DISCUSSION PAPER SERIES

IZA DP No. 5062

Understanding Rising Income Inequality in Germany

Martin Biewen
Andos Juhasz

July 2010

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Understanding Rising Income Inequality in Germany

Martin Biewen
University of Tübingen,
DIW Berlin and IZA

Andos Juhasz
University of Tübingen

Discussion Paper No. 5062
July 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 5062
July 2010

ABSTRACT

Understanding Rising Income Inequality in Germany*

We examine the causes for rising income inequality in Europe's most populous economy. From 2000 to 2006, Germany experienced an unprecedented rise in net equivalized income inequality and poverty. At the same time, unemployment rose to record levels and there was evidence for a widening distribution of labour market returns, as well as that of other market incomes. Other factors that possibly contributed to the rise in income inequality were changes in the tax system, changes in the household structure (in particular the rising share of single parent households), and changes in other socio-economic characteristics (e.g. age or education). We address the question of which factors were the main drivers of the observed inequality increase. Our results suggest that most of the increase can be explained by both changes in employment outcomes and in market returns, and, to a similar extent, by changes in the tax system. Changes in household structures and other household characteristics seem to have played a much smaller role. Put into an international perspective, our results suggest that rising income inequality in non-Anglo-Saxon countries is the likely result of *both* increasing inequality in market returns *and* increasing inequality in employment outcomes, as well as of idiosyncratic changes such as tax reforms.

JEL Classification: D31, C14, I30

Keywords: income inequality, poverty, unemployment, kernel density estimation

Corresponding author:

Martin Biewen
Department of Economics
University of Tübingen
Mohlstr. 36
72074 Tübingen
Germany
E-mail: martin.biewen@uni-tuebingen.de

* We are grateful to seminar participants in Rauischholzhausen and at the 8th International German Socio-Economic Panel User Conference for helpful comments and discussions. All errors are our own. The data used in this paper were made available by the German Socio-Economic Panel Study (GSOEP) at the German Institute for Economic Research (DIW), Berlin.

1 Introduction

There has been a clear trend of increasing income inequality in industrialized countries over the past three decades, although with differences in the timing and intensities across countries (see OECD (2008)). This trend was first observed in Anglo-Saxon countries such as the United States, where pronounced changes in the wage and earnings distribution in the 1980s and 1990s sparked a large body of literature examining the possible causes of increasing inequalities in labour market returns (see e.g. Bound/Johnson (1992), Levy/Murnane (1992), Murphy/Welch (1992), Juhn et al. (1993), and DiNardo et al. (1996)). The fact that many of these changes could not be observed to the same extent in most of the less flexible European labour markets, led Krugman (1994) to formulate the hypothesis, that the dramatic increase in wage inequality in the United States - which was generally seen as the result of skill-biased technological progress - and growing unemployment in the less flexible European labour markets, are 'two sides of the same coin'.[2] An interesting implication of Krugman's hypothesis is that both rising wage inequality in Anglo-Saxon countries and rising unemployment in continental European countries have similar effects on the distribution of overall personal incomes. Rising wage inequality directly translates into rising inequality in personal incomes because wages are a major part of overall income, and rising unemployment increases personal income inequality if the former wage incomes of the unemployed are only imperfectly replaced by unemployment benefits.

Germany is one of the countries whose wage and income distribution seemed remarkably stable for a long time.[3] While this view has recently been questioned for the distribution of wages,[4] there is a consensus that the pronounced changes in the structure of wages that were observed in other countries reached Germany with considerably delay. In Germany, wage inequality started to grow in a clear way from the mid-1990s onwards, although the changes were less drastic than those observed in countries such as the United States (see Kohn (2006), Genandt/Pfeiffer (2007), Dustmann et al. (2009), Fuchs-Schündeln et al. (2010), Antonczyk et al. (2010a, 2010b)). The distribution of overall incomes remained quite stable until the end of the 1990s, but witnessed a sharp increase in inequality and poverty beginning in 2000. This increase was accompanied by a steep increase in unemployment which, along with the changes in the wage distribution, probably

[2]See e.g. Puhani (2008) for direct tests of the Krugman hypothesis.

[3]See Steiner/Wagner (1998), Biewen (2000), Prasad (2004). Strictly speaking, this applies only to West Germany. The East German income distribution changed considerably following the reunification of the country.

[4]See Dustmann et al. (2009).

also contributed to the widening of the overall distribution of disposable personal incomes. Rising unemployment and rising wage dispersion are not the only factors that may have been responsible for the increase in overall inequality. Other factors include demographic changes, changes in living arrangements, changes in characteristics such as age or educational qualifications, and changes in the tax- and transfer system (see OECD (2008)).

While a large number of studies has focussed on such individual factors, surprisingly little is known about the relative importance of the different factors for the observed changes in the overall distribution of incomes. Changes in the distribution of incomes have been well-documented for many countries (see e.g. OECD (2008) and the references therein), but it is unclear which of the many possible candidates are the main drivers of distributional change. This is all the more surprising as the final distribution of disposable incomes is what seems most interesting from a policy point of view. For example, in view of the Krugman hypothesis, it seems highly relevant to know whether rising income inequality in Germany is more the result of a widening wage distribution, or the result of rising unemployment. This requires a comprehensive view of the income distribution including its different economic, social and institutional determinants such as demographic aspects, employment outcomes, remuneration of market activities, taxes and government transfers.[5]

In this paper, we provide a detailed examination of the main causes for rising income inequality in Germany in a unified framework. Building on previous work by Hyslop/Mare (2005) for New Zealand and Daly/Valetta (2006) for the United States, we use the semi-parametric kernel density reweighting methodology originally developed by DiNardo et al. (1996) in order to decompose the unprecedented increase in inequality and poverty in Germany from 2000 to 2006 into a number of components. We consider in particular i) changes in the distribution of household types, ii) changes in the distribution of socio-economic attributes such as age or educational qualifications, iii) changes in employment outcomes conditional on such characteristics, iv) changes in market returns to characteristics including changes in labour market returns, and v) changes in the tax system. Our results complement previous studies on the German income distribution,[6] which documented some of the developments considered here, but which did not attempt to quantify their relative importance for the overall development of the distribution. Our findings suggest

[5] Such a comprehensive view of the income distribution has also been adopted in a recent study by Checchi/Garcia-Penalosa (2010).

[6] See e.g. Hauser/Becker (2003), Federal Government of Germany (2008), German Council of Economic Experts (2009), and Grabka/Frick (2010).

that roughly three quarters of the increase in inequality and poverty between 2000 and 2006 are accounted for by rising unemployment, rising inequality in market returns, and by changes in the tax system. Each of the three factors seems to have contributed a roughly equal share to the overall increase. By contrast, changes in the distribution of household types, for example the increasing share of single person and lone parent households, and other developments such as changing educational qualifications and immigration seem to have played a minor role.

The remainder of this paper is structured as follows. Section 2 provides an informal discussion of the increase in inequality and poverty in Germany between 2000 and 2006 and its possible causes. In section 3, we describe our methodological setup. Section 4 discusses some data and specification issues. In section 5, we present our empirical analysis. Section 6 concludes.

2 Possible sources of increasing inequality

In this section, we provide an informal discussion of the increase in inequality and poverty in Germany over the period 2000 to 2006. This increase is shown in figures 1 and 2.[7] Before the year 2000, inequality in equivalized incomes remained remarkably stable (in the first half of the 1990s), or was even slightly decreasing (in the second half of the 1990s), see e.g. Biewen (2000), and Grabka/Frick (2010)).

— Figures 1 and 2 about here —

As measured by commonly used indices, inequality and poverty increased considerably between 2000 and 2006. For example, the ratio between the 90 percent and the 10 percent quantile increased from roughly 3.3 in 2000 to 3.9 in 2006, while the Gini increased from .26 to .30. Similarly, the percentage of individuals below the commonly used poverty line of 60 percent of median equivalized income rose from 12 percent in 2000 to 16.5 percent in 2006. For sample size reasons, we will pool in our analysis the years 1999/2000 and 2005/2006, which is well justified given the roughly constant level of inequality and poverty in the years we pool. Also note that

[7] Our income concept is yearly equivalized post-government personal income, which is calculated as the sum of income from all sources in a given household (including government transfers), net of taxes and social security contributions. The resulting value is then divided by an equivalence scale and distributed equally among household members. More details on the definition of our variables are given in section 4.

the inequality increase took place during a period of stagnating mean income (see last graph of figure 2).

What are possible candidates to explain this considerable increase in inequality and poverty? As indicated above, the distribution of personal disposable incomes is the highly complex result of a multitude of economic, institutional and social processes such as household formation processes, employment outcomes, returns to employment and capital, private and public transfers, and taxes and social security contributions. We focus on the following factors which appear to be the most likely candidates to explain the observed inequality increase.

Rising unemployment

As figure 3 shows, the period 2000 to 2005 was one of steep unemployment growth. At the peak in 2005, there were almost 5 million registered unemployed in Germany. Increasing unemployment is expected to increase inequality in disposable incomes because a growing fraction of the population loses employment and therefore wage income. Unemployment benefits are typically much lower than previous wage income, and they decrease or cease to be paid as unemployment continues. Given that the increase in unemployment over the period 2000 to 2005 was considerable, we expect this increase to have large impact on income inequality. The contribution to the increase in overall inequality will be the larger, the more unemployment growth is concentrated in the lower part of the income distribution. This blends well with the hypothesis that skill-biased technological progress especially affects the employment prospects of low-skilled worker (see the discussion above). Also note that during the period of the unemployment increase, overall employment was stagnating (see figure 3).

— Figure 3 about here —

Increasing dispersion of market returns

A second possible source of increasing inequality is increasing inequality in market returns, especially labour market returns. This has been the focus of many previous studies. The common perception is that the effects of skill-biased technological progress, which is seen as the main cause for the widening wage distribution in Anglo-Saxon countries since the 1980s,[8] reached the German labour market with a delay. In Germany, wage inequality started to grow in a clear way

[8] See e.g. Bound/Johnson (1992), Levy/Murnane (1992), Murphy/Welch (1992), Gosling et al. (2000).

from the mid-1990s onwards, see Kohn (2006), Genandt/Pfeiffer (2007), Dustman et al. (2009), Fuchs-Schündeln et al. (2010), Antonczyk et al. (2010a, 2010b). The common perception is that wage inequality increased both between and within skill groups, and that increases at the top are well explained by skill-biased technical progress, while increases in the lower tail of the distribution are better explained by additional factors such as deunionization and supply side effects (Dustman et al. (2009), Antonczyk et al. (2010b)).

Given the fact that labour market incomes are only one component of overall income and that they are transformed by the tax and transfer system, it is unclear to what extent the observed widening of labour market returns contributed to the increase in overall income inequality. Moreover, wage income is not the only form of market income. Other forms of market income include income from self-employment and capital income. We include these forms of income in our analysis by considering market incomes from all possible sources. Increases in inequality in these sources may also have contributed to the overall inequality increase over the period 2000 to 2006. For example, there is evidence that wealth inequality increased over this period, implying that capital incomes also grew more unequal (see German Council of Economic Experts (2009), and Frick/Grabka (2009)).

Changes in the tax system

As in many other countries, the German tax schedule experienced several changes between 2000 and 2006. The main changes are summarized in table 1. Tax rates were generally reduced, but reductions were somewhat higher at the top of the distribution. Given that some of the changes were considerable, it seems likely that these changes had some impact on the final distribution of disposable income.

— Table 1 about here —

Changes in the household structure

There are clear trends in the way household structures change in industrialized countries.[9] In particular, there is a trend towards smaller households and towards untypical household forms such as single parents. Given that incomes are pooled within households and given the fact that different household types systematically differ in their average income, changes in the distribution of household types may also have a potentially large effect on the overall income distribution.

[9] See OECD (2008), ch. 2.

— Figure 4 about here —

Figure 4 displays the evolution of population shares of a number of household types over the period under consideration. The figure shows that the population shares of traditional household forms such as couples with or without children are in decline, while the population shares of single adult households, lone parents, and pensioner households are steadily increasing. However, it also seems clear that population shares only change very gradually so that their effects on the distribution of income over a relatively short time span is probably limited when compared to those of the more severe changes in unemployment and wage dispersion. The effect of the secular decline of household size on the income distribution in Germany was studied by Peichl et al. (2010). Not explicitly considering other influences on the income distribution, they find that the effect of declining household sizes is indeed limited, even over a period of 20 years. Nevertheless, it seems necessary to account for such changes when studying the effect of other factors such as unemployment and market returns.

Changes in other socio-economic attributes

There are, apart from the household form, a number of other characteristics whose change over time may have a potentially large influence on the income distribution. These are in particular changes in the age structure of the population (increasing share of the elderly, and the decreasing shares of children and young persons), changes in educational qualifications (secular skill-upgrading), and other changes in the composition of the population, e.g. due to immigration. We explicitly take account of these changes in our analysis by modeling and quantifying this 'characteristics'-effect in our decompositions.

Other changes

We will capture distributional changes induced by factors other than the ones listed above in the 'residual' of our decomposition analysis. It turns out that the unexplained 'residual' of our analysis is relatively small so that the factors listed above successfully account for most of the inequality increase from 2000 to 2006. What may be the such other factors? A factor which may also have a potentially large impact on the distribution of net incomes are changes in Germany's highly complex system of government transfers. Due to the high complexity of these transfers,[10] we refrain from explicitly modeling changes in these transfers. This seems unproblematic as there

[10] They include transfers for children, students, mothers, the disabled, the unemployed, housing allowances, general social assistance, to name only the most important ones.

were only minor changes in these transfers over the period 2000 to 2005.[11] The fact that the residual component of our decomposition is relatively small confirms our conjecture that changes in the transfer system had only a small effect on the development of the income distribution over the years 2000 to 2006.

3 Estimation of counterfactual income densities

Following DiNardo et al. (1996) and Hyslop/Mare (2005),[12] we use a semiparametric decomposition technique to decompose the inequality increase from 1999/2000 ('period 0') to 2005/2006 ('period 1') into different components that are attributable to the factors listed above. The basic idea of this decomposition technique is that of a shift-share analysis, in which observations are reweighted according to whether they are over- or underrepresented in a counterfactual situation. Counterfactual situations are obtained by holding some aspects of interest fixed at the period 0 level, while changing others to the period 1 level. The method has its limitations in that it cannot account for interactions between the different factors in the form of behavioural reactions or general equilibrium effects. Nevertheless, it is generally believed that counterfactual reweighting and simulation exercises convey important information about the main drivers of distributional changes.[13]

Stage 1: Changes in the distribution of household types

As a first stage we consider the effect of shifts in the composition of the population with respect to a number of household types (we distinguish between the six household types mentioned above, see figure 4). The counterfactual income distribution in which everything is as in period 0, but the distribution of household types is shifted to that of period 1 is given by

$$f_0(y|t_h = 1) = \sum_{j=1}^{6} w_{1j} f_{0j}(y), \tag{1}$$

[11] However, a major reform, the so-called 'Hartz-Reform' was enacted in 2005. Due to several transitional rules, the full effects of the 'Hartz-Reform' on the income distribution are probably measurable only after 2006. Due to the high complexity of the changes and due to reasons of data availability for the years after 2006 (new data become available with a considerable delay because incomes are asked retrospectively for the previous year) along with the need to pool years, we defer an analysis of these changes to future research.

[12] For a similar application, see Daly/Valetta (2006).

[13] See Fortin et al. (2010).

where y denotes net equivalized personal income, w_{1j} is the population share of household type j in period 1, and f_{0j}(y) the income distribution of individuals from household type j in period 0. Analogously, $f_0(y|t_h = 0)$ would be the factual income distribution in period 0, where w_{1j} is replaced by the factual population shares w_{0j}.

Stages 2 and 3: Changes in socio-economic attributes and employment outcomes

The second and third stages of our decompositions account for changes in the distribution of socio-economic attributes x (e.g. the age and educational composition of the household, see below for more details) and changes in household employment outcomes e conditional on these attributes x. For example, the counterfactual income density for individuals living in household type j in which everything is as in period 0 but the distribution of socio-economic attributes x and the distribution of household employment outcomes e conditional on socio-economic attributes are as in period 1, is given by

$$f_{0j}(y|t_x = 1, t_e = 1) = \int_e \int_x f_{0j}(y|x, e)\, dF_{1j}(e|x)\, dF_{1j}(x) \tag{2}$$

$$= \int_e \int_x f_{0j}(y|x, e) \left[\frac{dF_{1j}(e|x)}{dF_{0j}(e|x)}\right] dF_{0j}(e|x) \left[\frac{dF_{1j}(x)}{dF_{0j}(x)}\right] dF_{0j}(x) \tag{3}$$

$$= \int_e \int_x \Psi_{e|x,j} \cdot \Psi_{x|j} \cdot f_{0j}(y|x, e)\, dF_{0j}(e|x)\, dF_{0j}(x). \tag{4}$$

Note that the counterfactual distribution $f_{0j}(y|t_x = 1, t_e = 1)$ is just a reweighted version of the factual distribution $f_{j0}(y)$ with reweighting factors $\Psi_{e|x,j}$ and $\Psi_{x|j}$. The factual distribution $f_{j0}(y)$ can be obtained by setting $\Psi_{e|x,j} = \Psi_{x|j} = 1$. Analogously, $f_{0j}(y|t_x = 1, t_e = 0)$ with $\Psi_{e|x,j} = 1$ is the counterfactual distribution where only the distribution of characteristics x is shifted to that of period 1 (while the conditional employment and everything else is held fixed at its period 0 level). Finally, $f_{0j}(y|t_x = 0, t_e = 1)$ with $\Psi_{x|j} = 1$ would be the distribution where only conditional employment outcomes are changed to the period 1 level, but everything else is held fixed at the period 0 level.

The reweighting factors $\Psi_{e|x,j}, \Psi_{x|j}$ can be rewritten as

$$\Psi_{x,j} = \frac{P_j(x|t = 1)}{P_j(x|t = 0)} = \frac{P_j(t = 1|x)}{P_j(t = 0|x)} \cdot \frac{P_j(t = 0)}{P_j(t = 1)}, \tag{5}$$

$$\Psi_{e|x,j} = \frac{dF_{1j}(e|x)}{dF_{0j}(e|x)} = \frac{P_{1j}(e|x)}{P_{0j}(e|x)}. \tag{6}$$

Equation (6) nicely illustrates how the reweighting procedure works. For example, if a particular household employment outcome e for an observation with household characteristics x

is less likely in period 1 compared to period 0, this observation will be down-weighted in period 0 when constructing the counterfactual density. Following Hyslop/Mare (2005), we define household employment outcomes e as an ordinal variable (for details, see below), so that reweighting factor $\Psi_{e|x,j}$ can be estimated using predictions from ordinal logit models $P_{1j}(e|x)$ and $P_{0j}(e|x)$. Analogously, reweighting factor $\Psi_{x|j}$ can be estimated using predictions from logit models $P_j(t=1|x), P_j(t=0|x)$ and the ratio of observational mass in period 0 and period 1.

Stage 4: Changes in market returns and changes in tax system

In stages 4 and 5 of our decomposition we consider changes in market returns to household characteristics $z=(e,x)$ and changes in the tax schedule. Counterfactual income y_0^{cf} in period 0 accounting for the expected change $\widehat{\Delta} y_{gross} = {z_0}'\hat{\beta}_{1j} - {z_0}'\hat{\beta}_{0j}$ in gross market income due to changes in the returns to household characteristics $\Delta\hat{\beta}_j = \hat{\beta}_{1j} - \hat{\beta}_{0j}$ and for changes in the tax schedule is given by

$$y_0^{cf} = y_{gross,0} + \widehat{\Delta} y_{gross} + y_{transf,0} - y_{sscontr,0} - \ tax_1(y_{gross,0} + \widehat{\Delta} y_{gross}) \tag{7}$$

where $y_{gross,0}$ denotes period 0 market incomes from all sources, $y_{transf,0}, y_{sscontr,0}$ period 0 government transfers and social security contributions, and $tax_1(\cdot)$ the tax schedule of period 1. The counterfactual income that results if only the tax schedule is changed to that of period 1 but market returns to household characteristics are fixed at the period 0 level is obtained by setting $\widehat{\Delta} y_{gross} = 0$.

In a similar way, the counterfactual income that results if only market returns are changed to the period 1 levels but the tax schedule is held fixed at its period 0 level is given by

$$y_0^{cf} = y_{gross,0} + \widehat{\Delta} y_{gross} + y_{transf,0} - y_{sscontr,0} - \ tax_0(y_{gross,0} + \widehat{\Delta} y_{gross}). \tag{8}$$

In short-hand notation, the changes to income in period 0 due to counterfactual variations of market returns and the tax schedule can be expressed as

$$y_0^{cf} = y_{net,0} + \widehat{\Delta} y_{gross} - \widehat{\Delta} t, \tag{9}$$

where $y_{net,0}$ is the factual net income of period 0.

Counterfactual densities incorporating stages 1 to 5

Combining equations (1) through (9) one can define counterfactual income densities that combine any desired set of counterfactual variations. For example, the overall income distribution in period

0 that results if household structures, employment outcomes, and market returns are fixed at their period 0 levels but the distribution of socio-economic attributes and the tax schedule are counterfactually set to their period 1 levels, is given by

$$f_0(y|t_h = 0, t_x = 1, t_e = 0, t_r = 0, t_t = 1) = f_0(y|0, 1, 0, 0, 1). \tag{10}$$

Following DiNardo et al. (1996), counterfactual densities $f_0(y|t_h, t_x, t_e, t_r, t_t)$ are estimated as

$$\hat{f}(y|t_h, t_x, t_e, t_r, t_t) = \sum_{j=1}^{6} \sum_{i=1}^{n_j} \theta_i \Psi_j \Psi_{x|j} \Psi_{e|x,j} K \left(\frac{y - (y_{net,0,i} + \hat{\Delta} y_i - \hat{\Delta} t_i)}{h} \right) \frac{1}{h}, \tag{11}$$

where θ_i denotes the sample weight of individual i, n_j is the number of individuals in household type j, $K(.)$ a kernel function, h a bandwidth, and $\Psi_j = w_{1j}/w_{0j}$. If a particular counterfactual variation is not desired, the corresponding weighting factor $\Psi_j, \Psi_{x|j}, \Psi_{e|x,j}$ is set to 1, or the corresponding shift factor $\widehat{\Delta} y_{gross}, \widehat{\Delta} t$ is set to zero, respectively.

Estimation of inequality and poverty indices

Given an estimated income density $\hat{f}(y)$, we use numerical integration methods to calculate the inequality and poverty indices shown in table 2.[14]

— Table 2 about here —

4 Data and specification issues

We base our analysis on data from the German Socio-Economic Panel (GSOEP) for the years 1999 to 2006. As indicated above we pool the years 1999/2000 ('period 0') and 2005/2006 ('period 1') in order to increase sample sizes and to make our analysis less dependent on particular years. The total sample sizes are 45,085 individuals for period 0 and 50,427 individuals for period 1. Our data refers to individuals (including children). All our calculations are weighted with the appropriate sample weights.[15] Our main income variable is real annual equivalized personal net income which is calculated from annual net household income. Annual net household income is given by

$$\text{net income} = \text{gross income} + \text{transfers} - \text{social security contributions} - \text{taxes}. \tag{12}$$

[14]For the definition and properties of these indices, see Cowell (2000)

[15]We do not use the so-called 'high-income sample' G because the validity of its sample weights is questionable.

Our data set comprises all of the listed components of net income. Taxes were calculated by the data provider, the DIW Berlin, using the official rules.[16] In order to compute the individual income of the members of a given household, household net income is divided by the sum of equivalence weights defined by the so-called OECD equivalence scale (the household head receives a weight of 1, additional household members over 14 years receive a weight of .5, household aged 14 years or less receive a weight of .3). In a robustness analysis, we consider two alternative equivalence scales to see whether our results depend on this particular choice (see below). Following recommendations and practice of the Statistical Office of the European Commission, we set the poverty line to 60 percent of the median of equivalized personal incomes in a given year. Note that our definitions are the same as the ones used in the official 'Report on Poverty and Richness' published by the federal government.[17]

As indicated above, we define six different household types: i) single pensioner households (65 years or older), ii) multiple pensioner households (at least one household member is 65 years or older and no household member is under 55), iii) single adults without children, iv) multiple adults without children, v) single adults with children, and vi) multiple adults with children. As socio-economic household attributes x we consider the number of adults in the household, the fraction of female adults in the household, the fraction of adult household members with different educational qualifications (university degree, high school and/or vocational training, no such degree or qualification), the fraction of adult household members with non-German nationality, the fraction of adult household members with disabilities, the fraction of married adults in the household, the fraction of household members in different age groups (0-3 years, 4-11 years, 12-17 years, 18-30 years, 31-50 years, 51-64 years, 65 or older), and a dummy indicating whether the household resided in East Germany (see table 6 in the appendix for details).

Employment outcomes e are defined in an ordinal way: i) no part-time or full-time workers in the household, ii) no full-time workers but at least one part-time worker, iii) one full-time worker but no part-time workers iv) one full-time worker and at least one part-time worker, v) at least two full-time workers. We estimate the probability for each household employment outcome e conditional on socio-economic attributes x using ordinal logit models. All estimations are carried out for each household type separately (see table 7 in the appendix for details). In order to estimate gross market returns, we regress gross household income from all sources on socio-

[16] See Grabka (2009) for more details on the definition of the different income variables.

[17] See Federal Government of Germany (2008). The only difference is that we do not consider imputed rental values as income.

economic attributes, employment outcome categories, and a full set of interactions. We drop regressors that are grossly insignificant in both periods. Again, all regressions are carried out separately for each household type (see table 8 in the appendix for details).

The tax schedule is estimated using a flexible polynomial in household gross income along with suitable interactions with variables such as marital status or children (i.e. we regress the household tax variable as given in the data on a polynomial in gross income and interactions with other characteristics). Our regressions fit the tax values given in the data extremely well, making us confident that we use the correct tax schedule. The regressions are only carried out for nonzero tax values. A household pays no tax if its gross income is below the sum of personal tax exemptions. When calculating counterfactual tax values, we first check whether this is the case. We then calculate positive tax values using the estimated tax schedule only if household gross income exceeds the sum of personal tax exemptions and impute a value of zero otherwise.

Note that our analysis refers to inequality in net income *between individuals* (not households). All data are individual data but individuals are attributed the characteristics and the (equivalized) incomes of the households they live in. Incomes are expressed in year 2000 Euros (except for tax calculations which require nominal incomes). For expositional reasons we consider log equivalized incomes, which we appropriately transform back when calculating the inequality and poverty indices in table 2.

5 Empirical results

This section presents our empirical results. The decomposition of the overall distribution into the different components representing the population subgroups described by household types is illustrated in figure 5 (note that all figures show distributions of *logged* incomes). It clearly emerges that multiple adult households with or without children represent by far the largest portion of the overall distribution, while single person households and pensioner households play a smaller role. The contribution of single parents households is particularly small, which is due to their relatively small population share (see figure 4). One can also see that the incomes of individuals in multiple adult households are mostly located in the upper part of the distribution, while those of individuals in multiple adult households with children and pensioner households are mostly located in the middle and the lower parts of the overall distribution.

5.1 Overall change in density from 1999/2000 to 2005/2006

Figure 6 shows how the overall shape of the (log) income distribution changed from 1999/2000 ('period 0') to 2005/2006 ('period 1'). The picture that emerges is one of increasing spread, i.e. the distribution in 2005/2006 has a lower peak and fatter tails than the one in 1999/2000. However, the widening of the distribution is not symmetric, the changes seem particularly pronounced in the lower tail of the distribution, implying that low incomes were particularly affected by increasing inequality.

5.2 'Ceteris paribus' effects of individual factors

Next, we consider 'ceteris paribus' effects of the different factors, i.e. we change only one factor at a time to its period 1 level, but hold everything else fixed at the level of the base period 0. We believe that such an exercise comes close to what one has in mind when asking about the 'effect' of a particular factor on the overall change. For example, the dashed line in figure 7 shows the income distribution that would prevail if the distribution of household types was changed to that of period 1, but everything else was held fixed at its period 0 level. The figure shows that changes in household structures alone did not contribute much to the change in the distribution between 1999/2000 to 2005/2006 (there was a slight shift of mass from the middle to the lower part of the distribution, but this effect was rather small).

In a similar way, figure 8 confirms that changing only the distribution of socio-economic attributes to its period 1 level but leaving everything else as in period 0, did not have any perceivable effect on the overall distribution. By contrast, figure 9 shows the effect of changing only conditional employment outcomes but leaving everything else constant. This considerably stretches the lower and middle part of the distribution to the left, suggesting that changes in unemployment and part-time employment affected in particular individuals in the middle and lower part of the distribution. High income households appeared to be largely unaffected by such changes. Figure 10 shows the effect of a ceteris paribus change in market returns. This also has a considerable effect. The distribution is shifted to the right, but more so for middle and higher incomes (i.e. middle and high income households benefited more from changes in the market remuneration of household characteristics). Finally, figure 11 presents the ceteris paribus effect of changes in the tax schedule. These also shifted the distribution to the right, but much more so for middle and especially for high incomes. This suggests that middle and high incomes benefited overproportionally from reduced tax rates, while the density in the very low end of the distribution remained practically constant (these households usually do not pay any tax at all).

The analysis of such ceteris paribus changes by household type reveals interesting additional patterns. For example, figure 12 shows the effects of a ceteris paribus change in conditional employment outcomes by household type. Obviously, individuals in pensioner households are only marginally affected by such changes (there is generally very little employment in these households), see top row of figure 12. The household types most affected by changes in employment probabilities are single adult households and multiple adult households with children. By contrast, multiple adult households without children were practically unaffected by changes in conditional employment prospects. Single parents household were also hit unfavourably by changing employment probabilities, presumably in connection with changes in part-time work. Figure 13 presents the results for a ceteris paribus change in market returns. The most interesting feature of this figure is that changes in market returns did not only affect household types with a close attachment to the labour market but also pensioner households whose market income is composed of interest, dividend and rent income rather than of wage income. This suggests that increasing inequality in gross incomes was not only the result of a widening wage distribution, but also the result of a widening distribution of capital income. All other household types were hit in a similar way by a distributional shift of the middle and upper part to the right. For these household types, changes in returns on the labour market were probably more relevant than changes in returns to capital. Finally, figure 14 displays the effects of a ceteris paribus change in the tax schedule. Pensioner households were only affected by such changes in the upper part of the distribution as most ordinary lower pensions were completely exempt from taxation during the period under consideration. The results for the other household types indicate that in particular the well-off multiple adult households without children benefited from the reductions in tax rates between 2000 and 2005.

Going back to the overall distribution, table 3 summarizes what percentage of the overall increase in inequality as measured by various inequality and poverty indices can be explained by changing one factor at a time. The numbers largely confirm the findings from the graphical analysis. Only a small percentage of the overall inequality increase can be explained by isolated changes in the distribution of household types (around 7 percent on average) or by changes in socio-economic attributes (around 5 percent on average). Changes in conditional employment outcomes explain on average 26 percent of the total increase (column 3), changes in market returns on average 24 percent (column 4), and changes in the tax system on average 26 percent (column 5). The general conclusion is that changes in employment outcomes, changes in market returns, and changes in the tax schedule explain the inequality increase in equal parts, while changes in household structures and socio-economic attributes play only a minor role. It is also interesting to look at individual contributions in different parts of the distribution. For example, changes in conditional employment outcomes contribute to an especially large extent to the increase in the poverty head count $FGT(0)$ and the poverty gap measure $FGT(1)$, while changes in market returns and changes in the tax schedule contribute more strongly to the increase of indices that measure inequality at the top of the distribution (see $p9050$).

A drawback of the ceteris paribus analysis presented so far is that the percentages contributed by each factor do not add up to the complete change. In the next section, we therefore present an exact decomposition whose contributions add up, and which also allows one to assess the importance of residual factors.

5.3 Exact decomposition of the increase in inequality and poverty

We proceed in the usual fashion,[18] and decompose the inequality increase 1999/2000 to 2005/2006 into a sequence of incremental changes that result when changes of the individual factors are accumulated. Using this idea, the change in inequality between 1999/2000 and 2005/2006 can be decomposed as

$$
\begin{aligned}
I\left(\hat{f}_1(y)\right) - I\left(\hat{f}_0(y)\right) &= \left[I\left(\hat{f}_0(y|1,0,0,0,0)\right) - I\left(\hat{f}_0(y|0,0,0,0,0)\right)\right] && (13)\\
&+ \left[I\left(\hat{f}_0(y|1,1,0,0,0)\right) - I\left(\hat{f}_0(y|1,0,0,0,0)\right)\right] && (14)\\
&+ \left[I\left(\hat{f}_0(y|1,1,1,0,0)\right) - I\left(\hat{f}_0(y|1,1,0,0,0)\right)\right] && (15)\\
&+ \left[I\left(\hat{f}_0(y|1,1,1,1,0)\right) - I\left(\hat{f}_0(y|1,1,1,0,0)\right)\right] && (16)\\
&+ \left[I\left(\hat{f}_0(y|1,1,1,1,1)\right) - I\left(\hat{f}_0(y|1,1,1,1,0)\right)\right] && (17)\\
&+ \left[I\left(\hat{f}_1(y)\right) - I\left(\hat{f}_0(y|1,1,1,1,1)\right)\right], && (18)
\end{aligned}
$$

where $I(\cdot)$ is one of the inequality or poverty indices in table 2. The overall inequality change $I\left(\hat{f}_1(y)\right) - I\left(\hat{f}_0(y)\right)$ is decomposed into parts contributed by changes in the household structure (13), changes in socio-economic characteristics (14), changes in conditional employment

[18]See e.g. DiNardo et al. (1996) and Hyslop/Mare (2005).

outcomes (15), changes in market returns (16), changes in the tax schedule (17), and an unexplained residual (18).

Table 4 shows the contributions of each of the factors as a percentage of the overall inequality increase. For example, some 5.31 percent of the increase of the Gini coefficient from 1999/2000 to 2005/2006 are attributable to changes in household structures. The results largely reproduce the findings from the ceteris paribus analysis. Changes in household structures and socio-economic attributes contribute relatively little (around 7 and 5 percent on average), whereas changes in employment outcomes explain some 25 percent, changes in market returns 27 percent, and changes in the tax schedule around 21 percent of the overall inequality increase on average. The unexplained residual amounts to 20 percent on average, implying that most of the inequality increase is successfully accounted for by the factors considered above.

5.4 Sensitivity analysis

Decomposition (13) to (17) contains an element of arbitrariness in that the order in which the different factors are cumulated could also have been different (e.g. one could have started with changing conditional employment outcomes instead changing the distribution of household types).[19] In order to assess how sensitive our decomposition results are with respect to the order chosen, we carried out the decomposition in all possible $5! = 120$ orders. The results of this exercise are shown in table 5 in the appendix. It turns out that the contributions of the different factors are reasonably stable when the order of the decomposition is changed. Moreover, we believe that the sequential order used in (13) to (17) is more reasonable than many of the other possible orderings for two reasons. First, in decomposition (13) to (17) factors are basically changed in the order of their 'pre-determinedness', i.e. household type and socio-economic attributes are chosen before employment outcomes, market incomes are the result of household characteristics and employment outcomes, and taxes are both the result of market incomes and household characteristics. Second, the order used in (13) to (17) essentially reproduces the contributions that result from the ceteris paribus analysis, which is appealing on a-priori grounds, but which has the disadvantage of non-additivity.

We also carried out further sensitivity checks, in particular we varied the bandwidth used in our

[19] Biewen (2001) illustrates the problems of possible path dependencies in sequential decompositions such as the one considered here.

density estimations and the equivalence scale used to make incomes comparable across household types. A combination of graphical inspection and Silverman's rule of thumb led us to use a fixed bandwidth of .175 throughout our analysis.[20] Our numerical results change only little if we vary the bandwidth between .1 and .3, and qualitative results remain unchanged. The same applies if we use two alternative equivalence scales (we used the so-called Luxembourg scale which deflates household incomes by the square root of household size, and another scale which assigns a weight of 1 to the household head, and weights of .7 and .5 to additional household members over 14 years, and up to 14 years, respectively).

6 Conclusion

This paper addressed the question of which factors were behind the recent increase in personal income inequality in Germany. Using a variant of DiNardo et al. (1996)'s semi-parametric reweighting methodology, we decompose the increase in income inequality and poverty from 1999/2000 to 2005/2006 into components explained by i) changes in the distribution of household types, ii) changes in the distribution of socio-economic characteristics, iii) changes in employment probabilities conditional on characteristics, iv) changes in market returns to characteristics, and v) changes in the tax system. Our results suggest that most of the inequality increase can be explained by both changes in employment outcomes and in market returns, and, to a similar extent, by changes in the tax system. Changes in household structures and other household characteristics seem to have played a much smaller role. Put into an international perspective, our results suggest that rising income inequality in non-Anglo-Saxon countries is the likely result of *both* increasing inequality in market returns *and* increasing inequality in employment outcomes, as well as of idiosyncratic changes such as tax reforms.

[20] Hyslop/Mare (2005) use a similar fixed bandwidth.

7 References

Antonczyk, D., B. Fitzenberger, and K. Sommerfeld (2010a): Rising Wage Inequality, the Decline of Collective Bargaining, and the Gender Wage Gap, *IZA Discussion Paper No. 4911*, forthcoming in *Labour Economics*.

Antonczyk, D., T. DeLeire, and B. Fitzenberger (2010b): Polarization and Rising Wage Inequality: Comparing the U.S. and Germany, *IZA Discussion Paper No. 4842*, Institute for the Study of Labor, Bonn.

Biewen, M. (2000): Income inequality in Germany during the 1980s and 1990s, *Review of Income and Wealth 46*, pp. 1 - 20.

Biewen, M. (2001): Measuring the Effects of Socio-Economic Variables on the Income Distribution: An Application to the East German Transition Process, *Review of Economics and Statistics 8*, pp. 185 - 190.

Bound, J. and G. Johnson (1992): Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations, *American Economic Review 82*, pp. 371 - 392.

Checchi, D., and C. Garcia-Penalosa (2010): Labour Market Institutions and the Personal Distribution of Income in the OECD, *Economica 77*, pp. 413-450.

Cowell, F. A. (2000): Measurement of Inequality, in: Atkinson, A.B. and F. Bourguignon (eds.), Handbook of Income Distribution, Volume 1, Elsevier Science B.V.

Daly, M. C. and Valetta, R. G. (2000): Inequality and poverty in the United States: the effects of changing family behavior and rising wage dispersion, *Economica 73*, pp. 75 - 98.

DiNardo, J., N. Fortin, and T. Lemieux (1996): Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach, *Econometrica 64*, pp. 1001 - 44.

Federal Government of Germany (2008): Lebenslagen in Deutschland. Der 3. Armuts- und Reichtumsbericht der Bundesregierung, Berlin.

Fortin, N., T. Lemieux, and S. Firpo (2010): Decomposition Methods in Economics, *NBER Working Paper 16045*, forthcoming in: Handbook of Labor Economics, Vol. 4.

Frick, J., and M. M. Grabka (2009): Gestiegene Vermögensungleichheit in Deutschland, *DIW*

Wochenbericht 4/2009, pp. 54 - 67.

Fuchs-Schündeln, N., D. Krüger, and M. Sommer (2010): Inequality Trends for Germany in the Last Two Decades: A Tale of Two Countries, *Review of Economic Dynamics 13*, 2010, pp. 103-132.

German Council of Economic Experts (2009): Securing the future through responsible economic policies, Annual Report 2009/10, Wiesbaden.

Gosling, A., St. Machin, and C. Meghir (2000): The Changing Distribution of Male Wages, 1966 - 1992, *Review of Economic Studies 67*, pp. 635 - 666.

Grabka, M. (2009): Codebook for the PEQUIV File 1984-2008: CNEF Variables with Extended Income Information for the SOEP, *Data Documentation 45*, DIW Berlin.

Grabka, M. M., and J. Frick (2010): Weiterhin hohes Armutsrisiko in Deutschland, *DIW Wochenbericht 7/2010*, pp. 1 - 11.

Haisken-DeNew, J. P. and Frick, J. R. (2005): Desktop Companion to the German Socio-Economic Panel Study (GSOEP), DIW Berlin.

Hauser, R., and I. Becker (2003) (eds.): Reporting on Income Distribution and Poverty. Perspectives from a German and a European Point of View, Springer, Berlin.

Hyslop D. R., and Mare, D. C. (2005): Understanding New Zealand's Changing Income Distribution 1983-98: A Semiparametric Analysis, *Economica 72*, pp. 469 - 495.

Juhn, C., K.M. Murphy, and B. Peirce (1993): Wage Inequality and the Rise in Returns o Skill, *Journal of Political Economy 101*, pp. 410 - 441.

Katz, L. F., Murphy K. M. (1992): Changes in Relative Wages, 1963-1987: Supply and Demand Factors, *Quarterly Journal of Economics 107*, pp. 35 - 78

Kohn, K. (2006): Rising Wage Dispersion, After All! The German Wage Structure at the Turn of the Century, *IZA Discussion Paper No. 2098*, Institute for the Study of Labor, Bonn.

Krugman, P. (1994): Past and Prospective Causes of High Unemployment, *Economic Review, Fourth Quarter 1994*, Federal Reserve Bank of Kansas City, pp. 23 - 43.

Levy, F., and R.J. Murnane (1992): U.S. Earnings Levels and Earnings Inequality: A Review of

Recent Trends and Proposed Explanations, *Journal of Economic Literature 30*, pp. 1333 - 1381.

Murphy, K.M., and F. Welch (1992): The Structure of Wages, *Quarterly Journal of Economics 107*, pp. 285 - 326.

OECD (2008): Growing unequal? Income distribution and poverty in OECD countries, OECD, Paris.

Peichl, A., N. Pestel, H. Schneider (2010): Does Size Matter? The Impact of Changes in Household Structure on Income Distribution in Germany, *IZA Discussion Paper No. 4770*, Institute for the Study of Labor, Bonn.

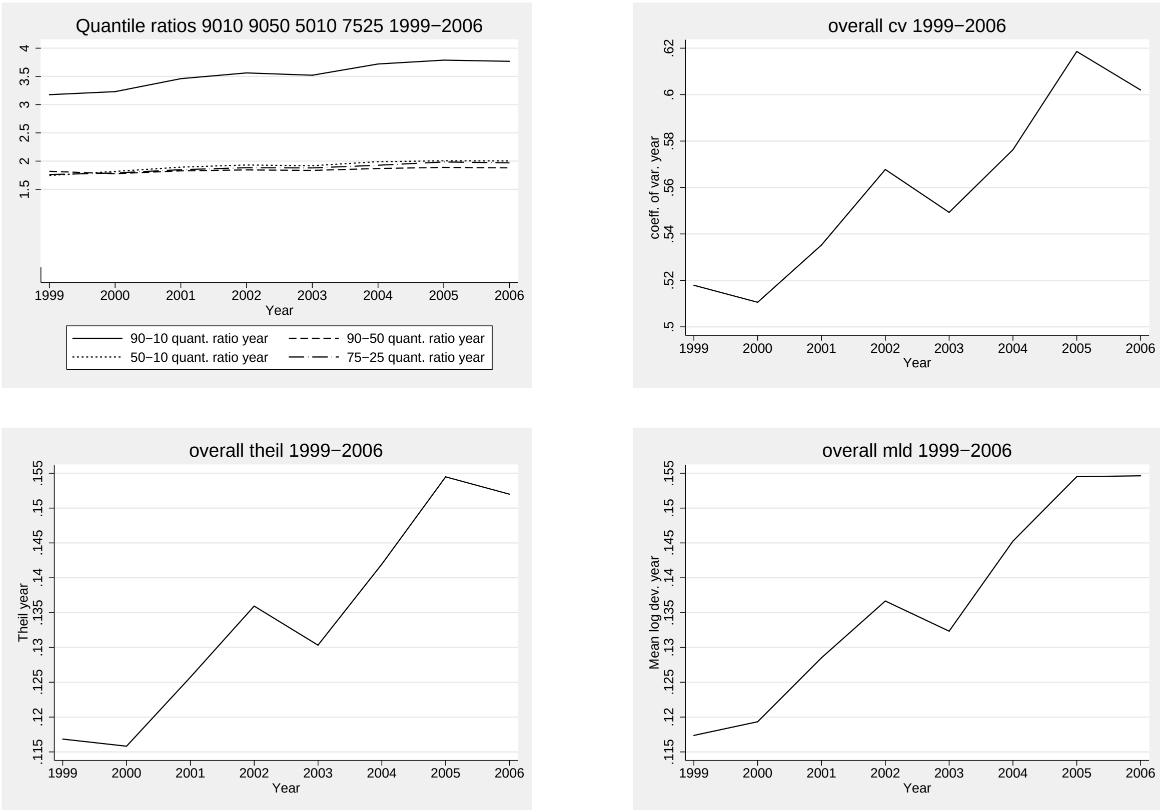
Pfeiffer, F. and Gernandt, J. (2007): Rising Wage Inequality in Germany, *Journal of Economics and Statistics 227*, pp. 358-380.

Prasad, E. S. (2004): The Unbearable Stability of the German Wage Structure: Evidence and Interpretation, *IMF Staff Papers No. 51*, International Monetary Fund.

Puhani, P. A. (2003): Transatlantic Differences in Labour Markets: Changes in Wage and Non-Employment Structures in the 1980s and 1990s, *German Economic Review 9*, pp. 312-338.

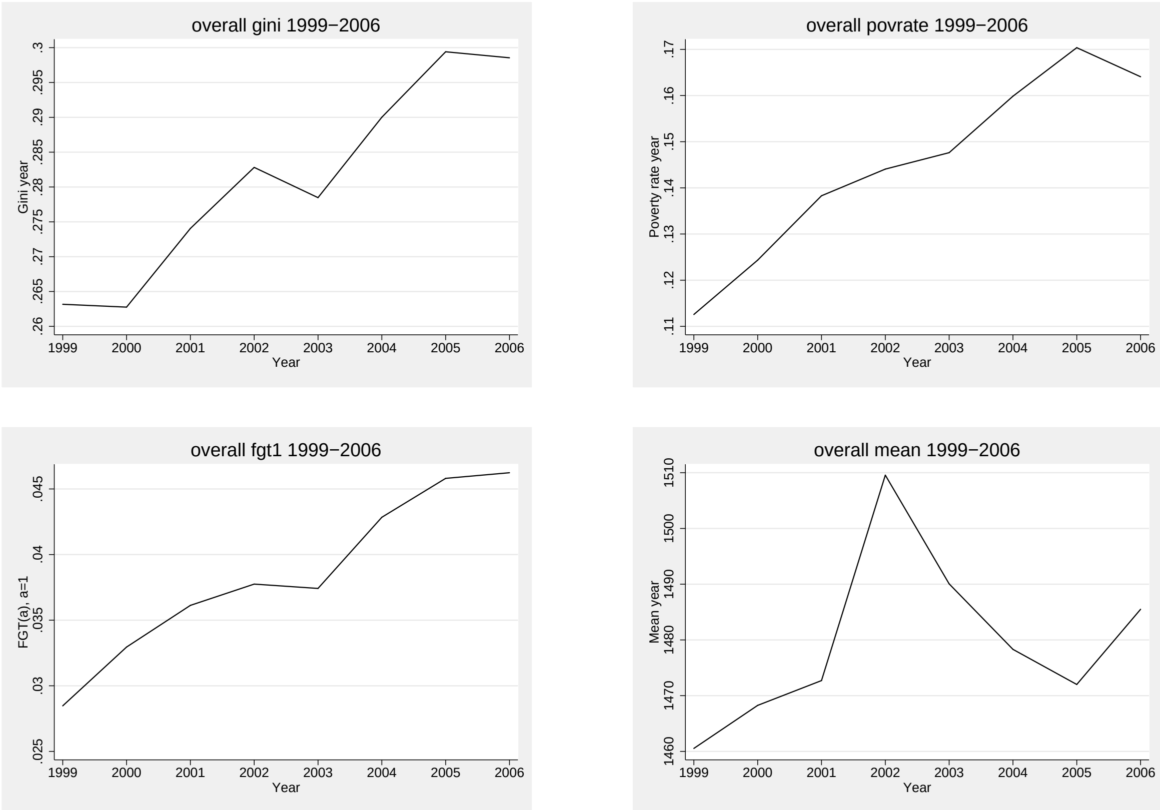
8 Figures

Figure 1 – Trends in inequality and poverty 1999-2006



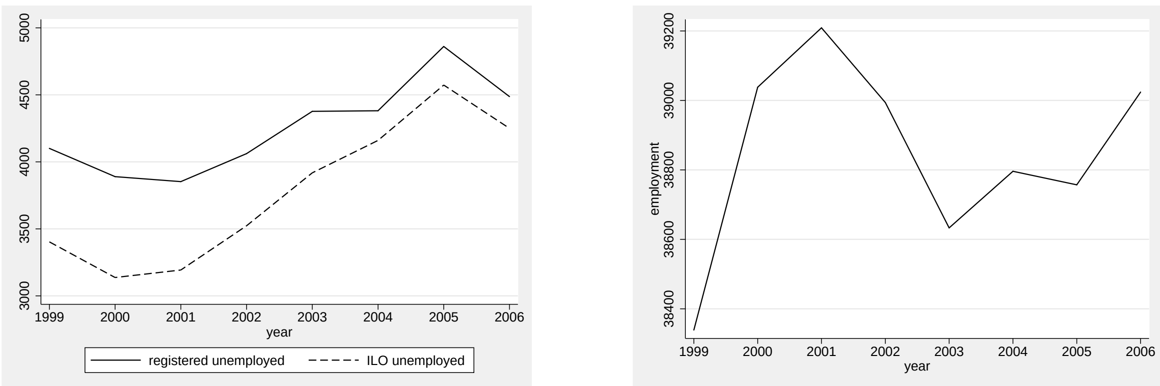
Source: GSOEP. See text for the definition of inequality measures and income variables.

Figure 2 – Trends in inequality and poverty 1999-2006



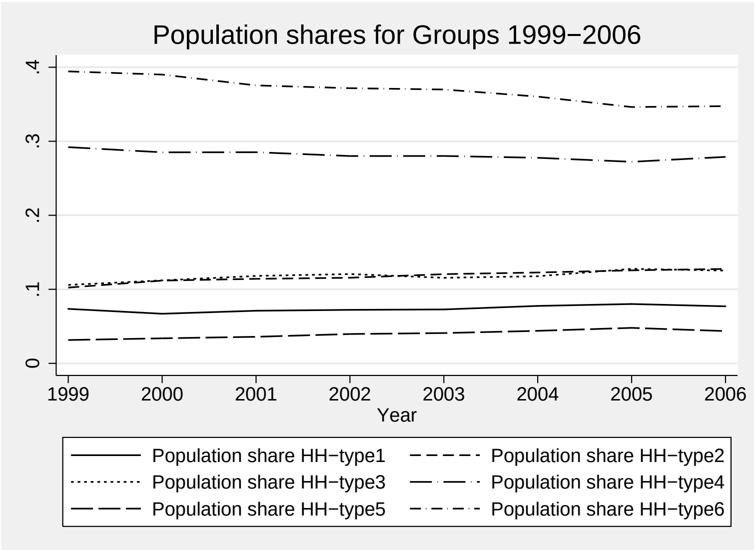
Source: GSOEP. See text for the definition of inequality measures and income variables.

Figure 3 – Trends in employment and unemployment 1999-2006



Source: German Federal Employment Office

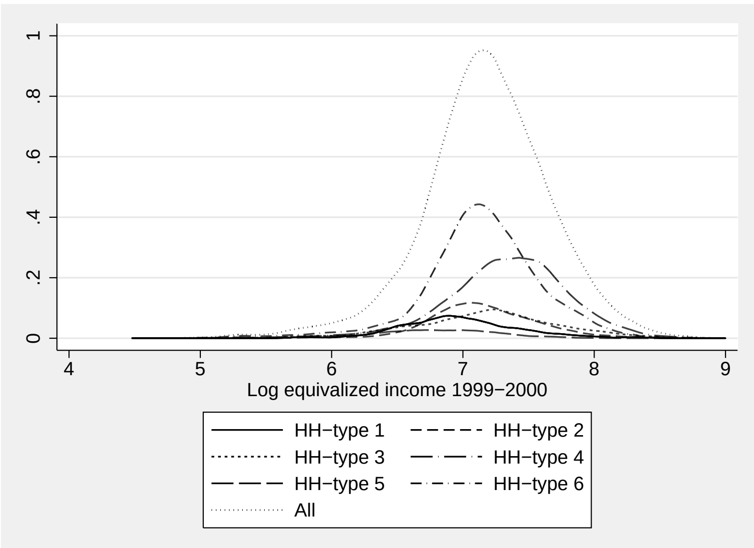
Figure 4 – Trends household structures 1999-2006



HH-type 1 = Single pensioners, HH-type 2 = Multiple pensioners, HH-type 3 = Single adults w/o kids, HH-type 4 = Multiple adults w/o kids, HH-type 5 = Single adults with kids, HH-type 6 = Multiple adults with kids

Source: GSOEP, own calculations

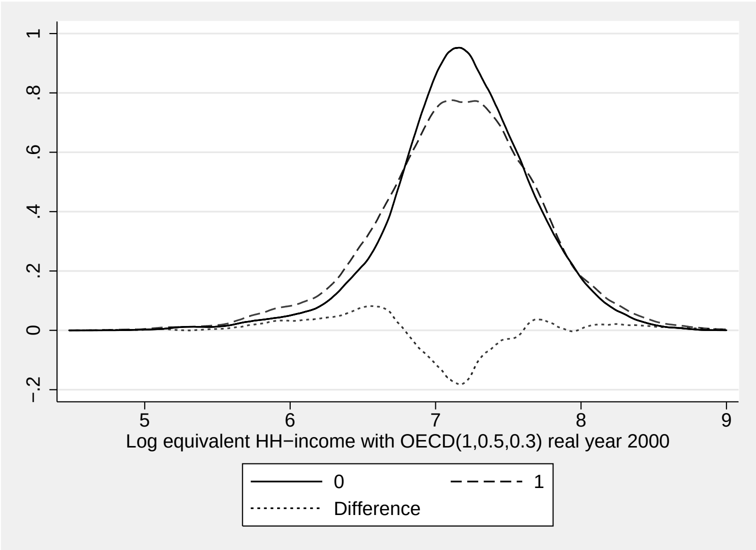
Figure 5 – Decomposition of the overall distribution by household types



HH-type 1 = Single pensioners, HH-type 2 = Multiple pensioners, HH-type 3 = Single adults w/o kids, HH-type 4 = Multiple adults w/o kids, HH-type 5 = Single adults with kids, HH-type 6 = Multiple adults with kids

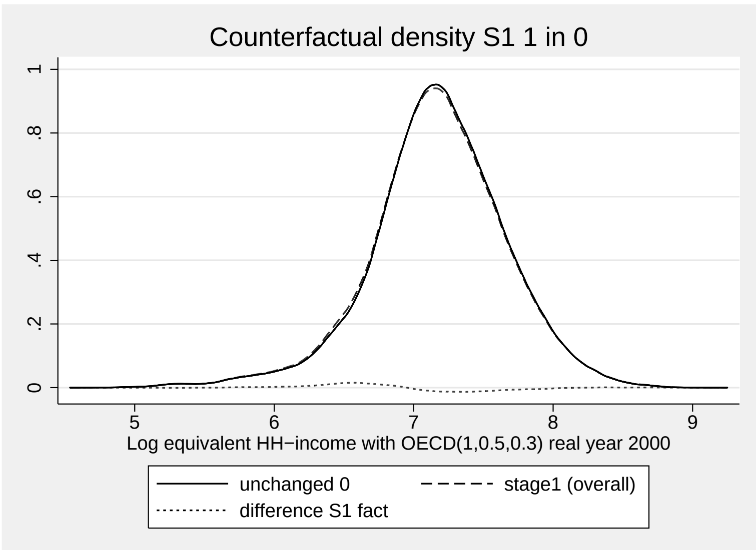
Source: GSOEP, own calculations

Figure 6 – Overall change in density from 1999/2000 ('period 0') to 2005/2006 ('period 1')



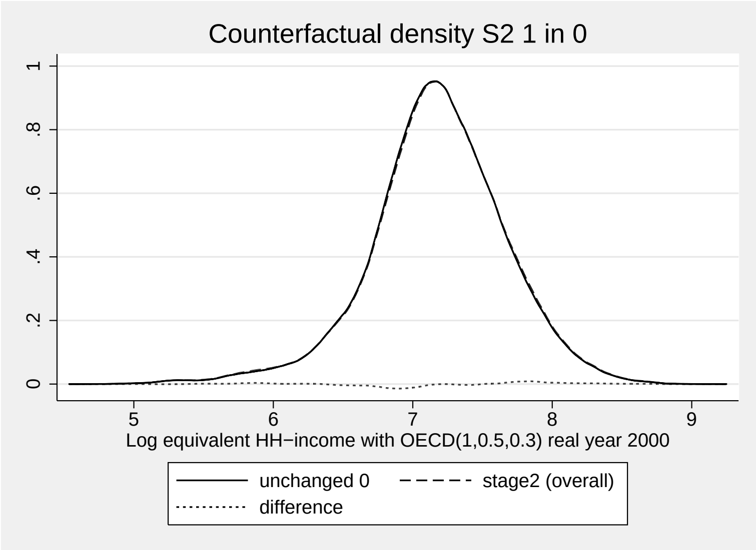
Source: GSOEP, own calculations

Figure 7 – Counterfactual income distribution if only the distribution of household types is changed (dashed line) vs. factual distribution (bold line).



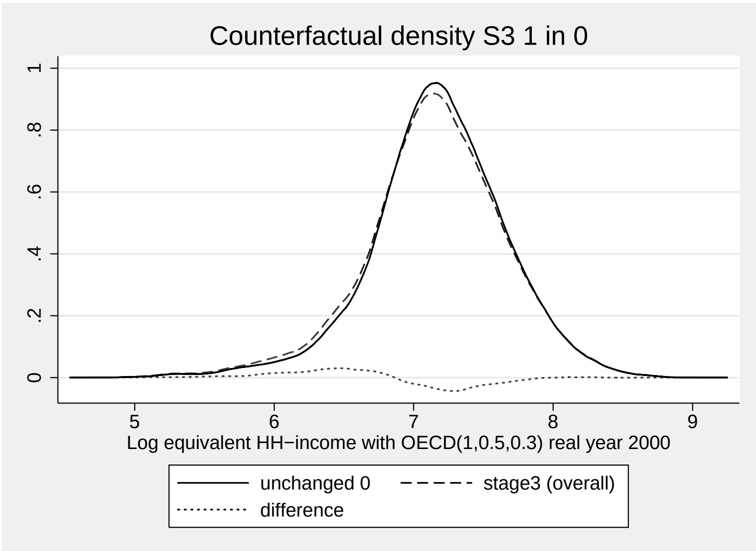
Source: GSOEP, own calculations

Figure 8 – Counterfactual income distribution if only the distribution of socio-economic attributes is changed (dashed line) vs. factual distribution (bold line).



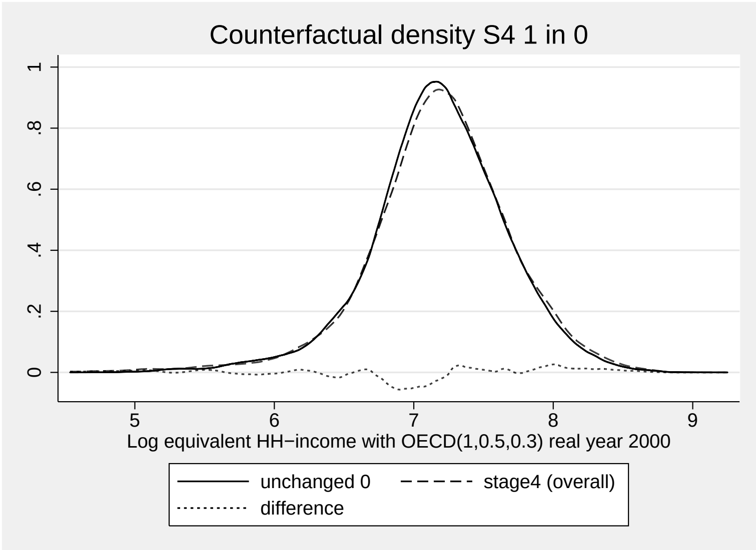
Source: GSOEP, own calculations

Figure 9 – Counterfactual income distribution if only conditional employment outcomes are changed (dashed line) vs. factual distribution (bold line).



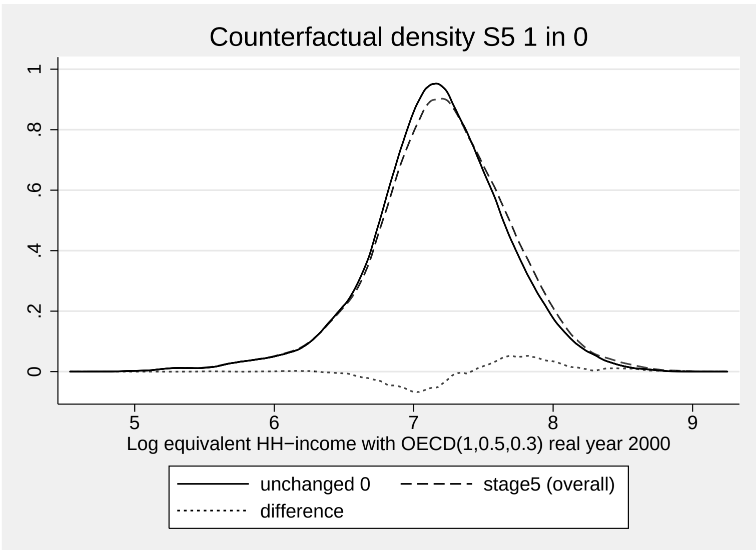
Source: GSOEP, own calculations

Figure 10 – Counterfactual income distribution if only market returns are changed (dashed line) vs. factual distribution (bold line).



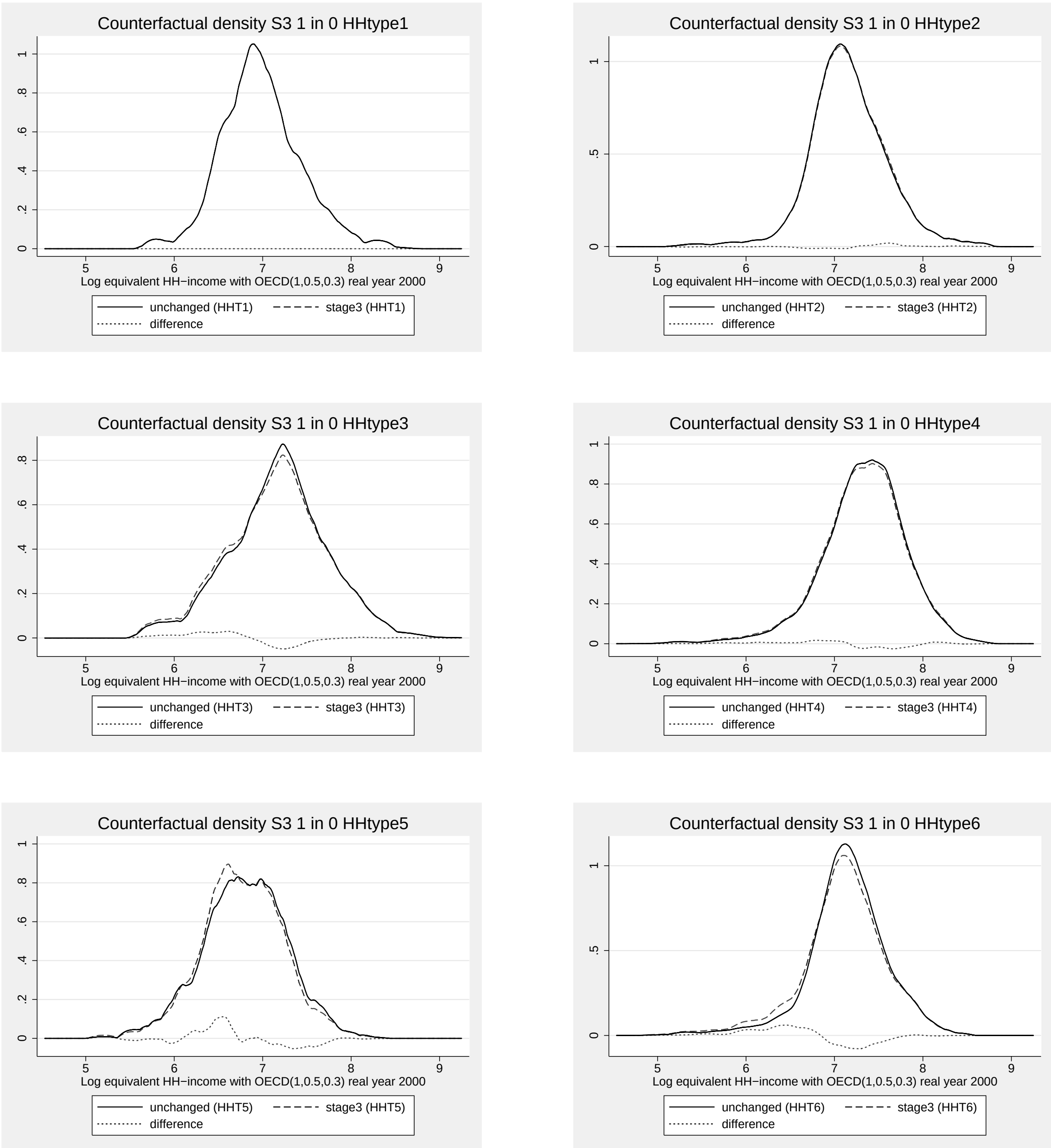
Source: GSOEP, own calculations

Figure 11 – Counterfactual income distribution if only the tax schedule is changed (dashed line) vs. factual distribution (bold line).



Source: GSOEP, own calculations

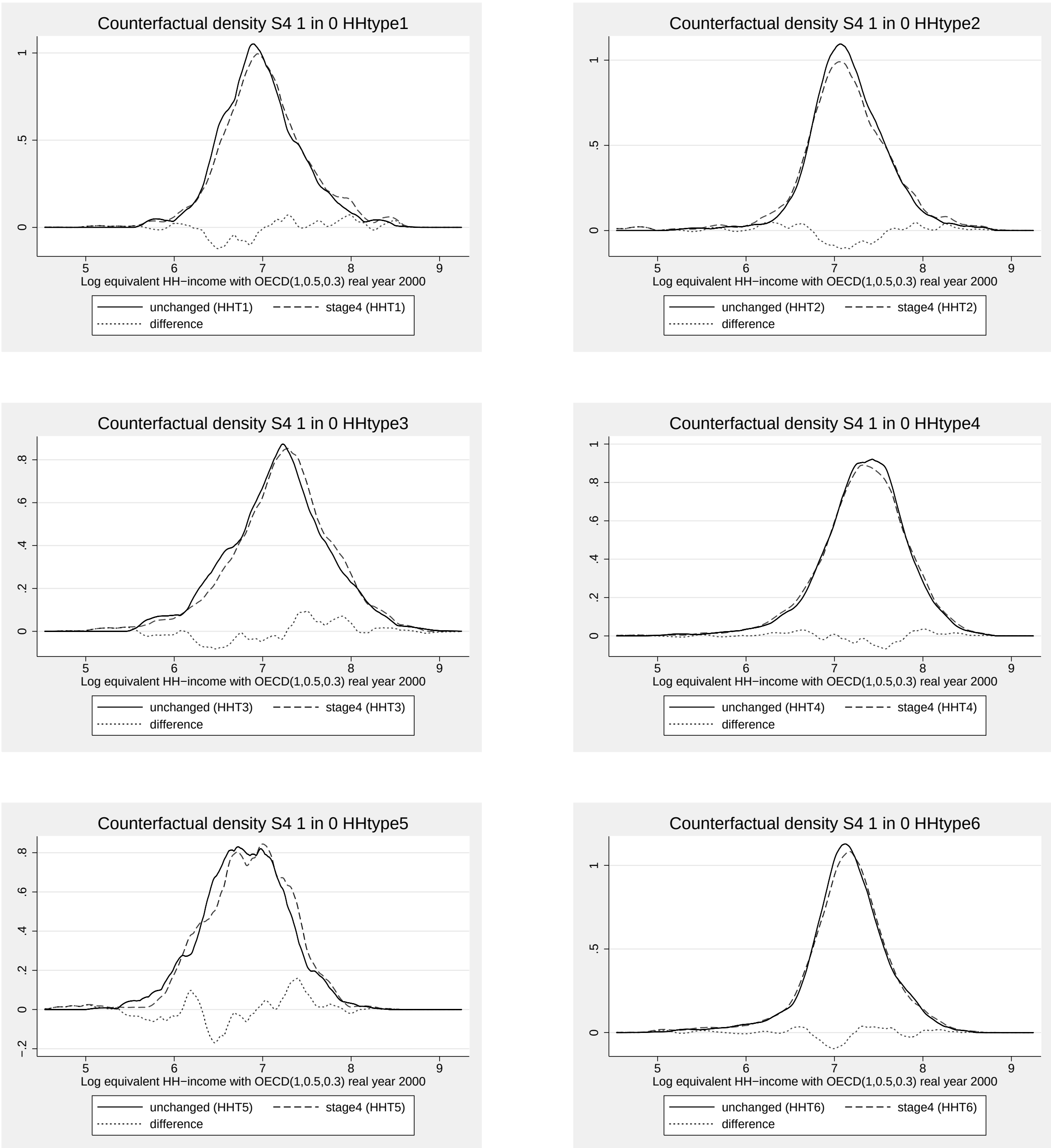
Figure 12 – Counterfactual income distribution if only conditional employment outcomes are changed (dashed line) vs. factual distribution (bold line), by household type.



HH-type 1 = Single pensioners, HH-type 2 = Multiple pensioners, HH-type 3 = Single adults w/o kids, HH-type 4 = Multiple adults w/o kids, HH-type 5 = Single adults with kids, HH-type 6 = Multiple adults with kids

Source: GSOEP, own calculations

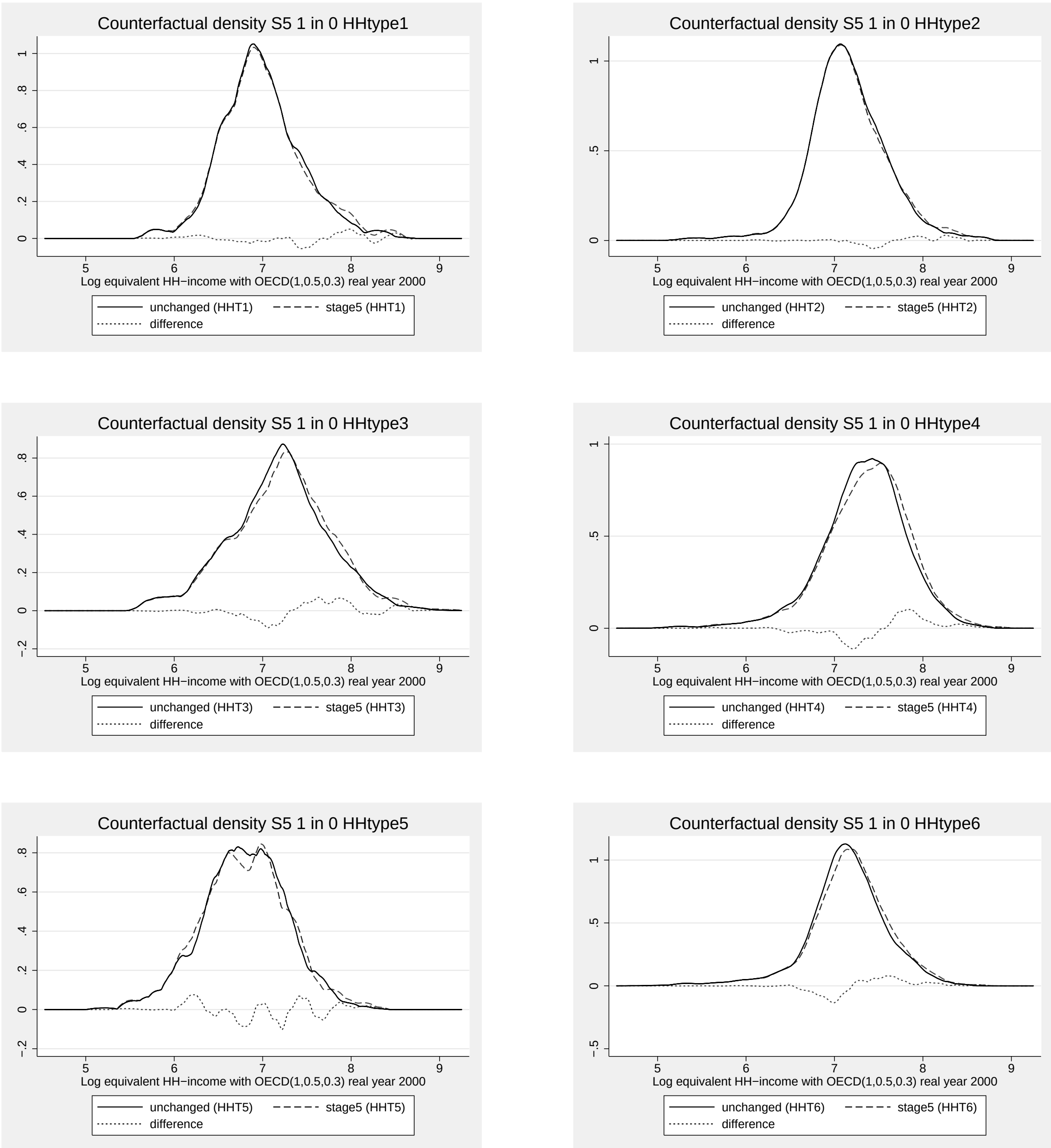
Figure 13 – Counterfactual income distribution if only market returns are changed (dashed line) vs. factual distribution (bold line), by household type.



HH-type 1 = Single pensioners, HH-type 2 = Multiple pensioners, HH-type 3 = Single adults w/o kids, HH-type 4 = Multiple adults w/o kids, HH-type 5 = Single adults with kids, HH-type 6 = Multiple adults with kids

Source: GSOEP, own calculations

Figure 14 – Counterfactual income distribution if only the tax schedule is changed (dashed line) vs. factual distribution (bold line).



HH-type 1 = Single pensioners, HH-type 2 = Multiple pensioners, HH-type 3 = Single adults w/o kids, HH-type 4 = Multiple adults w/o kids, HH-type 5 = Single adults with kids, HH-type 6 = Multiple adults with kids

Source: GSOEP, own calculations

9 Tables

Table 1 – Changes in the German tax schedule

Year	Basic Allowance	Min. Marginal Tax Rate	End of Progression Zone	Max. Marginal Tax Rate
1999	6,681 EUR	23.9%	61,376 EUR	53%
2000	6,902 EUR	22.9%	58,643 EUR	51%
2001	7,206 EUR	19.9%	54,998 EUR	48.5%
2002/2003	7,235 EUR	19.9%	55,008 EUR	48.8%
2004	7,664 EUR	16.0%	52,152 EUR	45%
2005/2006	7,664 EUR	15.0%	52,152 EUR	42%

Source: German Federal Ministry of Finance

Table 2 – Inequality and poverty indices

Index	Abbr.	Estimator
Quantile ratio 90/10	p9010	$\widehat{\text{p9010}}(\hat{f}) = \hat{q}_{90}/\hat{q}_{10}$
Quantile ratio 90/50	p9050	$\widehat{\text{p9050}}(\hat{f}) = \hat{q}_{90}/\hat{q}_{50}$
Quantile ratio 75/25	p7525	$\widehat{\text{p5010}}(\hat{f}) = \hat{q}_{75}/\hat{q}_{25}$
Quantile ratio 50/10	p5010	$\widehat{\text{p5010}}(\hat{f}) = \hat{q}_{50}/\hat{q}_{10}$
Coefficient of variance	CV	$\widehat{\text{cv}}(\hat{f}) = sd(\hat{f})/\mu(\hat{f})$
Theil's measure	Theil	$\widehat{\text{theil}}(\hat{f}) = \int \frac{y}{\mu(\hat{f})} \log\left(\frac{y}{\mu(\hat{f})}\right) \hat{f}(y)\,dy$
Mean log deviation	MLD	$\widehat{\text{mld}}(\hat{f}) = -\int \log\left(\frac{y}{\mu(\hat{f})}\right) \hat{f}(y)\,dy$
Gini coefficient	Gini	$\widehat{\text{gini}}(\hat{f}) = \int y(2\hat{F}(y)-1)\hat{f}(y)\,dy$
Forster, Greer, Thorbecke	FGT(α)	$\widehat{\text{FGT}}(\hat{f},\alpha) = \int_{\{y<p(\hat{f})\}} \left(\frac{p(\hat{f})-y}{p(\hat{f})}\right)^{\alpha} \hat{f}(y)dy,\ \alpha \geq 0$

Note: $FGT(0)$ = poverty headcount, $FGT(1)$ = poverty gap measure, $p(\hat{f})$ = poverty line

Table 3 – Ceteris paribus effects as percentage of overall inequality increase

	Percentage of the overall inequality increase explained by ceteris paribus change of									
	Household Structure (1)		**Socio-economic attributes (2)**		**Employment outcomes (3)**		**Return on attributes (4)**		**Tax system (5)**	
p5010	7.39	(2.82)	4.42	(3.26)	36.09	(9.10)	17.84	(11.70)	22.37	(5.81)
p7525	6.80	(2.29)	3.39	(2.78)	24.02	(5.14)	13.65	(9.15)	27.50	(6.26)
p9010	8.93	(2.45)	6.93	(2.92)	34.54	(7.23)	23.12	(9.54)	25.18	(6.06)
p9050	13.47	(4.92)	13.47	(7.26)	33.87	(13.21)	37.28	(21.22)	33.87	(20.36)
Cv	8.20	(2.21)	5.22	(2.88)	19.28	(4.14)	20.76	(6.99)	26.02	(5.98)
Theil	8.33	(2.24)	6.54	(2.79)	22.99	(4.50)	27.51	(8.56)	24.21	(5.51)
Mld	3.90	(2.23)	3.82	(2.96)	23.03	(5.53)	21.56	(9.98)	26.52	(5.30)
Gini	5.31	(2.44)	3.33	(3.08)	22.95	(5.27)	13.93	(9.52)	28.24	(5.59)
Fgt0	7.72	(2.58)	4.19	(2.76)	31.58	(6.87)	20.83	(10.67)	21.25	(5.28)
Fgt1	4.03	(2.54)	7.33	(3.81)	35.14	(9.11)	30.96	(13.93)	21.80	(5.88)

Source: GSOEP, own calculations. The numbers in parentheses are bootstrap standard errors which correctly take into account the longitudinal sample design and the clustering of individuals in households.

Table 4 – Exact decomposition of inequality increase

	Results of sequential decomposition attributable to										
	Household Structure (1)		**Socio-economic attributes (2)**		**Employment outcomes (3)**		**Return on attributes (4)**		**Tax system (5)**		**Residual**
p5010	7.39	(2.82)	5.96	(3.26)	30.48	(8.76)	23.62	(16.49)	24.30	(5.64)	8.25
p7525	6.80	(2.29)	3.42	(2.79)	22.54	(5.33)	14.15	(10.66)	19.08	(3.22)	33.30
p9010	8.93	(2.45)	6.04	(2.96)	30.16	(7.11)	29.61	(13.03)	20.59	(3.77)	4.67
p9050	13.47	(4.92)	6.77	(7.24)	30.80	(12.70)	41.89	(24.98)	10.62	(8.59)	-3.55
Cv	8.20	(2.21)	4.66	(3.04)	16.96	(4.19)	22.76	(7.80)	20.92	(5.24)	26.50
Theil	8.33	(2.24)	5.07	(2.70)	19.92	(4.54)	31.41	(9.80)	19.88	(4.66)	15.36
Mld	3.90	(2.23)	5.81	(2.70)	23.30	(5.43)	28.85	(12.1)	19.64	(4.65)	18.47
Gini	5.31	(2.44)	5.54	(2.79)	23.17	(4.99)	17.71	(10.91)	17.77	(4.71)	30.48
Fgt0	7.72	(2.58)	5.34	(2.73)	26.67	(6.64)	20.23	(12.34)	19.81	(3.97)	20.24
Fgt1	4.03	(2.54)	8.21	(3.79)	30.40	(9.07)	39.38	(17.19)	23.09	(5.01)	-5.11

Source: GSOEP, own calculations. The numbers in parentheses are bootstrap standard errors which correctly take into account the longitudinal sample design and the clustering of individuals in households.

10 Appendix

Table 5 – Results from all 120 possible sequential decompositions

	Marginal relative change attributable to				
	Household Structure (1)	**Socio-economic attributes (2)**	**Employment outcomes (3)**	**Return on attributes (4)**	**Tax system (5)**
			c5010 (Total Change = .231)		
Primary order	7.39	5.96	30.48	23.62	24.30
Mean	4.57	8.16	37.08	20.11	23.37
Sd	2.57	2.50	2.91	2.37	1.26
Min	1.61	4.54	31.64	15.67	21.37
Max	7.77	11.62	43.90	23.98	25.60
			c7525 (Total Change = .200)		
Primary order	6.80	3.42	22.54	14.15	19.80
Mean	6.27	4.51	24.00	9.12	22.79
Sd	1.18	0.88	1.25	5.12	5.21
Min	3.48	3.42	20.70	1.80	15.69
Max	7.25	5.44	26.65	15.93	30.14
			c9010 (Total Change = .623)		
Primary order	8.93	6.04	30.16	29.61	20.59
Mean	6.61	9.09	35.61	21.12	22.90
Sd	2.02	1.94	2.64	3.91	3.93
Min	4.21	6.04	29.82	13.80	16.69
Max	9.47	12.15	41.27	29.61	30.50
			c9050 (Total Change = .101)		
Primary order	13.47	6.77	30.80	41.89	10.62
Mean	11.02	11.02	32.47	23.44	22.06
Sd	2.31	2.32	3.39	12.99	12.59
Min	3.31	3.31	26.31	9.93	6.63
Max	13.51	13.51	36.91	41.89	40.45
			CV (Total Change = .085)		
Primary order	8.20	4.66	16.96	22.76	20.92
Mean	7.13	5.34	19.07	18.53	23.44
Sd	0.89	0.66	0.83	2.82	2.82
Min	5.43	4.07	16.96	14.36	19.81
Max	8.54	6.23	20.40	23.70	27.25
			Theil (Total Change = .035)		
Primary order	8.33	5.07	19.93	31.42	19.89
Mean	6.60	6.45	23.58	26.15	21.86
Sd	1.21	0.79	1.77	2.77	2.63
Min	4.73	4.95	19.93	21.42	18.13
Max	8.33	7.55	26.87	32.07	26.08
			Mld (Total Change = .034)		
Primary order	3.90	5.82	23.31	28.86	19.64
Mean	4.62	7.02	26.10	19.80	23.99
Sd	1.09	1.87	2.27	5.16	5.18

Min	3.49	3.82	23.03	12.64	17.86
Max	6.85	9.01	30.29	29.14	32.15
		Gini (Total Change = .035)			
Primary order	5.32	5.54	23.17	17.72	17.77
Mean	6.21	6.25	24.03	9.07	23.96
Sd	0.96	1.67	1.07	6.77	6.60
Min	5.11	3.33	22.95	1.14	16.69
Max	8.47	8.06	26.57	18.28	33.05
		Fgt0 (Total Change = .043)			
Primary order	7.72	5.34	26.67	20.23	19.81
Mean	5.44	6.22	28.78	20.40	20.41
Sd	1.92	2.21	1.85	2.12	0.99
Min	2.03	2.69	25.13	15.95	18.54
Max	7.87	9.35	32.56	24.63	21.92
		Fgt1 (Total Change = .138)			
Primary order	4.03	8.21	30.40	39.38	23.09
Mean	1.49	9.42	38.05	35.23	22.55
Sd	2.56	1.85	4.51	4.04	0.93
Min	-2.38	6.99	30.40	29.94	20.09
Max	4.44	13.00	46.49	40.94	24.23

Source: GSOEP, own calculations

Table 6 – Variable names

hhemp	household employment outcome
	(0 = 'no ft/no pt', 1 = 'no ft/at least 1 pt',
	2 = '1 ft/no pt', 3 = '1 ft/at least 1 pt',
	4 = 'at least 2 ft')
hhemp_d0-hhemp_d4	household employment category dummies
hhadult	number of adults in the household
f_ad_fem	fraction of adult hh-members female
f_ad_for	fraction of adult hh-members foreigner
f_ad_mar	fraction of adult hh-members married
f_ad_uni	fraction of adult hh-members with university degree
f_ad_abv	fraction of adult hh-members with high school degree
	and/or vocational training
f_ad_dis	fraction of adult hh-members with disabilities
f_ad_03	fraction of hh-members aged 0-3 years
f_ad_11	fraction of hh-members aged 4-11 years
f_ad_17	fraction of hh-members aged 12-17 years
f_ad_30	fraction of hh-members aged 18-30 years
f_ad_50	fraction of hh-members aged 31-50 years
f_ad_65	fraction of hh-members aged 51-65 years
f_ad_99	fraction of hh-members aged 65 years or older
e	East Germany

Table 7 – Ordinal logit models

	Household Type 2		*Household Type 3*		*Household Type 4*		*Household Type 5*		*Household Type 6*	
Variable	1999/2000	2005/2006	1999/2000	2005/2006	1999/2000	2005/2006	1999/2000	2005/2006	1999/2000	2005/2006
number of adults	1.875	1.745			.420	.499			1.063	1.329
	(.317)	(.355)			(.070)	(.083)			(.110)	(.125)
f_ad_fem			-.159	-.550	-1.136	-.660	-2.183	-1.410		
			(.140)	(.142)	(.296)	(.322)	(.538)	(.776)		
f_ad_age50			1.079	1.037	.567	.771	-.711	.999	.756	.759
			(.171)	(.185)	(.177)	(.211)	(.385)	(.485)	(.183)	(.266)
f_ad_age64			-.533	-.168	-1.586	-.773	-2.589	.579	-.349	-.586
			(.173)	(.179)	(.191)	(.200)	(1.433)	(.728)	(.364)	(.410)
f_ad_age99	-2.629	-3.736			-3.515	-2.377				
	(.563)	(.616)			(.448)	(.344)				
f_ad_uni	1.268	1.040	1.153	1.794	2.068	2.053	1.969	1.814	1.523	2.426
	(.387)	(.340)	(.250)	(.260)	(.198)	(.233)	(.404)	(.523)	(.246)	(.274)
f_ad_abv			.560	.818	1.435	1.785	.349	.955	1.428	2.014
			(.196)	(.211)	(.151)	(.175)	(.302)	(.349)	(.204)	(.249)
f_ad_dis	-.900	-.992	-1.003	-.687	-1.042	-1.177			-1.067	-1.546
	(.485)	(.423)	(.224)	(.210)	(.183)	(.217)			(.404)	(.366)
f_ad_mar			.411	.476	.094	.278	-.472	-.123	.858	.882
			(.274)	(.245)	(.137)	(.143)	(.286)	(.312)	(.231)	(.277)
f_ad_for			-.312	-.962	.571	-.034	1.368	.797	-.146	-.415
			(.336)	(.354)	(.205)	(.231)	(.564)	(.546)	(.212)	(.219)
f_ch_11							2.523	2.326	.447	.931
							(.728)	(.946)	(.155)	(.161)
f_ch_17							3.639	2.731	1.292	1.123
							(.756)	(.976)	(.154)	(.167)
e	-0.810	-0.697	-.864	-.688	-.231	-.397	-.086	-.664	.489	.395
	(.288)	(.261)	(.146)	(.152)	(.093)	(.103)	(.295)	(.349)	(.127)	(.167)
/cut1	2.439	2.179	-.367	.086	-1.014	.078	.940	2.685	2.374	4.104
/cut2	2.931	2.682	-.080	.389	-.558	.612	.938	3.951	2.978	4.755
/cut3					1.138	2.348			5.973	7.426
/cut4					1.819	3.158			7.352	9.084
Pseudo R^2	0.167	0.170	0.123	0.119	0.113	0.086	0.181	0.115	0.086	0.103
Number of clusters	1349	1700	2127	2122	3744	3350	422	447	3471	2716

Source: GSOEP, own calculations. Standard errors account for clustering of observations in households.

Table 8 – Regression of market incomes on household characteristics (# denotes interaction effects)

	Household Type 1			
Variable	1999/2000		2005/2006	
hhemp_d1	2.410	(0.478)	2.268	(0.239)
hhemp_d2	1.960	(0.481)	2.250	(0.400)
f_ad_fem	0.234	(0.232)	-0.089	(0.189)
f_ad_abv	0.716	(0.187)	0.515	(0.167)
f_ad_uni	1.617	(0.378)	1.333	(0.300)
e	-1.135	(0.217)	-1.310	(0.173)
_cons	3.291	(0.271)	3.872	(0.205)
R^2	0.131		0.147	
Number of clusters	1339		1868	
	Household Type 2			
Variable	1999/2000		2005/2006	
hhemp_d1	2.133	(0.199)	1.202	(0.191)
hhemp_d2	2.866	(0.162)	2.287	(0.151)
hhemp_d3	3.074	(0.206)	2.873	(0.276)
hhemp_d4	3.790	(0.222)	2.524	(0.273)
hhadult	0.315	(0.215)	0.026	(0.174)
f_ad_fem	0.232	(0.834)	-0.082	(0.664)
f_ad_for	0.228	(0.423)	0.143	(0.320)
f_ad_mar	0.633	(0.369)	-0.216	(0.409)
f_ad_age99	-0.187	(0.329)	-1.116	(0.299)
f_ad_uni	1.491	(0.295)	1.858	(0.280)
f_ad_abv	0.239	(0.235)	0.619	(0.250)
f_ad_dis	0.336	(0.210)	0.248	(0.210)
e	-1.305	(0.156)	-1.531	(0.159)
_cons	2.751	(0.998)	5.213	(0.895)
R^2	0.286		0.295	
Number of clusters	1264		1622	
	Household Type 3			
Variable	1999/2000		2005/2006	
hhemp_d1	-0.326	(0.363)	0.750	(0.366)
hhemp_d2	1.000	(0.248)	2.104	(0.299)
f_ad_fem	-0.551	(0.198)	0.192	(0.196)
f_ad_for	-0.558	(0.505)	1.048	(0.289)
f_ad_age50	-1.307	(0.369)	-0.744	(0.273)
f_ad_age64	-1.266	(0.216)	-1.286	(0.256)
f_ad_uni	0.478	(0.397)	1.058	(0.342)
f_ad_abv	-0.023	(0.243)	0.578	(0.271)
f_ad_dis	-0.938	(0.224)	-0.500	(0.264)
e	-0.508	(0.197)	-0.563	(0.198)
hhemp#				
c.f_ad_fem				
1	0.505	(0.219)	0.049	(0.273)
2	0.464	(0.198)	-0.313	(0.198)
hhemp#				

c.f_ad_for				
1	0.129	(0.537)	-1.851	(0.402)
2	0.433	(0.515)	-0.997	(0.298)
hhemp#				
c.f_ad_age50				
1	1.841	(0.383)	1.155	(0.312)
2	1.511	(0.367)	0.906	(0.272)
hhemp#				
c.f_ad_age64				
1	1.792	(0.256)	1.628	(0.298)
2	1.498	(0.222)	1.539	(0.262)
hhemp#				
c.f_ad_uni				
1	0.634	(0.486)	0.004	(0.394)
2	-0.051	(0.404)	-0.593	(0.355)
hhemp#				
c.f_ad_abv				
1	0.617	(0.372)	-0.162	(0.333)
2	0.176	(0.255)	-0.415	(0.282)
hhemp#				
c.f_ad_dis				
1	0.667	(0.290)	0.786	(0.342)
2	0.897	(0.235)	0.557	(0.271)
hhemp#c.e				
1	0.196	(0.240)	0.154	(0.290)
2	0.137	(0.203)	0.182	(0.202)
_cons	6.519	(0.232)	5.427	(0.290)
R^2	0.609		0.572	
Number of clusters	2002		1980	

Household Type 4

Variable	1999/2000		2005/2006	
hhemp_d1	0.123	(0.887)	0.167	(0.643)
hhemp_d2	0.061	(0.749)	0.308	(0.596)
hhemp_d3	0.293	(0.759)	0.658	(0.596)
hhemp_d4	0.601	(0.746)	0.754	(0.583)
f_ad_age50	-1.529	(0.509)	-1.395	(0.452)
f_ad_age64	-2.306	(0.401)	-0.829	(0.316)
f_ad_age99	-3.387	(0.829)	-1.211	(0.766)
f_ad_uni	1.990	(0.343)	1.166	(0.440)
e	-1.268	(0.212)	-0.858	(0.197)
hhemp#				
c.hhadult				
0	-0.194	(0.242)	-0.173	(0.134)
1	-0.160	(0.108)	-0.215	(0.074)
2	-0.082	(0.033)	-0.137	(0.039)
3	-0.107	(0.038)	-0.141	(0.031)
4	-0.124	(0.023)	-0.144	(0.025)
hhemp#				
c.f_ad_fem				
0	-0.229	(0.570)	-0.610	(0.682)

1	-0.737	(0.631)	-0.036	(0.293)
2	-0.029	(0.149)	-0.090	(0.155)
3	0.021	(0.180)	0.127	(0.183)
4	-0.051	(0.122)	0.077	(0.187)
hhemp#				
c.f_ad_for				
0	-1.245	(0.495)	-1.041	(0.570)
1	-0.157	(0.168)	-0.577	(0.365)
2	-0.233	(0.087)	-0.202	(0.112)
3	-0.459	(0.180)	-0.416	(0.287)
4	-0.216	(0.069)	-0.183	(0.090)
hhemp#				
c.f_ad_mar				
0	-0.195	(0.330)	-0.206	(0.276)
1	0.164	(0.161)	0.202	(0.189)
2	-0.059	(0.115)	0.134	(0.076)
3	0.080	(0.060)	0.238	(0.106)
4	0.015	(0.039)	0.043	(0.049)
hhemp#				
c.f_ad_age50				
1	1.431	(0.606)	1.237	(0.511)
2	1.719	(0.522)	1.597	(0.459)
3	1.710	(0.520)	1.410	(0.471)
4	1.699	(0.512)	1.703	(0.455)
hhemp#				
c.f_ad_age64				
1	1.892	(0.478)	0.602	(0.407)
2	2.493	(0.417)	1.129	(0.332)
3	2.480	(0.414)	0.890	(0.344)
4	2.435	(0.405)	1.102	(0.323)
hhemp#				
c.f_ad_age99				
1	2.557	(1.055)	0.986	(0.862)
2	2.371	(0.985)	0.727	(0.804)
3	3.054	(0.858)	0.348	(0.960)
4	2.817	(0.904)	0.849	(0.844)
hhemp#				
c.f_ad_uni				
1	-0.952	(0.378)	-0.211	(0.468)
2	-1.285	(0.347)	-0.636	(0.456)
3	-1.480	(0.353)	-0.660	(0.449)
4	-1.540	(0.346)	-0.658	(0.444)
hhemp#				
c.f_ad_dis				
0	0.233	(0.338)	-1.081	(0.320)
1	-0.054	(0.200)	-0.229	(0.254)
2	-0.422	(0.209)	0.016	(0.105)
3	-0.047	(0.139)	0.196	(0.167)
4	-0.010	(0.111)	0.211	(0.093)
hhemp#c.e				
1	0.968	(0.262)	0.513	(0.237)

2	0.814	(0.216)	0.417	(0.209)
3	0.864	(0.218)	0.535	(0.205)
4	0.968	(0.214)	0.647	(0.203)
_cons	7.525	(0.738)	7.171	(0.571)
R^2	0.613		0.535	
Number of clusters	3659		3247	

Household Type 5				
Variable	1999/2000		2005/2006	
hhemp_d1	1.009	(0.696)	1.005	(0.654)
hhemp_d2	1.676	(0.524)	2.426	(0.658)
f_ad_age50	0.083	(0.381)	-0.063	(0.486)
f_ad_age64	-1.634	(1.816)	-1.071	(1.150)
f_ad_uni	-1.401	(0.638)	0.117	(0.459)
f_ad_dis	0.391	(0.347)	-0.681	(0.286)
f_ch_age11	-0.581	(0.595)	1.168	(0.830)
f_ch_age17	0.515	(0.432)	0.549	(0.902)
e	-0.191	(0.347)	-0.740	(0.565)
hhemp#				
c.f_ad_age50				
1	0.458	(0.494)	0.777	(0.563)
2	0.140	(0.415)	0.222	(0.553)
hhemp#				
c.f_ad_age64				
1	2.735	(1.854)	1.730	(1.224)
2	2.483	(1.849)	1.132	(1.209)
hhemp#				
c.f_ad_uni				
1	1.043	(0.793)	0.384	(0.494)
2	1.912	(0.651)	0.436	(0.482)
hhemp#				
c.f_ch_age11				
1	0.568	(0.701)	-0.673	(0.870)
2	0.843	(0.613)	-0.933	(0.892)
hhemp#				
c.f_ch_age17				
1	-0.685	(0.582)	-0.128	(0.935)
2	-0.556	(0.472)	-0.404	(0.960)
hhemp#c.e				
1	0.411	(0.406)	0.753	(0.597)
2	-0.191	(0.373)	0.244	(0.587)
_cons	5.295	(0.510)	4.610	(0.571)
R^2	0.449		0.427	
Number of clusters	387		407	

Household Type 6				
Variable	1999/2000		2005/2006	
hhemp_d1	0.553	(1.283)	1.694	(1.690)
hhemp_d2	0.782	(1.165)	3.296	(1.451)
hhemp_d3	0.848	(1.132)	4.106	(1.490)

hhemp_d4	1.144	(1.120)	4.941	(1.444)
hhadult	-0.008	(0.325)	0.233	(0.254)
f_ad_fem	-2.396	(1.107)	1.099	(1.916)
f_ad_for	-0.098	(0.430)	-1.097	(0.472)
f_ad_mar	-0.032	(0.431)	-0.029	(0.379)
f_ad_age50	-0.056	(0.336)	0.228	(0.398)
f_ad_age64	-1.831	(0.795)	-0.435	(0.738)
f_ad_uni	-0.284	(1.000)	1.316	(0.668)
f_ad_abv	0.340	(0.453)	0.593	(0.367)
f_ch_age11	-0.083	(0.420)	0.066	(0.407)
f_ch_age17	0.286	(0.479)	0.425	(0.442)
e	-0.352	(0.035)	-0.301	(0.053)
hhemp#				
c.hhadult				
1	-0.149	(0.364)	-0.019	(0.289)
2	-0.111	(0.331)	-0.176	(0.259)
3	-0.034	(0.327)	-0.258	(0.275)
4	-0.124	(0.326)	-0.475	(0.258)
hhemp#				
c.f_ad_fem				
1	1.601	(1.186)	-0.557	(1.967)
2	1.985	(1.166)	-1.505	(1.937)
3	2.473	(1.125)	-1.815	(1.942)
4	2.249	(1.117)	-1.488	(1.942)
hhemp#				
c.f_ad_for				
1	0.161	(0.512)	0.492	(0.477)
2	0.009	(0.436)	0.856	(0.482)
3	-0.166	(0.436)	0.450	(0.516)
4	0.057	(0.434)	1.141	(0.485)
hhemp#				
c.f_ad_mar				
1	-0.875	(0.484)	0.025	(0.476)
2	0.000	(0.450)	0.235	(0.401)
3	0.050	(0.437)	0.032	(0.394)
4	-0.026	(0.436)	0.130	(0.398)
hhemp#				
c.f_ad_age50				
1	0.344	(0.438)	0.178	(0.591)
2	0.295	(0.347)	0.085	(0.414)
3	0.137	(0.349)	0.250	(0.420)
4	0.221	(0.347)	-0.386	(0.430)
hhemp#				
c.f_ad_age64				
1	1.976	(0.869)	1.002	(0.836)
2	1.980	(0.810)	1.059	(0.771)
3	2.064	(0.807)	1.011	(0.759)
4	1.998	(0.809)	0.213	(0.766)
hhemp#				
c.f_ad_uni				
1	1.701	(1.070)	0.114	(0.705)

2	1.119	(1.005)	-0.595	(0.683)
3	1.049	(1.006)	-0.518	(0.695)
4	0.908	(1.006)	-0.625	(0.685)
hhemp#				
c.f_ad_abv				
1	0.445	(0.520)	0.211	(0.458)
2	-0.061	(0.461)	-0.299	(0.387)
3	-0.107	(0.461)	-0.429	(0.414)
4	-0.145	(0.460)	-0.458	(0.389)
hhemp#				
c.f_ad_dis				
0	-0.971	(0.483)	-0.769	(0.772)
1	-0.735	(0.353)	0.451	(0.274)
2	-0.061	(0.104)	0.019	(0.144)
3	-0.239	(0.098)	-0.149	(0.154)
4	-0.132	(0.133)	-0.312	(0.233)
hhemp#				
c.f_ch_age11				
1	0.364	(0.472)	-0.281	(0.503)
2	0.088	(0.434)	-0.021	(0.412)
3	-0.049	(0.429)	-0.171	(0.413)
4	0.149	(0.431)	-0.018	(0.421)
hhemp#				
c.f_ch_age17				
1	0.094	(0.516)	-0.602	(0.515
2	-0.304	(0.490)	-0.472	(0.458)
3	-0.310	(0.487)	-0.463	(0.447)
4	-0.121	(0.490)	-0.324	(0.455)
_cons	6.514	(1.111)	3.323	(1.416)
R^2	0.446		0.564	
Number of clusters	3442		2676	

Standard errors account for clustering of observations in households.

Source: GSOEP, own calculations.