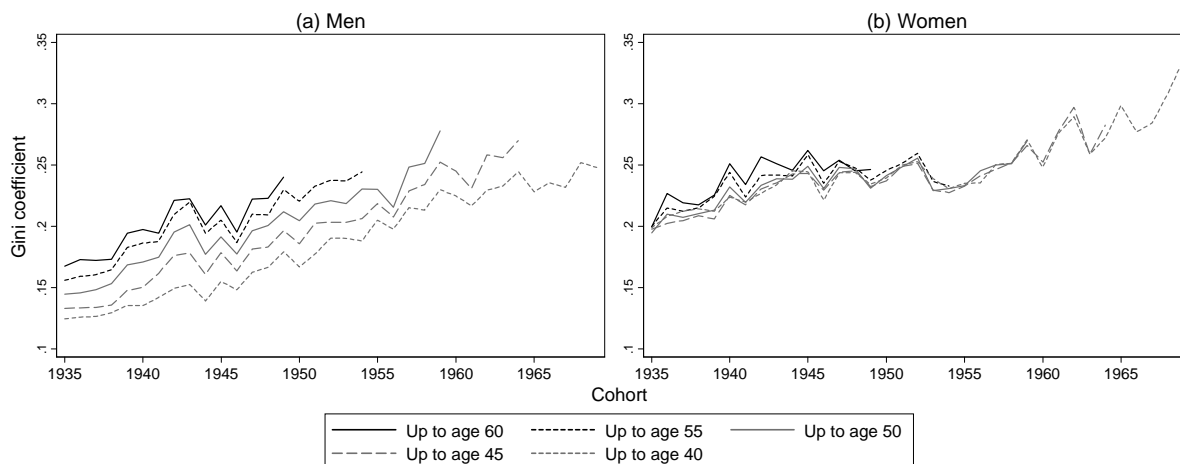

Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.
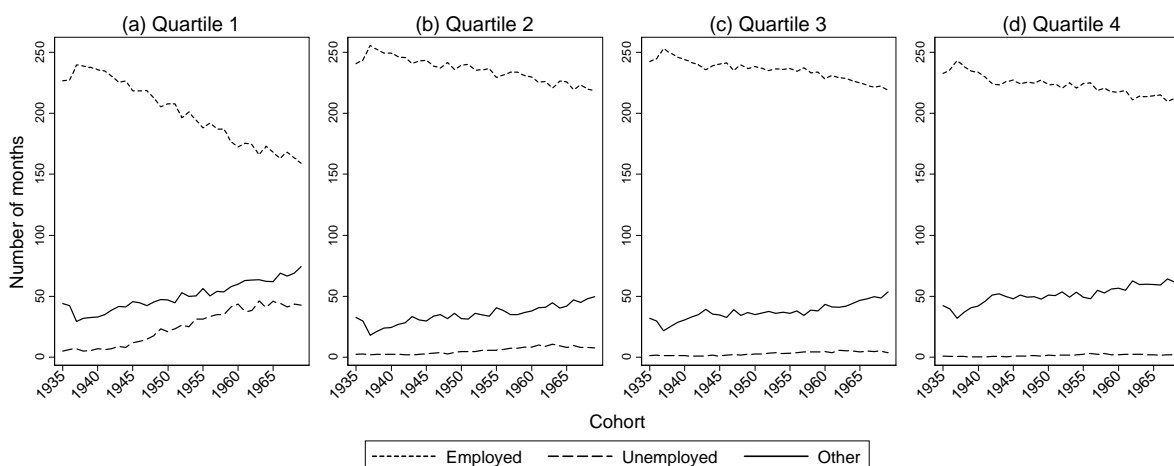
III.6 NPV after 25

All graphs in this section are based on federal bond discounting.

Figure III.49: Gini coefficients of UAX for cohorts 1935-1969, men and women

Note: The NPV is based on annual earnings from age 25 to age *X* instead of age 17 to age *X*.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.50: Months of employment status up to age forty by quartile of UA-40 for cohorts 1935-1969, men


Note: Earnings quartiles based on UA-40 with federal bond discounting. The NPV is based on annual earnings from age 25 to age *X* instead of age 17 to age *X*.
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

III.7 Marginal employment "minijobs"

All graphs in this section are based on federal bond discounting.

Figure III.50: Gini coefficients of UAX with earnings from marginal employment set to zero for cohorts 1935-1969, men and women
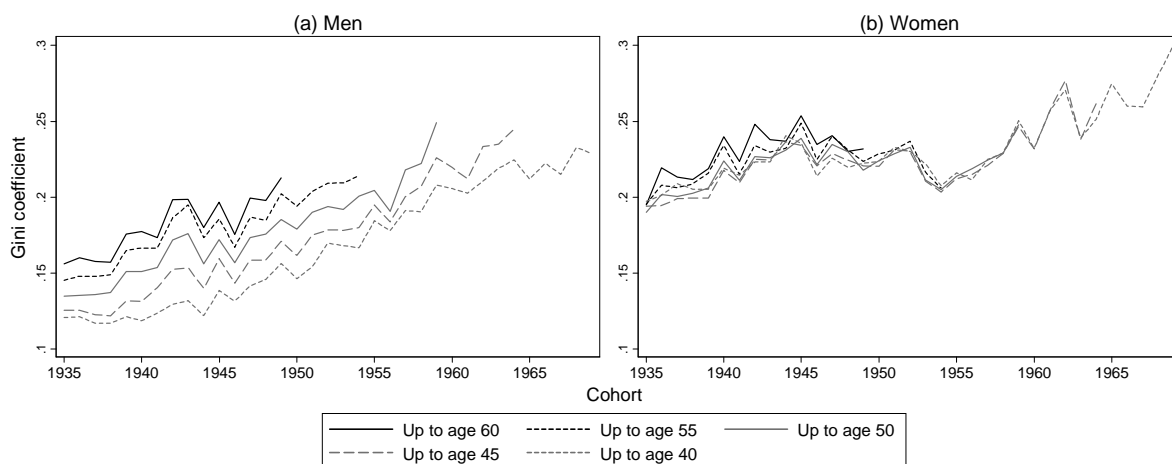
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure III.51: Gini coefficients of UAX with earnings from marginal employment set to missing for cohorts 1935-1969, men and women
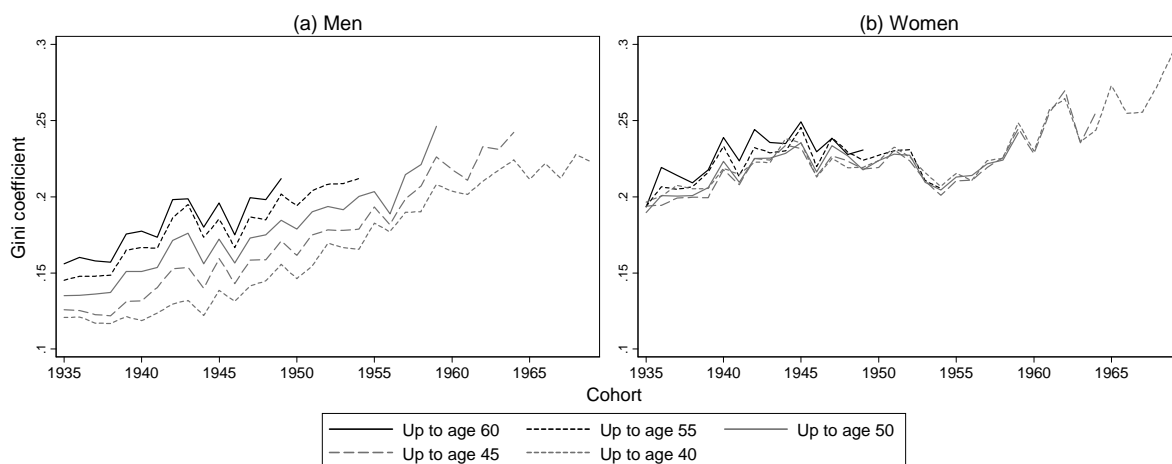
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Appendix IV: Decomposing the rise of inequality

We consider two cohorts, "parents" and "children", and the distribution of their UAX at the same X, e.g. when the cohorts are forty-five. Let a denote the actually measured Gini coefficient of the UAX-distribution of parents and let A denote that coefficient for the children. The difference

$$A - a \qquad (\text{IV.1})$$

is to be decomposed into two parts: the increase of inequality due to the rise of unemployment and the increase due to changes in the wage structure.

Let's denote by h the hypothetical Gini coefficient of the UAX-distribution of parents had they lived under full employment. Empirically, we obtain h by imputing earnings to parents for the few months when they were unemployed. The ratio of a to h captures the inequality increase due to "quasi-full-employment" instead of full employment.

In order to gauge the effect on the inequality increase $A - a$ due to the rise of unemployment, we have to estimate the hypothetical Gini coefficient of the UAX-distribution of children in case they had lived under the same situation of "quasi-full-employment" as their parents. We do this by computing the hypothetical Gini coefficient of the UAX-distribution of children had they lived under full employment and by assuming that the inequality-increasing effect of having "quasi-full-employment" rather than full employment is symmetric to the effect we found for the parents' cohort.

Let's denote by H the hypothetical Gini coefficient of the UAX-distribution of children had they lived under full employment. Empirically, we obtain it by imputing earnings to children for the months they were unemployed. The hypothetical Gini coefficient of the UAX-distribution had they lived under the same "quasi-full-employment" as their parents is:

$$(a/h) \cdot H. \qquad (\text{IV.2})$$

The difference $(a/h)H - a$ is thus the increase in UAX inequality between the two cohorts that we attribute to changes in the wage structure while $A - (a/h)H$ is the increase we attribute to the rise of unemployment.

In the main text of the paper we refer to the shares of the inequality increase that can be attributed to the two factors. The share due to changes in the wage structure is:

$$\frac{(a/h)H - a}{A - a} = \frac{(H - h)/h}{(A - a)/a}. \qquad (\text{IV.3})$$

This is the formula we use in the main text of our paper: growth rate of the Gini in the hypothetical situation of full employment divided by the growth rate of the actually measured Gini.

Figure IV.1: Effect of the imputation of earnings for times of unemployment on total inequality for cohorts 1935-1969, men
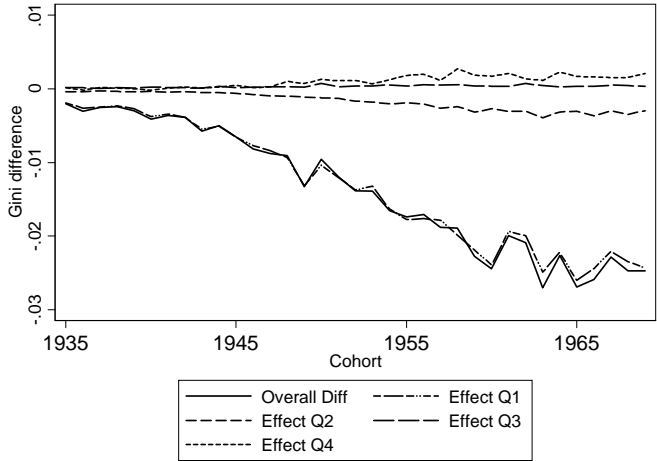
Note: Earnings quartiles based on up-to-age 40 earnings with federal bond discounting without imputation of earnings for times of unemployment. Gini difference denotes the difference compared to the Gini coefficient of UA-40 without imputation of times of unemployment if times of unemployment are imputed across all quartiles ("Overall Diff") or for each quartile separately while leaving the other quartiles unchanged ("Effect Q1", "Effect Q2", "Effect Q3", "Effect Q4").
Source: FDZ-RV − VSKT2002, 2004-2009_Bönke, own calculations using weighted data.

Figure VI.1 decomposes the effect on the Gini of UA-40 if earnings are imputed for times of unemployment. To measure the influence of the imputation on total inequality reduction by quartile, first the total effect of the imputation is calculated (also see Figure 13 in Section 5) and the difference to the actually observed UA-40 Gini (solid black line labelled "Overall Diff") is computed. Then, earnings are imputed for unemployment spells for each earnings quartile separately while leaving the UA-40 in the other three quartiles unaltered. Thereby, earnings quartiles are based on the original UA-40 distribution. The effect on overall inequality for imputing in the first quartile only is labelled "Effect Q1" and so forth for quartile two, three and four in the above figure. This exercise reveals that almost the whole inequality reduction stems from the imputation of unemployment spells in the first quartile.