

Income Distribution and the Business Cycle in Germany: A Semiparametric Approach*

Andos JUHÁSZ**

Abstract: The question of how the distribution of income is influenced by the state of the economy is important for understanding the economic mechanisms linking micro- and macro-level variables. There is no generally applicable theory on how the distribution of household incomes is influenced by changed macro-economic conditions. This paper adapts an empirical approach in order to investigate this relationship by employing a semi-parametric double-index model without restrictions on the shape of the link function between indices of household- and macro-level variables. We conclude that the link function is non-trivial and non-additive, meaning that the influence of macroeconomic conditions (i.e. GDP, inflation, government expenditure or unemployment) is likely to vary with household characteristics. The influence found is small, but in parts significant, drawing a comprehensive picture of potential link-mechanisms.

* This is an open-access article distributed under the terms of the Creative Commons license Attribution-Noncommercial-NoDerivs 3.0 Unported. Distribution and reproduction are permitted, provided that the original authors and JPAID as the source are credited. The author is grateful to Martin Biewen for his comments and suggestions. All errors are the author's own. The data used in this paper (SOEP v27, 1996-2010) was made available by the German Socio-Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW), Berlin and the GENESIS Data Base of the German Federal Statistical Office, Wiesbaden.

** Department of Economics, University of Tübingen, Tübingen, Germany.
Email: andos.juhasz@uni-tuebingen.de

Keywords: Income Distribution, Business Cycle, Double Index Model, Semiparametric Estimation

The question of how the level of incomes and their dispersion behave in different macro-economic conditions (as given by the GDP, inflation, government expenditure and unemployment among others) goes back at least to the 1950s, when Kuznets (1955) stated an inverse u-shaped relation between the inequality of incomes and the state of economy. He argued that emerging economies that transform from agricultural societies to industrial ones become increasingly unequal at first, and, when the transformation shifts towards a service and knowledge oriented society, human capital becomes more important relative to physical capital, and the inequality of incomes decreases. Since then, many analyses have been conducted on this question and a large theoretical and empirical literature has developed. While the theoretical work generally identified a diversity of functional channels through which income distribution and the business cycle are connected to each other, the overall effect¹ remained unclear and most likely depends on the economy being analyzed.²

So far, the majority of the literature has used summary measures, i.e. different quantiles (e.g. Blinder & Esaki 1978; Jäntti 1994), point measures of inequality, or shape parameters under distributional assumptions (e.g. Thurow 1970; Mount & Salem 1974; Jäntti & Jenkins 2010) to establish a connection between macro-economic states and the income distribution. There was a strong development in the econometric strategies involved, starting with (system) regression approaches (e.g. Blinder

¹ For the sake of linguistic diversity, when we talk about “effects” or “impact” in this paper, we do not necessarily use it in a causal manner, but also with the background of descriptive evidence as an expression for general association. This especially applies to the “results” section.

² See e.g. Easterly and Fischer (2001); Berthold et al. (2010).

& Esaki 1978) to time series settings involving cointegration of the potentially non-stationary time series of macro-economic variables and inequality measures (e.g. Parker 1998; Jäntti & Jenkins 2010). However, these papers discarded detailed information on the income distribution when they collapsed the income distribution into summary statistics (e.g. of inequality).

When household panel data is available, we have the option to make use of detailed panel information to analyze the macro-economic influence on incomes in tandem with micro-economic factors. While numerous papers on the relationship between household level variables and the income structure exist, it also makes sense to follow a hybrid approach that uses both micro- and macro-level information and their possible interactions to analyze their impact on the distribution of incomes at the micro-level. The assumption that the effect of macro-economic variables may be household specific is not new. For example, Blank (1989) found that elderly households and households with a female head were less responsive to the business cycle than others.

In this paper, we adapt a hybrid-level approach following the strategy of Farré and Vella (2008), who use Spanish data to analyze the impact of micro- and macro-level variables jointly with a semi-parametric double index model (Ichimura & Lee 1991). With several technical and data-related adaptations, we apply the method to German data between 1996 and 2010 extracted from the German Socio-Economic Panel (SOEP), to analyze potential mechanisms between GDP, inflation, government expenditure and unemployment on the one hand and income distribution on the other.

Making use of cross-sectional information, the double-index model allows for arbitrary functional relationships between indices of micro- and macro-level variables to describe their joint influence on the income distribution and allows us to employ a counterfactual approach to simulate different macro-economic states and to analyze their impacts on the distribution of incomes. We find that the impact of changing macro-economic conditions on the distribution of incomes is small, but

in parts significant, drawing a comprehensible picture of potential mechanisms. In particular, we find that the level of government expenditure has a positive impact on income levels, while the level of inflation is inversely related to it. We find further that both a negative GDP shock and a lower inflation are in line with lower inequality, and that higher government expenditure is also associated with lower inequality. As regards unemployment, a significant association with the income distribution could not be given.

The rest of this paper is organized as follows. In the next, second section, we summarize the theoretical role of macro-economic relationships regarding the shape of the income distribution and give a summary of existing econometric approaches. Third, we describe the double-index model in detail, then give an overview of the data to be used for estimation. Following this, we report our results in detail on the estimated relationship between income distribution and the business cycle and discuss some counterfactual experiments. The sixth and final section concludes.

Literature

The theoretical role of macro-economic conditions

In this subsection, we partly follow Parker (1998) and Berthold et al. (2010), who summarize theoretical impacts of macro-economic variables on income distribution. In particular, our research concentrates on four factors to describe the state of the macro-economy: GDP, inflation, government final consumption expenditure,³ and unemployment.

³ “General government final consumption expenditure consists of expenditure incurred by government in its production of non-market final goods and services [...] and market goods and services provided as social transfers in kind” (OECD 2010).

As mentioned above, general theories on a possible relationship between the distribution of incomes and macro-economic circumstances go back at least to the 1950s. In his 1955 paper, Kuznets states an inverse u-shaped relationship between income inequality and a country's general level of wealth, e.g. as measured by the gross domestic product. He argues that emerging economies transforming from agricultural to industrial societies become increasingly unequal as profits belong to those who provide capital for the transformation. When the transformation shifts towards a service and knowledge oriented society, human capital becomes increasingly important relative to real capital, leading to a more equal society. Nevertheless, these mechanisms are highly idealized. Above all, they do not take into account effects of globalization or adaptation to international structures. The overall effect of changing GDP remains unclear.

Blinder and Esaki (1978) conducted an early study investigating how inflation and unemployment may relate to income inequality. Inflation is historically referred to as the "the cruelest tax" affecting classes with fixed incomes like retirees in countries where pensions are not adjusted to inflation or only adjusted to inflation with delay,⁴ employees and workers with non-adjusted and non-renegotiated contracts and the poor, depending on possibly static welfare payments. Furthermore, as stated in Palmer and Barth (1977), under a progressive tax regime (as in Germany), poor income earners bear the greatest proportional increase in the tax burden, when cold progression occurs. Nevertheless, this secondary effect of inflation should have an equalizing effect in general as the tax burden generally increases, leading to higher income redistribution.⁵ On the other hand, there are also beneficiaries of non-anticipated positive inflationary shocks, namely borrowers with fixed interest rate contracts, making their lenders the losers of this complementary relationship. From the capitalists' point of view, as

⁴ See Minarik (1979).

⁵ See Heer and Süßmuth (2003).

pointed out by Schultz (1969) and Buse (1982), there are two kinds of production related effects. The first is the so called demand-pull inflation, where the rise of the prices exceeds the rise in production costs. For example, this may be the case for firms with fixed contracts that disconnect the firm from the effect of a higher price level on the production side. In this case, capitalists benefit from inflation. The other kind is the cost-push inflation, the opposite of the former one, leading to decreasing profits. Due to the existence of manifold channels, an overall effect remains a priori undetermined.

When turning to unemployment, we can identify different structural impacts on income distribution in theory. The fact that unemployment benefits are lower than ordinary income does not imply a general statement with respect to inequality. Higher unemployment among the high earners for example, would lead to lower inequality, while the same circumstance in lower income groups has the opposite effect. The empirical fact that the risk of unemployment is higher for lower income groups leads to the assumption that an unemployment shock results in higher inequality. An additional secondary labor market effect is that, with higher unemployment, the income structure of the lower income recipients also shifts downwards due to a shift in the supply-demand equilibrium on the job market, while higher income groups may be less affected. This secondary effect amplifies the overall impact on inequality. Another issue that points in the opposite direction is the fact that, in absolute terms, lower income recipients benefit much more from the social safety net of unemployment insurance. When analyzing incomes at the household level, a further effect may occur. As micro-economic units with potentially more than one earning member, households have the option of in-unit compensation in the event of employment loss. As pointed out in Parker (1996), members of household units compensate for this event by activating earning potential that Parker refers to as “secondary worker effect” or “added worker effect” as in McGinnity (2002). This effect may lower unemployment effects considerably. The overall effect though might still be a tendency to higher inequality when

unemployment rises.

Finally, the effect of government expenditure is somewhat more straightforward. In a welfare state, the state's redistribution mechanism is central to achieve social compensation and is an active tool for keeping the effect of disequalizing forces within limits. The effect of higher government expenditure in welfare states should lower income inequality.

Empirical approaches

Empirical studies to reveal relationships between income inequality and different aspects of the macro-economic business cycle have existed at least since the late 1960s.⁶ The prominent approach of Blinder and Esaki (1978) uses a quantile income share technique. They regress the quantile shares of the income distribution on the aggregate rate of unemployment and inflation. By additionally controlling for general time effects, they identify the cyclical effects in a seemingly unrelated regressions (SUR) setting. To allow for a more general error structure, Jäntti (1994) uses an extension of this model based on a generalized least squares (GLS) approach. As an alternative, other authors use different summary measure approaches and regress measures of inequality, such as the Gini coefficient, on determinants of the business cycle. Additionally, there exists a large number of papers using different shape parameters as dependent variables, after making a parametric distributional assumptions. Here, authors derive the impact of macro-economic variables by using functional relationships between inequality measures and distribution parameters.⁷

As Parker (1998; 2000) points out, despite their initial popularity, the above models have important theoretical drawbacks. The first one is that, in many cases, only a few explanatory variables were included.

⁶ For more details on the history of such studies see Parker (1998).

⁷ The most recent example is Jäntti and Jenkins (2010).

Although this strategy led to more powerful inference when data was sparse, it is likely that such models also suffer from omitted variable bias, making results questionable. In addition to this, these models are also likely to be affected by the non-stationarity of the variables involved.

If variables involved in the estimation have unit roots, it is appropriate to use a cointegration setting as pioneered by Engle and Granger (1987) or Johansen (1988). Jäntti and Jenkins (2010), for example, use a parametric model, namely the Singh-Maddala distribution, to approximate the size distribution of incomes, and apply a cointegration approach to the shape parameters that have an unlimited support.

In general, while time series analysis is a very important tool to show relationships between income inequality and macro-economic variables for different time frames, these approaches also have construction-based limitations. Depending on the information available, this may lead to strong inefficiencies. One source for such an inefficiency, given a sizeable panel data source, is to disregard unit-level information on the size distribution of incomes and to collapse it to point estimates like inequality measures, quantiles or shape parameters. When sufficient micro-data is available, methods from panel data analysis can be used as an alternative. A great advantage of these methods is to allow for different impacts by the observed macro-economic circumstances on different micro-level units. The idea that different income domains are affected differently by changed macro-economic environments was already captured to some extent by the above models, e.g. by looking at potentially different impacts on different quantiles or allowing for different impacts on different shape parameters that depend on some distribution moments of a higher order. But those models are not able to incorporate selective effects on micro-units depending on the vectors of micro-unit characteristics. However, in a panel context with dominating cross sectional information, it is possible to incorporate such dependencies and to identify more detailed (short-term) relationships. Such a data setting generally circumvents the problems as mentioned in the settings with dominating time dimensions as time series persistence does

not lead to biased estimates with a dominating cross sectional dimension.⁸ To approach the question from this direction, Farré and Vella (2008) use a semi-parametric double-index model from Ichimura and Lee (1991) to simultaneously model the relationship between household income, individual characteristics and the business cycle for Spanish data. As the differential effect of a changed macro-economic environment on different households is unclear a priori, the model they use allows for an arbitrary relationship between micro- and macro-economic indices.

Our paper follows the strategy of Farré and Vella (2008) in general to apply the double-index model to micro-level data from the German Socio-Economic Panel (SOEP). Nevertheless there are several differences in the data/variable selection and in the model setup. As regards data, we additionally split the population into a heterogeneous set of homogeneous household types to conduct specific estimations. Regarding the choice of variables, while Farré and Vella (2008) use personal level dependent micro-level variables of the respondent, we use variables that refer to characteristics of the respondent that are derived from the household level. In this we establish an analogy with the nature of the dependent variable. Additionally, we use household-type-specific setups, e.g. controlling for different age groups for elderly households and non-elderly households or a disability dummy for elderly households.⁹ Also, the country-specific characteristic of the formerly divided Germany is accounted for by using a dummy variable.¹⁰ The selection of the macroeconomic variables followed the choices of Farré and Vella (2008) and was based on linear pre-estimations. The variables in scope delivered the highest R^2 in our pre-estimations. Several lags and powers of the variables were looked at. As regards methodology, we extended the

⁸ See Wooldridge (2002: 175).

⁹ Choosing a disability dummy for the elderly only was inspired by linear pre-estimations.

¹⁰ This is a spacial characteristic that was also looked at in a different manner by Farré and Vella (2008) when looking at city sizes.

model by estimating a second level conditional heteroskedasticity to be taken into account for counterfactual analyses and by using locally adapted bandwidths as suggested in Klein and Vella (2006). Additionally we use for the estimation of standard errors a different sub-sampling method with rate of convergence estimation.

Methodology

Following Farré and Vella (2008), we use a special case of semi-parametric multiple-index models from Ichimura and Lee (1991) that allows for unrestricted relationships between two indices representing micro and macroeconomic variables. Let

$$y_{it} = g(I_{1it}, I_{2c(i,t)}) + u_{it} \quad (1)$$

be the relationship between the household income y_{it} of individual i at time t , belonging to some cluster $c(i,t)$ (in our application: month of the questionnaire $\times$ federal state $\times$ year), and let g be an unknown function, satisfying some regularity conditions as given in Ichimura and Lee (1991). Let further be

$$I_{1it} = x'_{it}\beta \quad (2)$$

a linear combination of micro-level (i.e. household-level) variables $x_{it} = (x_{1it}, x_{2it}, \dots, x_{mit})'$ and an unknown column parameter vector $\beta \in \mathbb{R}^m$, and

$$I_{2c(i,t)} = z'_{c(i,t)}\gamma \quad (3)$$

a linear combination of macro-level variables $z_{c(i,t)} = (z_{1c(i,t)}, z_{2c(i,t)}, \dots, z_{nc(i,t)})'$ and an unknown column parameter vector $\gamma \in \mathbb{R}^n$ and let u_{it} be an additive, zero-mean, idiosyncratic error term, uncorrelated with I_{1it} and $I_{2c(i,t)}$ for $i \in \{1, \dots, N\}$ and $t \in \{1, \dots, T\}$.

Model estimation

To estimate the parameter vectors γ and β we minimize the objective function

$$Q(\beta, \gamma) = \frac{1}{NT} \sum_{t=1}^T \sum_{n=1}^N \omega_{it} \{y_{it} - E[y_{it} | x'_{it}\beta, z'_t\gamma]\}^2 \quad (4)$$

with observation weight w_{it} and conditional expectation $E[y_{it} | x'_{it}\beta, z'_t\gamma]$.¹¹ As we do not make any assumptions on the particular functional form of the conditional expectation, we estimate it with a two dimensional Nadaraya–Watson (Nadaraya 1964; Watson 1964) style estimator

$$\hat{E}[y_{it} | x'_{it}\beta, z'_t\gamma] = \frac{\sum_{j=1}^T \sum_{l=1}^N \omega_{jl} K_{2,H}(x'_{it}\beta - x'_{jl}\beta, z'_t\gamma - z'_l\gamma) y_{il}}{\sum_{j=1}^T \sum_{l=1}^N \omega_{jl} K_{2,H}(x'_{it}\beta - x'_{jl}\beta, z'_t\gamma - z'_l\gamma)} \quad (5)$$

Here, in particular, we choose the kernel function to be given by

$$K_{2,H}(\zeta_1, \zeta_2) = \frac{1}{2\pi} |H|^{-\frac{1}{2}} e^{-\frac{1}{2}(\zeta_1, \zeta_2) H^{-1} (\zeta_1, \zeta_2)'} \quad (6)$$

for a bandwidth matrix H .

Fixed bandwidth estimation

As suggested for example by Härdle and Müller (1997), we choose an initial

$$H_1 = \Sigma^{-\frac{1}{2}} \cdot \tilde{h}_j = \Sigma^{-\frac{1}{2}} \cdot \left(\frac{4}{q+2} \right)^{1/(q+4)} n^{-1/(q+4)} \sigma_j \quad (7)$$

¹¹ From now on, we will refer to the cluster index $c(i, t)$ simply by t for notational convenience.

with $\Sigma^{-\frac{1}{2}}$ as the variance-covariance matrix of (ζ_1, ζ_2) and with $\tilde{h}_j$ derived from the generalized rule of thumb plug-in bandwidth of Scott (1992) with σ_j as j -th diagonal element of Σ and n as the sample size. A standard estimation of the variance-covariance matrix delivers

$$\hat{H}_1 = \hat{\Sigma}^{-\frac{1}{2}} \cdot n^{-1/6} \hat{\sigma}_j \quad (8)$$

where we used the fact that the dimension is $q=2$.¹²

Local adaptive bandwidth estimation

While this bandwidth is a good choice asymptotically, we use an additional multi-step procedure to estimate locally adapted bandwidths to reduce the finite sample bias of the Nadaraya-Watson estimator, as suggested by Klein and Vella (2006).

Abramson (1982) suggests the local bandwidth adaptation

$$\hat{\lambda}_{it}^1 = \frac{\sqrt{G(\sum_{j=1}^T \sum_{l=1}^N \omega_{jl} K_{2, \hat{H}_1} (x'_{it}\beta - x'_{jl}\beta, z'_t\gamma - z'_l\gamma))}}{\sum_{j=1}^T \sum_{l=1}^N \omega_{jl} K_{2, \hat{H}_2} (x'_{it}\beta - x'_{jl}\beta, z'_t\gamma - z'_l\gamma)} \quad (9)$$

with $G()$ as the geometric mean over the estimated densities based on the bandwidth matrix $\hat{H}_1$. Defining now

$$\hat{H}_2(i, t) = \hat{H}_1 \cdot \hat{\lambda}_{it}^1 \quad (10)$$

we receive a locally adapted bandwidth matrix. In a second and final step, we estimate

$$\hat{\lambda}_{it}^2 = \sqrt{\frac{G(\sum_{j=1}^T \sum_{l=1}^N \omega_{jl} K_{2, \hat{H}_2} (x'_{it}\beta - x'_{jl}\beta, z'_t\gamma - z'_l\gamma))}{\sum_{j=1}^T \sum_{l=1}^N \omega_{jl} K_{2, \hat{H}_2} (x'_{it}\beta - x'_{jl}\beta, z'_t\gamma - z'_l\gamma)}} \quad (11)$$

¹² This strategy is equivalent to a Mahalanobis transformation of the data and using a diagonal bandwidth matrix.

based on $\hat{H}_2(i, t)$ and receive

$$\hat{H} = \hat{H}(i, t) = \hat{H}_1(i, t) \cdot \hat{\lambda}_{it}^2 \quad (12)$$

which is used in the final estimator in (5). After estimating the link function $\hat{g}$ and the coefficients $(\hat{\beta}, \hat{\gamma})$ we receive

$$\hat{y}_{it} = \hat{g}(\hat{I}_{1it}, \hat{I}_{2t}) = \hat{g}(x_{it}'\hat{\beta}, z_t'\hat{\gamma}) \quad (13)$$

Estimated conditional heteroskedasticity

As regards the disturbance term u_{it} , we can either estimate

$$\hat{u}_{it} = u_{it} \quad (14)$$

or estimate u_{it} by the Nadaraya-Watson estimator applied to the second central moment as

$$\begin{aligned} \hat{\sigma}^2(y_{it}|I_{1it}, I_{2t}) &:= \widehat{Var}[y_{it}|x'_{it}\beta, z'_t\gamma] = \hat{E}[y_{it}^2|x'_{it}\beta, z'_t\gamma] - \hat{y}_{it}^2 \\ &= \frac{\sum_{j=1}^T \sum_{l=1}^N \omega_{jl} K_{2,H}(x'_{it}\beta - x'_{jl}\beta, z'_t\gamma - z'_l\gamma) \cdot y_{it}^2}{\sum_{j=1}^T \sum_{l=1}^N \omega_{jl} K_{2,H}(x'_{it}\beta - x'_{jl}\beta, z'_t\gamma - z'_l\gamma)} - \hat{y}_{it}^2 \quad (15) \end{aligned}$$

Once the conditional standard deviations are known, we can simulate $\tilde{u}_{it}$ as conditional (and even as counterfactual) heteroskedasticity under normality assumption and receive

$$\tilde{y}_{it} = \hat{y}_{it} + \tilde{u}_{it}$$

our final model fit.

Counterfactual analysis

Once the model fits are estimated, the resulting distribution can be derived, e.g., with standard weighted kernel density estimation as

$$\hat{f}_t(\tilde{y}_{it}) = \frac{1}{N} \sum_{j=1}^N \frac{w_{jt} K_h(\tilde{y}_{it} - \tilde{y}_{jt})}{\sum_{l=1}^N \omega_{lt}} \quad (16)$$

with a Kernel function K and a bandwidth h for any point of time $t \in \{1, \dots, T\}$. If we are interested in a counterfactual macro-economic condition, we may simply change the particular value(s) and calculate the model fit

$$\hat{y}_{it}^c := \hat{E}[y_{it} | I_{1it}, I_{2t}^c] = \frac{\sum_{j=1}^T \sum_{l=1}^N \omega_{jl} K_{2,H}(\hat{I}_{1it} - \hat{I}_{1jl}, \hat{I}_{2t}^c - \hat{I}_{2l}) y_{jl}}{\sum_{j=1}^T \sum_{l=1}^N \omega_{jl} K_{2,H}(\hat{I}_{1it} - \hat{I}_{1jl}, \hat{I}_{2t}^c - \hat{I}_{2l})} \quad (17)$$

and analogously the counterfactual error term as shown in (15). Subsequently, based on $\tilde{y}_{it}^c$, we can estimate the counterfactual distribution and derive arbitrary measures of counterfactual location or inequality.

Data and Descriptive Analysis

Household level variables

In order to obtain micro-economic variables, we use the German Socio-Economic Panel (SOEP) provided by the German Institute for Economic Research (DIW) in Berlin.¹³ The SOEP is a representative yearly panel study of private households in Germany that has been running since 1984. We focus our analysis on data between 1996 and 2010. The dependent variable of our model is the equivalized monthly household income at survey time in 2005 prices. We use unbalanced panel data for our estimations, accounting for clustered household structures when calculating standard errors. As an equivalence scale, we use the so called “new OECD scale” to account for household economies of scale. This scale assigns the weight 1 to the first household member, 0.5 to any further member above 14 years of age and 0.3 to children

¹³ For details see Haisken-DeNew and Frick (2005) or Wagner et al. (2007).

under 14. For our estimation strategy, we assume that there may be structural differences between different household types regarding the effects to be estimated. For this reason we define six different household types. The first household type is defined to consist of single elderly households, the second one of multiple elderly households allowing for one spouse to be 55 or older. The third household type is singles without children, where children are family members with at most 17 years of age, and the fourth type is multiple adults without children. The fifth and the sixth group are single and non-single households with children. In order to obtain sufficiently large numbers of observations for our non-parametric estimation, we merge the above six types into three, labeling them as 12, 34 and 56. We chose this notation to clarify that we will still differentiate among the merged types by including a dummy variable in our estimations.¹⁴

Following the above structure, we define slightly different sets of micro-level explanatory variables for each household type. We use household level variables because the income variable is derived from the household level. In all household types we control for the share of adult household members with labor market participation, the share of members with high school degree, the share of members with university degree, and the share of females. Additionally, we use a dummy variable indicating membership in the first of the two merged household types. Finally, we include a dummy for East Germany. For household type 12, we use in addition the share of family members in the age group of below 65 years of age and the share of persons with disability. For household types 34 and 56, we use the share of adults in the age category 31 to 50 and the age category 51 to 64 and take 18 to 30 years of age as the reference category.¹⁵

¹⁴ A table of reference can be found in Appendix B.

¹⁵ It is our understanding that the choice of control variables of relevance in this setting is not exhaustive. The limit of using 8 micro-level and 4 macro-level control variables is, as described below, a computational one.

Sample selection

For reasons of numerical stability, we do not make use of 2.5 percent of the lowest and highest incomes. This strategy avoids numerical convergence problems in sparse data areas, when executing our iteration procedure. As incomes of self-employed household members are often misreported, we do not take into account households with self-employment either.¹⁶ The final full sample consists of 282,018 observations with 41,286 different individuals in 17,860 households. For reasons of computational feasibility, we can only use an 80 percent random subsample of the data.

Macro-economic variables

As macro-economic variables we use the logarithmic, lagged, real GDP on the level of the federal state (Bundesland), the lagged inflation rate, the government final consumption expenditure and the month-specific male ILO unemployment rate for the years 1996 to 2010. These variables were taken from the Genesis Data Base of the German Federal Statistical Office and the World Bank Database, respectively. The choice of these variables is quite similar to the one suggested by Farré and Vella (2008), but there are some differences. While we are aware of the fact that this choice is not exhaustive and that other choices of time and functional transformations may also make sense, we limited the number of macro-economic variables to four to keep numerical calculations in reasonable time limits. Also, we based the specific time and functional pre-transformations of the variables on their goodness-of-fit in linear pre-estimations. In the case of unemployment, we particularly concentrate

¹⁶ Generally, self-employed household members might react to the business cycle even stronger than others. Nevertheless, for the correctness of our estimates, we excluded such households. Results have to be interpreted in an according manner.

on male unemployment as it usually depicts the labor market situation better than general unemployment.

Figure 1 below shows the time series of macro-economic variables between 1991 and 2010. After German reunification in 1991, the time series of inflation shows a particularly strong fall until 1996. The time period before 1996 is also characterized by a rise in unemployment and government expenditure. While we generally welcome variation in the macro-economic data, we must assume that these time series were related to a unique structural change following reunification so that we do not include the years 1991 to 1995 in our estimations.

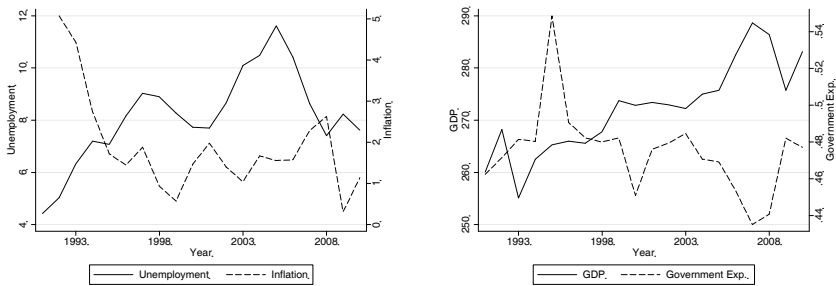


Figure 1. Time series: Male ILO unemployment and inflation rate (left); GDP per capita and government expenditure per capita without pensions (right). Source: Genesis data base, Federal Statistical Office.

There was little change in inflation rates during the 1996-2000 period. They were stable at around 2 percent and somewhat lower after the 2008 crisis. The unemployment rate was around 8 percent with the exception of the years 2003 to 2007 where they peaked at somewhat below 12 percent describing a total variational range of about 4 percent. After the crisis, the unemployment level in Germany stayed exceptionally stable. While the GDP was steadily rising until 2008, it fell strongly by about 5 percent during the crisis. The total range of the time series was around 10 percent of its starting value. The level of government expenditure on the other hand was relatively stable until 2005 and fell

slowly before the crisis. Afterwards it rose by around 6 percent onto the level before 2005.

Turning now to the distribution of incomes in the same time period, Figure 2 shows relevant summary measures to analyze the underlying dynamics. The time series can be divided into four sub-periods. In the first one, between 1991 and 1996, inequality fell substantially as mean income in East Germany converged to that in West Germany. The years 1996 to 2000 were mainly characterized by stagnation in income inequality, while the period from 2000 to 2006 was characterized by a steep rise. Around the crisis year 2008 inequality was about constant.¹⁷

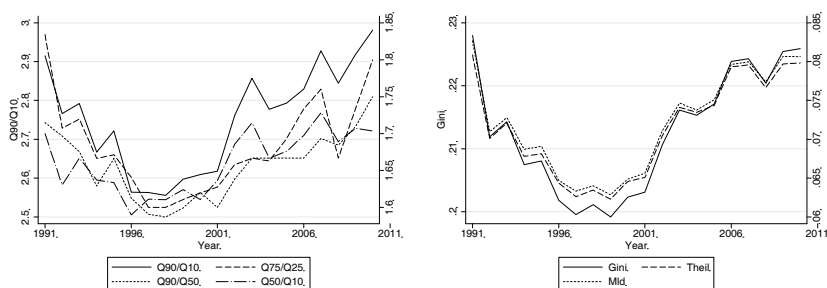


Figure 2. Time series: Quantile ratios of the income distribution (left); Gini, Theil measure and the mean log deviation (right). Own calculations based on data after sample selection.

While there are some parallels between the macro-economic time series and those describing income inequalities, there are at least as many differences, meaning that statistical relationships, if they exist, are not straight forward enough to be seen on a descriptive level and thus require a more rigorous statistical approach.

¹⁷ Here we refer to the time series based on our sample selection. As we dropped the highest 2.5% of observations, our picture slightly differs from the one including it. Including the highest 2.5% of incomes shows a rather falling inequality after the 2008 crisis that strongly affected the right end of the income distribution.

Results

Income distribution and macro-economic states

Before we turn to the interpretation of the outcomes in detail, we begin with some general remarks on the interpretation of the results, which is more difficult than in linear models. For the following reasons, the interpretation has to be more cautious. Firstly, all variables are standardized,¹⁸ so the interpretation refers to standard deviations. Furthermore, all micro-economic coefficients are to be understood relative to the size of the coefficient of the (standard deviation of the) share of adult household members in employment. All macro-economic coefficients are measured relative to the coefficient of GDP. This is because in each index one coefficient has to be normalized to one for identification. In addition to this, there are further restrictions. For example, if the coefficient of high school degree is 1.5, for example, we cannot interpret its magnitude as 1.5 times higher than the one of the constrained coefficient, because the marginal effect still depends of the shape of the function g . While this is less important for the micro-economic variables showing a monotonic behavior for the overall income level as depicted in Figure 3, the macro-economic behavior direction is not always monotonic. Thus in some cases, a seemingly higher coefficient may also come along with a changed sign in the overall partial effect. So, what we may say generally in the above example is that, *ceteris paribus*, one standard deviation of the fraction of adult high school graduates in the family has the same association as 1.5 standard deviations change in the fraction of employed adults in the family.

¹⁸ See identification restrictions of the Double Index Model, Ichimura and Lee (1991).

Table 1.

Double Index Model Estimation Results: Micro- and Macro-Index Components

Dependent Variable: Income	HH-Type 12		HH-Type 34		HH-Type 56	
	coeff.	P-value	coeff.	P-value	coeff.	P-value
Micro-Level Variables						
Employed ^(a)	1		1		1	
High School ^(a)	4.491***	0.005	0.428***	0.000	0.5200***	0.000
University ^(a)	8.413***	0.003	0.851***	0.000	1.3867***	0.000
East Germany	-4.176***	0.003	-0.344***	0.000	-0.1465**	0.035
Age 31-50 ^(a)			0.343***	0.000	0.2831***	0.000
Age 51-64 ^(a)	0.185	0.499	0.431***	0.000	0.1696***	0.000
Disabled ^(a)	0.466	0.296				
HH-Type 1/3/5 ^(b)	-0.510	0.438	-0.578***	0.000	-0.5965***	0.000
Female ^(a)	-1.136*	0.057	-0.115***	0.000	-0.1424***	0.004
Macro-Level Variables						
ln(GDP) ^(c)	1		1		1	
Inflation ^(c)	-0.018	0.654	-0.045	0.267	-0.1952***	0.000
Government Exp. ^(d)	-0.063*	0.094	0.080	0.131	0.1535***	0.002
Unemployment ^(e)	0.012	0.763	0.009	0.817	-0.0228	0.621

Note. P-values are based on clustered subsampling with sample size correction and estimated rate of convergence.

* p<0.1 ** p<0.05 *** p<0.01

^(a)share of adults in the HH, ^(b)according to HH pool, ^(c)lagged ^(d)share of GDP, ^(e)male ILO

All variables are standardized. HH-Type 12: Elderly, HH-Type 34: non-elderly w/o children, HH-Type 56: non-elderly with children. Own calculations.

Index coefficients

Table 1 shows the estimated results based on the double-index model. To analyze the results of micro-economic variables, we take an example-wise look at the results for elderly households. We see that this household type shows positive and highly significant¹⁹ coefficients of

¹⁹ For more details on the calculation of the confidence intervals see Appendix A.

schooling. The fact that the coefficients are so high, relative to the coefficient of the share of employed household members, shows that employment income plays a subordinated role in elderly households. The association with living in the eastern part of Germany comes with a coefficient that has the size of high school degree with a negative sign and is also highly significant, in line with the fact that pensioners are financially less well off in the eastern part of Germany. While the share of females has a significantly negative coefficient, the other variables are insignificant. Households without children and households with children show a qualitatively similar picture, but there, all coefficients are highly significant. Qualitatively, signs are in line with a priori expectations.

As regards macro-economic associations with the income distribution, the coefficient of inflation is highly insignificant for elderly households. This can be explained by the fact that the dominant source of income is pension payment, which is inflation adjusted. Following the same argument, incomes are also unaffected by changes in the unemployment rate. If we take a look at the surface plots in Figure 3, we see that GDP may have an implicitly negative association as a constrained variable. While we cannot be sure about the significance of GDP in the model, it is likely that the influence is not significant and cannot be distinguished from small positive values. This point is important, however, for the interpretation of the association with government expenditure. At first glance, we see a surprising negative sign, but adding this to the fact that the surface mainly falls with a higher macro-economic index, the resulting association with government expenditure is positive and significant at the 10 percent level.

As regards households without children, the signs of the coefficients of inflation and government expenditure are in line with theoretical expectations, even if the results turn out to be statistically insignificant. The sign of the former result is negative, and we may assume that higher inflation is a disadvantage to an important part of the population, e.g. due to more expensive access to the capital market or smaller real asset returns. On the other hand, households benefit from higher government

expenditure, which is in line with the positive coefficient. As regards the unemployment rate, the coefficient is almost zero and statistically not different from zero.

Turning to households with children we see that both the coefficient of inflation and the coefficient of government expenditure are significant at the 1 percent level. While the coefficient of inflation is negative, the one of government expenditure is positive. This is in line with our former results and with economic theory. The fact that we can show a significant influence here may be explained with the higher dependency of this group on fixed employment contracts as their degree of freedom to renegotiate might be limited due to the fact that they have dependent children. Also, a change in government expenditure may affect households with children more, as they are more frequently recipients of social transfers. Finally, the effect of employment is small and insignificant like in the other two cases.

As regards the general insignificance of unemployment in our model, we give four possible explanations. The first one is that a great part of the population is really unaffected by unemployment, never entering the state of unemployment, while the coefficient is fitted globally, making a lack of effect an understandable compromise. The second one is that Germany has a strong social net especially for those with lower income, who are at risk of becoming unemployed. In these cases, and especially in cases of shorter unemployment, unemployment payments compensate the loss of income to a large extent, while our dependent variable does not differentiate between unemployment payments and employment income as income sources. The third explanation lies in the developments on the German labor market, where even despite the financial crises, unemployment as measured with the ILO definition stayed surprisingly stable. Still, there was also generally a strong rise in part-time employment over the whole estimation period, and especially after the 2008 crisis, also a rise in temporary employment and labor leasing (Bundesagentur für Arbeit (BA) 2012) with lower hourly payment. In numbers, the part-time share in employment rose from the

level of 12.9 percent in 1996 to 19.4 percent in 2010 (BA 2011). An important part of this development was a substitute for regular full time employment. While these effects probably strongly influenced the distribution of household incomes, they were not captured in the unemployment definition. So mechanically, the two variables are to a large extent unrelated. Finally, the so-called “added worker effect” may have contributed to the intra-household reduction of the effect of unemployment in multiple adult settings.²⁰

Furthermore, our first point above also discloses an important limitation of the use of aggregate macro-economic variables with the double-index model that is more a model problem than a data-driven one. Although the double-index model is quite flexible in estimating household-specific impacts of macro-variables on the household level due to its nonparametric link function, it does not distinguish between households given the same micro-index level. If the index is not capable of dividing the population into subgroups, such as those influenced by unemployment and those which are generally not, estimation will rather result in a trade-off. If the size of the subgroup of households which are not affected by the unemployment dominates, this trade-off will indicate a non-response to unemployment at the population level.

²⁰ E.g. as a reaction to unemployment, the spouse may compensate income loss by participating in the labor market more intensively. See also McGinnity (2002) on this topic.

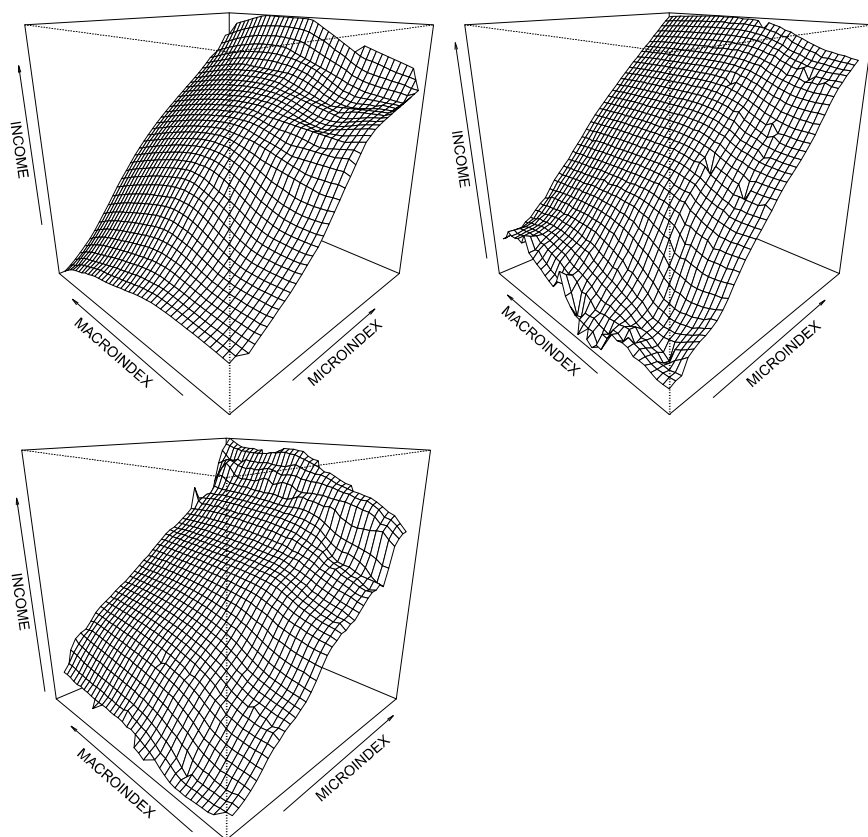


Figure 3. Regression surfaces based on double index model estimation for household types 12 of the elderly (upper left)(a), 34 (upper right)(b) and 56 (lower left)(c) of the non-elderly without and with children, respectively. Own calculations.

The link function

After analyzing the estimated coefficients, we take a closer look at the link function that was estimated nonparametrically along with the coefficients. Figure 3 shows regression surfaces that depict the overall shape of the function g for the three different household types. The upper right and the lower left surfaces of the household types 34 and 56

respectively increase in the direction of the micro-index with minimal local exceptions. For example, a household with higher characteristic values of variables with positive coefficients, such as a higher share of university degree holders in the household, also has a higher mean income as the overall micro-index value increases. The same is true for the macro-index direction. For example, if social expenditure increases *ceteris paribus*, the surface shows mostly a higher income level, although there are exceptions. In general though, the macro-direction has a much weaker increase and a much weaker role for household income than micro-economic characteristics. It is interesting to see that the coefficients in Table 1 do not provide this information, as this behavior is a sole property of the function g . Turning finally to the household type 12 in the upper left, we see that while the surface is leveling up in the micro-index dimension, it is rather leveling down with the higher macro-index. When looking at the signs of the estimated macro-level coefficients, the sign of the significant coefficient of government expenditure is not as expected, either. This may come from the fact that GDP itself, the constrained variable, has a statistically insignificant negative effect (and a true positive one), changing all other signs as they are measured in a relative manner.

To analyze interesting properties of g further, we take a look at different cross-sections of the surface. Figure 4(a)-(c) shows the cross sections derived from Figure 3(a)-(c) at different levels of the macro-economic index. While the dashed lines show the behavior of the function below the 25th centile of the macro-index for each of the household types, the solid line shows the behavior around the median and the dotted line the one above the 75th centile. As the lines for each of the household types are different across centiles, we may conclude that the macro-economic state plays a non-trivial role for the conditional distribution of incomes, so that $g(x, y) \neq g(x)$ for the link function g .

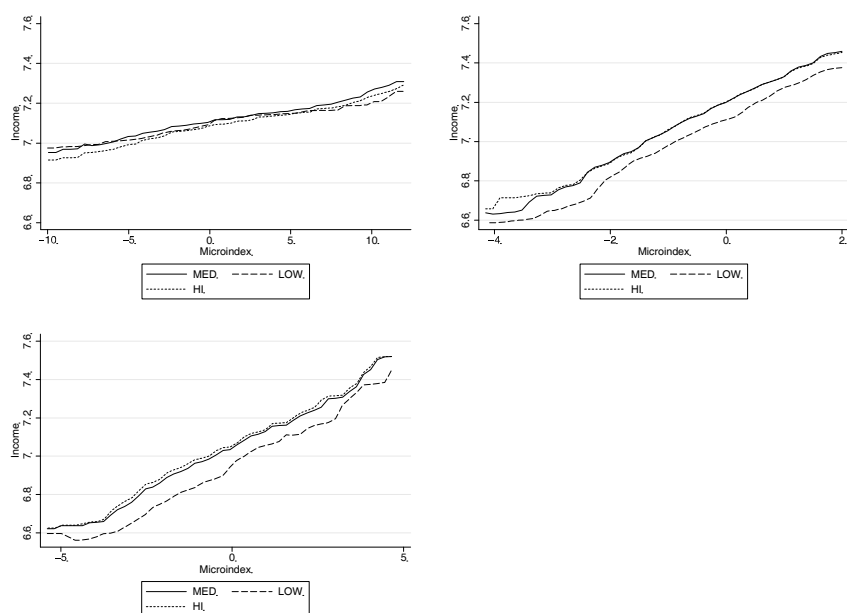


Figure 4. Average cross sections of the double-index regression fits below the 25th (LOW), between the 25th and 75th (MED) and above the 75th (HI) centile of the macro-indices. For household type 12 (upper left)(a), household type 34 (upper right)(b), household type 56 (lower left)(c), own calculations.

An additional conclusion that we may derive is that the lines of different levels are non-parallel. If the link function was of a some simpler form $g(x,y) = g_1(x) + g_2(y)$, then the lines were parallel, meaning that the micro- and macroeconomic arguments of the link function were additively separable. But as the lines show a non-parallel behavior, we expect that the two indices truly interact with each other. In a practical sense, this means that the magnitude of the macro-level impact on households varies with household characteristics. Additionally, we observe that the shapes of the cross-sections differ throughout household types, which gives empirical support to our differentiated approach.

As regards elderly households, we see in Figure 4 that both the micro-index and the macro-index have a limited effect on the expected

income. Additionally, we see that the lines resulting from the levels of the macro-index cross each other, suggesting a strongly non-linear behavior or a generally rather insignificant one (despite the significance of the sole component government expenditure). The other two household types show a clearer pattern. As regards the levels of income resulting from the levels of the macro-index, we observe that while above-average macro states have not much additional effect, negative shocks are associated with higher deviations from the median level. Furthermore, if we take a look at the micro-indices, we observe that, while lowest micro-indices go along with nearly constant income levels, in its other parts the relationship is not very far from linear.

Counterfactual analysis

While interpreting partial effects directly is not straight forward in this context, especially as we leave states of general equilibrium, estimation results may be used to analyze associations with the unconditional distribution level in 2010. The information given on the distribution level allows us to analyze arbitrary measures of location and dispersion. Similar to the results in the last section though, we do not find pronounced changes in the values of these measures under counterfactual circumstances. More interesting is the fact that despite their low magnitudes, the changes we find are mostly in line with a priori expectations and theory. We approach this point by taking a look at distributional effects when putting counterfactual macro-economic states into the estimated model and analyze the counterfactual fits. To receive clear effects, we use emphasized magnitudes of macro-economic states.

The counterfactual macro-economic settings are as follows. To analyze the behavior of inequality measures for changed GDP states, we adjust the GDP counterfactually by 20 percent. In case of inflation we use 3 percent, in the one of government expenditure 20 percent, and we set unemployment to zero and +10 percent as compared to the rate in 2010 in absolute terms. We emphasize here that this counterfactual

model is only able to capture *ceteris paribus* effects, giving a mechanical approach to analyze the link between the income distribution and the state of the world described by macro-economic factors. We will not be able to incorporate secondary effects and forces towards economic equilibria. Results should be interpreted in the light of these limitations.

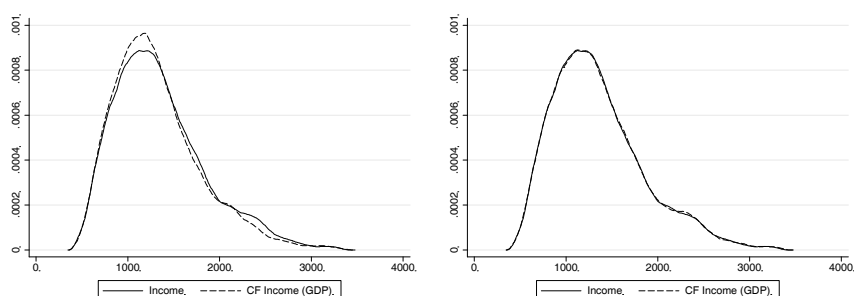


Figure 5. Unconditional income distributions and their counterfactual (cf.) counterparts for lower GDP (-20%) (left) and for higher GDP (+20%) (right) with estimated cf. heterogeneity for the year 2010. Own calculations.

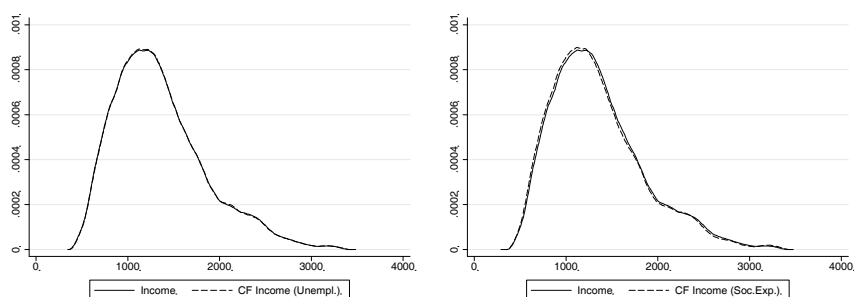


Figure 6. Unconditional income distributions and their counterfactual (cf.) counterparts for lower inflation (-3%) (left) and for higher inflation (+3%) (right) with estimated cf. heterogeneity for the year 2010. Own calculations.

Taking a look at Figure 5 and Table 2, we see that with a lower GDP the distribution shifts to the left (lower Median) and becomes less

dispersed (lower quantile ratios and inequality measures), while when we elevate the level of the GDP, there is a slight right slide of the distribution and a slightly higher dispersion, but the effects are much smaller. In more detail, when the GDP is lowered, the very left part of the distribution remains almost unchanged, while the rest of the distribution experiences a stronger effect. A possible explanation is that the loss of full-time income in the event of an economic downturn (e.g. working time reduction or salary cuts in case of performance-linked payment) is less pronounced for recipients of low incomes in absolute terms, given a strong financial compensation of the social security system in Germany. As we see, households in the very left part of the distribution with probably mostly independent income sources of the labor and capital market (e.g. social benefits) are not influenced by the state of the economy, while the influence is especially pronounced in the middle section. As for high income earners, people with higher incomes tend to have a higher degree of freedom to reallocate their income sources according to the actual state of the economy. In general, we observe that while short-term negative economic shocks have stronger negative effects on the income distribution, positive shocks are either more neutral or have a more delayed effect, underpinning the non-linearity of macro-economic influence. In addition, the effects over the distribution are strongly heterogeneous, meaning that differential effects of changes in macro-economic states on different households exist.

As regards inflation, we see in Figure 6 that lower inflation goes along with a slight right shift of the income distribution, and as the measures of dispersion in Table 2 indicates, inequality becomes slightly lower. We observe the opposite, but more pronounced effect in case of higher inflation, leading to a general left shift of the income distribution and to higher inequality.

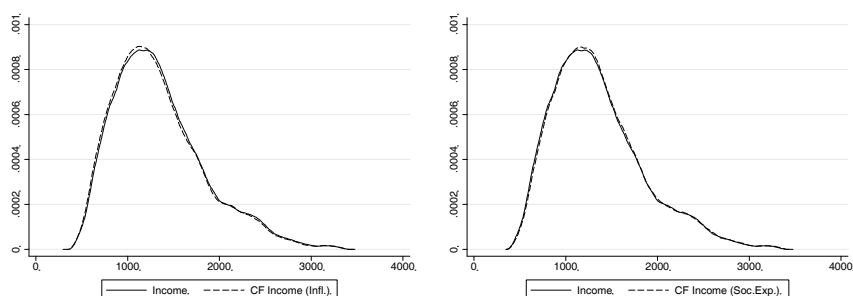


Figure 7. Unconditional income distributions and their counterfactual (cf.) counterparts for lower government expenditure (-20%) (left) and for higher government expenditure (+20%) (right) with estimated cf. heterogeneity for the year 2010. Own calculations.

The effect of changed government expenditure in Figure 7 is in line with general expectations. Lower government expenditure leads to a slight left shift of the income distribution, while a higher expenditure leads to a slight right shift. Additionally, as measures of inequality indicate, lower government expenditure is in line with higher inequality and vice versa. It might surprise at first glance that also the right part of the distribution benefits from government expenditure though (see e.g. the shift of the 90th quantile). The reason is that only one half of the government expenditure is allocated to the social security sector. Other components are spending with more equally distributed beneficiaries like education, free time and sports, health and others as summarized in Table 4 of Appendix B.

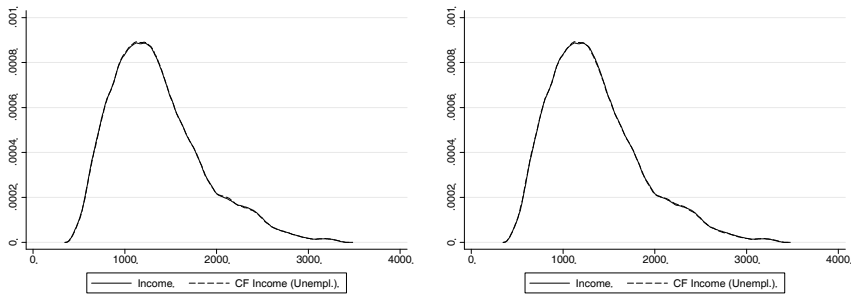


Figure 8. Unconditional income distributions and their counterfactual (cf.) counterparts for no unemployment (left) and for higher unemployment (+10%) (right) with estimated cf. heterogeneity for the year 2010. Own calculations.

Finally, Figure 8 shows the behavior of the income distribution when changing the unemployment rate to zero and when elevating it by 10 percent, respectively. We see almost no effect. This observation is in line with the very small coefficients of unemployment.²¹

Table 2.

Measures of the Income Distribution for Counterfactual Macro Variables for the Year 2010

	Income Factual	GDP (-20%)	Inflation (-3%)	Gov. Exp. (-20%)	Unemployment (0%)
10th Cent.	764.794	763.798	764.410	763.430	764.064
Median	1217.005	1215.375	1218.018	1216.288	1217.400
90th Cent.	1946.051	1943.486	1947.935	1947.104	1947.935
QR 90/10	1181.256	1179.688	1183.525	1183.674	1183.871
QR 50/10	452.210	451.576	453.607	452.857	453.335
QR 90/50	729.045	728.110	729.917	730.816	730.535
Coefficient of variation	0.3731	0.3671	0.3703	0.3776	0.3720

²¹ We have already given possible explanations for this result on page UNEMP.

Gini coefficient	0.2038	0.1991	0.2021	0.2057	0.2030
Theil index	0.0665	0.0640	0.0655	0.0679	0.0661
Mean Log Deviation	0.0676	0.0645	0.0665	0.0688	0.0670
	Income Factual	GDP (+20%)	Inflation (+3%)	Gov. Exp. (+20%)	Unemployment (+10%)
10th Cent.	764.794	764.062	763.430	764.690	763.928
Median	1217.005	1217.266	1216.110	1218.018	1217.265
90th Cent.	1946.051	1948.823	1947.177	1948.823	1948.188
QR 90/10	1181.256	1184.761	1183.747	1184.133	1184.259
QR 50/10	452.210	453.204	452.679	453.327	453.335
QR 90/50	729.045	731.556	731.066	730.805	730.923
Coefficient of variation	0.3731	0.3743	0.3762	0.3695	0.3737
Gini coefficient	0.2038	0.2044	0.2053	0.2016	0.2042
Theil index	0.0665	0.0669	0.0675	0.0652	0.0667
Mean Log Deviation	0.0676	0.0680	0.0685	0.0662	0.0679

Source. Own calculations.

Conclusion

This paper employs a double-index model to analyze the impacts of macro-economic variables on the size distribution of incomes in Germany between 1996 and 2010. Making use of cross-sectional information, the model allows for arbitrary functional relationships between indices of micro- and macro-level variables in order to trace out their joint influence on the income distribution. After estimating the relationship, we employ a counterfactual approach to simulate different macro-economic states and to analyze their impacts on the distribution of incomes.

The results suggest that while the measurable impact of changed macro-economic variables on micro-level incomes is small, it is in parts significant and draws a comprehensive picture of potential links between

micro- and macro-variables. In particular, we find that the level of government expenditure has a significant positive impact on the income levels of elderly households and those with children. The level of inflation is highly significant and negatively associated with the income level of households with children. A negative GDP shock and lower inflation are associated with lower inequality, and higher government expenditure corresponds to lower inequality. As regards unemployment, we find no evidence for an effect on the general level of incomes or on their inequality. This seems to be driven by the fact that the unemployment rate fails to reflect the diversity of structural changes that took place in the German labor market.

The fact that the impact of macro-economic states was found to be very small for Germany has manifold possible explanations. Firstly, as regards the model, the interaction of micro- and macro-level variables was only given at the level of indices, potentially depicting only a mixture of numerous single effects that might partly have different impacts on the population. This limitation is particularly pronounced in the case of unemployment, where only a small part of the population is strongly affected. This situation results in a seemingly strongly reduced effect in our estimation results. Nevertheless, as pronounced changes in macro-economic variables could reveal a pronounced effect on the size distribution of incomes in Spain as analyzed by Farré and Vella (2008), the model itself is capable of delivering significant results in general, but they may be limited due to more stable macro-economic circumstances, where more variation might be needed to be introduced given more granulated information on the state of the economy. While this does not make it easy to distinguish between the limitedness of the model and a lack of response of the size distribution of German incomes to changes in macro-economic states, there are indications that the latter may play a non-negligible role and the results of this paper do not contradict such an assumption. To take a single example, while the share of short-term labor contracts in Germany was around 11 percent in 2004, it was around 30 percent in Spain,²² which indicates a different general structure of the

labor market with a possibly stronger response to the business cycle. In general, differences in state legislatures, government systems and market structures are sources of heterogeneity of possible response mechanisms and may lead to very different results across countries.

References

- Abramson, I. S. (1982). On bandwidth variation in kernel estimates: A square root law. *Annals of Statistics*, 10(4), 1217-1223.
- Berthold, N., Brunner, A. & Zenzen, J. (2010). *Makroökonomische Rahmenbedingungen und die Einkommensverteilung: Welchen Einfluss hat die Finanzkrise?* (Wirtschaftswissenschaftliche Beiträge des Lehrstuhls für Volkswirtschaftslehre, insbes. Wirtschaftsordnung und Sozialpolitik, Nr. 108). Bayerische Julius-Maximilians-Universität Würzburg Wirtschaftswissenschaftliche Fakultät. Retrieved from: http://www.wiwi.uni-wuerzburg.de/fileadmin/12010400/diskussionsbeitraege/DP_108.pdf
- Biewen, M. & Juhasz, A. (2012). Understanding rising inequality in Germany, 1999/2000 - 2005/06. *Review of Income and Wealth*, 58, 62-647.
- Blank, R. M. (1989). Disaggregating the effect of the business cycle on the distribution of income. *Economics*, 56, 141-163.
- Blinder, A. S. & Esaki, H. Y. (1978). Macroeconomic activity and income distribution in the postwar United States. *The Review of Economics and Statistics*, 60(4), 604-609.
- Bundesagentur für Arbeit (2011). *Analyse Arbeitsmarkt Deutschland, Zeitreihen 2011*. Author: Germany. Retrieved from: <http://statistik.arbeitsagentur.de/Statistischer-Content/Statistische-Analysen/Analytikereports/Zentrale-Analytikereports/Jaehrliche-Analytikereports/Generische-Publikationen/Arbeitsmarkt-Deutschland-Zeitreihen/Analyse-Arbeitsmarkt-Deutschland-Zeitreihen-2011.pdf>, on 02.04.2013.
- Bundesagentur für Arbeit (2012). *Arbeitsmarkt Deutschland, Zeitarbeit Aktuelle Entwicklung*. Author: Germany. Retrieved from: <http://statistik.arbeitsagentur.de/Statistischer-Content/Arbeitsmarktberichte/Berichte-Broschueren/Arbeitsmarkt/Generische-Publikationen/Arbeitsmarkt-Deuts->

²² See Zachert (2004).

- chland-Zeitarbeit-Aktuelle-Entwicklung-1HJ2012.pdf, on 02.04. 2013.
- Buse, A. (1982). The cyclical behavior of the size distribution of income in Canada 1947-1978. *Canadian Journal of Economic*, 15, 189-204.
- Easterly, W., & Fischer, S. (2001). Inflation and the poor. *Journal of Money, Credit and Banking*, 33 (2,1), 160-178.
- Engle, R. F., & Granger, C. W. J. Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2), 251-276.
- Farré, L. & Vella, F. (2008). Macroeconomic conditions and the distribution of income in Spain. *LABOUR*, 22(3), 383-410.
- Geyer, C. J. (2006). *5601 notes: The subsampling bootstrap*. Retrieved from: <http://www.stat.umn.edu/geyer/5601/notes/sub.pdf>
- Haisken-DeNew, J. P. & Frick, J. R. (2005). *The desktop companion to the German socio-economic panel study* (Version 8.0). DIW: Berlin.
- Härdle, W., P. Hall & Ichimura, H. (1993). Optimal smoothing in single-index models. *Annals of Statistics*, 21(1), 157-178.
- Härdle, W., & Müller, M. (1997). *Multivariate and semiparametric kernel regression* (SFB 373 Discussion Papers 1997, 26). Humboldt University of Berlin, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes.
- Heer, B., & Süßmuth, B. (2003). *Cold progression and its effects on income distribution* (CESifo Working Paper No. 951). München: CESifo.
- Horowitz, J. (2009). *Semiparametric and nonparametric methods in econometrics*. New York: Springer-Verlag.
- Ichimura, H. & Lee, L. F. (1991). Semiparametric least squares (SLS) and weighted sls estimation of multiple index models: Single equation estimation, in W. Barnett, J. Powell & G. Tauchen (Eds.), *Nonparametric and semiparametric methods in economics and statistics*. Cambridge: Cambridge University Press, pp. 3-49.
- Jäntti, M. (1994). A more efficient estimate of the effects of macroeconomic activity on the distribution of income. *Review of Economics and Statistics*, 76(2), 372-378.
- Jäntti, M., & Jenkins, S. P. (2010). The impact of macroeconomic conditions on income inequality. *Journal of Economic Inequality*, 8, 221-240.
- Johansen, S. (1988). Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control*, 12, 231-254.
- Kuznets, S. (1955). Economic growth and income inequality. *The American*

Economic Review 45(1), 1-28.

- McGinnity, F. (2002). The labour-force participation of the wives of unemployed men: Comparing Britain and West Germany using longitudinal data. *European Sociological Review*, 18(4), 473-488.
- Minarik, J. J. (1979). The size distribution of income during inflation. *Review of Income and Wealth*, 25, 377-392.
- Mount, T. D., & Salem, A. B. (1974). A convenient descriptive model of income distribution: The gamma density. *Econometrica*, 42, 1115-1127.
- Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and its Applications*, 9(1), 141-142.
- Newey, W. K., & Stoker, T. M. (1993). Efficiency of weighted average derivative estimators and index models. *Econometrica* 61, 1199-1223.
- OECD. (2010), National accounts at a glance 2009. Paris: OECD Publishing. DOI: <http://dx.doi.org/10.1787/9789264067981-en>
- Palmer, J. L., & Barth, M. C. (1977). The distributional effects of inflation and higher unemployment. In M. Moon & E. Smolensky (Eds.), *Improving measures of economic well-being*. New York: Academic Press, pp. 201-239.
- Parker, S. (1996). *Explaining the determinants of US income inequality, 1948-90* (Economics Working Paper 163). Durham: University of Durham.
- Parker, S. C. (1998). Income inequality and the business cycle: A survey of the evidence and some new results. *Journal of Post Keynesian Economics*, 21(2), 201-225.
- Parker, S. C. (2000). Opening a can of worms: The pitfalls of time-series regression analyses of income inequality. *Applied Economics*, 32(2), 221-230.
- Politis, D. N., Romano, J. P., & Wolf, M. (1999). *Subsampling*. New York: Springer-Verlag.
- Scott, D. W. (1992). *Multivariate density estimation: Theory, practice, and visualization*. New York: Wiley.
- Singh, S. K., & Maddala, G.S. (1976). A function for the size distribution of incomes. *Econometrica*, 44(5), 963-970.
- Schultz, T. P. (1969). Secular trends and cyclical behavior of income distribution in the United States: 1944-1965. In L. Soltow (Ed.), *Six papers*

- on the size distribution of wealth and income*. New York: Columbia University Press/NBER.
- Thurow, L. C. (1970). Analyzing the American income distribution. *American Economic Review Papers and Proceedings*, 60, 261-269.
- Wagner, G. G., Frick, J. R., & Schupp, J. (2007). The German socio-economic panel study (SOEP): Scope, evolution and enhancements. *Schmollers Jahrbuch* 127(1), 139-169.
- Watson, G. S. (1964). Smooth regression analysis. *Sankhya - The Indian Journal of Statistics*, Series A. 26(4), 359-372.
- Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, Massachusetts: MIT Press.
- Zachert, U. (2004). *Kündigungsschutz, Befristung und Leiharbeit in Europa*. WSI Mitteilungen 3/2004.

Appendix A

For the calculation of the standard errors we use the method of subsampling. Subsampling is an alternative method to the ordinary bootstrap. Its advantage lies in the fact that the sample size used for the re-sampling may be lower than the original sample size. This is particularly important in the given case because non-parametric estimation in greater samples sizes is highly time-consuming and conventional bootstrap calculations would be far beyond reasonable time constraints. For more on subsampling see Politis et al. (1999).

For our model and data we used a subsample size of 12,000 observations for each household type and repeated the draw 200 times for each household type. As a direct estimation of the confidence intervals based on the observed standard deviation under a normality assumption would strongly overestimate the size of the confidence intervals, we additionally corrected the outcome for the sample size that we actually used for our full estimations. For this purpose, we estimated the overall convergence speed of our model as follows. For further details, see Geyer (2006).

Assuming that the speed of convergence follows some rate $\tau_n = n^\rho$ with some fixed ρ , Politis et al. (1999: Chapter 8) shows that

$$G_b^{*-1}(t) \approx b^{-\rho} F^{-1}(t)$$

with

$$\tau_n(\theta_b^* - \hat{\theta}_n) \sim G_b^*(t), \tau_n(\hat{\theta}_n - \theta) \sim F(t),$$

where θ_b^* is the subsampling estimator for some sample size $b < n$, $\hat{\theta}_n$ the estimator based on the full sample of size n and θ the true parameter. Furthermore, $G_b^*(t)$ is the distribution of the subsampling estimator and $F(t)$ the (unknown) limiting distribution function of the parameter estimate $\hat{\theta}_n$.

Now assuming that G and F are smooth and applying this

approximation to two different points t and s with $t > s$ plus taking the natural logarithm,²³ we receive

$$\log [G_b^{*-1}(t) - G_b^{*-1}(s)] \approx \rho \log(b) + \log[F_b^{*-1}(t) - F_b^{*-1}(s)]$$

For several sample sizes b_j and pairs of quantiles (s_i, t_i) , we can estimate ρ with simple OLS after taking the average over the pairs of quantiles for each sample size. The expression then simplifies to

$$\bar{y}_j \approx -\rho \log(b_j) + c(F)$$

with a constant depending on F . Independent of F , ρ can be estimated as

$$\hat{\rho} = -\frac{\text{cov}\{\bar{y}, \log(b)\}}{\text{var}\{\log(b)\}},$$

which is the final estimator proposed.

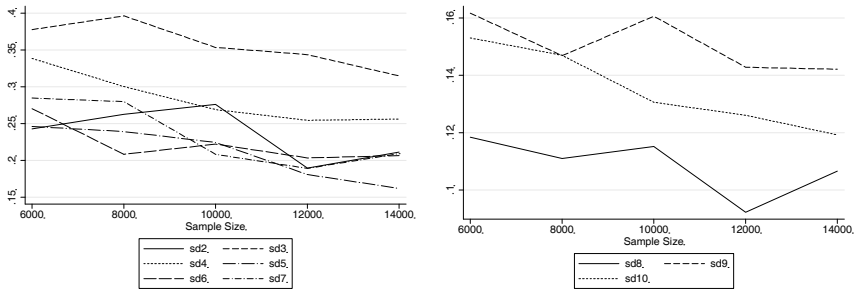


Figure 9. Micro-index (left), macro-index (right) coefficient estimators convergence for different sample sizes measured by the 90/10 quantile ratio for the 7 micro- and 3 macro-coefficients respectively. Own calculations.

To receive more stable estimates, we modified this approach slightly by using an additional mild assumption. We assume that the rate of convergence is the same for all micro- and for all macro-coefficient

²³ We have to take the difference first, so that the logarithm is always defined.

estimators respectively and pool the estimated quantile differences together over all entries of the coefficients β and γ as defined in Section 2. As demonstrated by the empirical convergence estimators in Figures 9(a) and 9(b), our assumption is in line with empirical findings.²⁴ Using 120 observations for each of the sample sizes 6000, 8000, 10000, 12000 and 14000 and calculating the quantile differences of $q_{75}-q_{25}$, $q_{70}-q_{30}$, $q_{65}-q_{35}$, $q_{60}-q_{40}$, $q_{55}-q_{45}$ we estimated a convergence rate of $\hat{\rho}_{micro}=0.306$ and $\hat{\rho}_{micro}=0.399$ with clustered confidence interval $[0.278; 0.333]$ and $[0.376; 0.422]$ respectively over the coefficient pool. The difference comes from the different nature of the two variable sets as macro-variables have much less volatile outcomes. The rates are significantly lower than the theoretical rate $\tilde{r}=0.5$ with an optimal weighting matrix based on a known conditional heterogeneity structure. For more details on the efficient estimation of index models, see Hall and Ichimura (1991), Newey and Stoker (1993) or Horowitz (2009).

²⁴ The dispersions were measured by the 90-10 quantile ratio to minimize the effect of very few extreme values that were probably caused by local optima of the iterations.

Appendix B

Table 3.
Household Types

HH-Type-Group #	HH-Type #	Description
12	1	Single elderly
12	2	Multiple elderly
34	3	Single w/o children
34	4	Multiple w/o children
56	5	Single with children
56	6	Multiple with children

Note. Own definitions.

Table 4.
The components of German final government expenditure in 2010

Government Expenditure	Billion Euro	% of total
General Public Administration	155.20	13.03
Defense	26.23	2.20%
Public Security and Order	40.20	3.38%
Economic Matters	118.90	9.98%
Environment Protection	17.04	1.43%
Housing and Municipal Infrastructure	15.72	1.32%
Health Care	178.48	14.99%
Free time, Sport, Culture and Religion	20.39	1.71%
Education	107.00	8.98%
Social Security	511.81	42.97%
Total	1190.97	100.00%

Source: German Federal Statistical Office, Wiesbaden 2013, Version: 22. March 2013.

