ifo WORKING PAPERS

347
2021

February 2021

The Impact of Labor Market Reforms on Income Inequality: Evidence from the German Hartz Reforms

Lea Immel

ifo INSTITUTE

Leibniz Institute for Economic Research
at the University of Munich

Imprint:

ifo Working Papers
Publisher and distributor: ifo Institute – Leibniz Institute for Economic Research at the University of Munich
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49(0)89 9224 0, Telefax +49(0)89 985369, email ifo@ifo.de
www.ifo.de

An electronic version of the paper may be downloaded from the ifo website:
www.ifo.de

ifo Working Paper No. 347

The Impact of Labor Market Reforms on Income Inequality: Evidence from the German Hartz Reforms*

Abstract

In this paper, I study the distributional consequences of the German Hartz reforms, a set of four comprehensive labor market reforms implemented in Germany between 2003 and 2005. I exploit regional variation in the intensity German counties were affected by the reforms to estimate the causal effect of the Hartz reforms on income inequality. My results suggest that the Hartz reforms have led to a small increase in income inequality. Testing for potential transmission channels, I find that the increase in inequality is partly due to a direct monetary impact of the last Hartz reform, Hartz IV, on the income of households relying on government transfers. Another part can be explained by a rise in the share of transfer recipients. In contrast, the Hartz reforms did not increase income inequality via a rise in labor supply or part-time work, via an increase in (full-time) wage inequality nor via an increase in the number of income earners per household.

JEL Code: D31, J08, J31, J68
Keywords: Labor market policy, income inequality, wages, Germany, Hartz reforms

Lea Immel
ifo Institute – Leibniz Institute for
Economic Research
at the University of Munich
Poschingerstr. 5
81679 Munich, Germany
Immel@ifo.de

* I thank Clemens Fuest, Christian Holzner, Florian Neumeier, Andreas Peichl, and Dominik Sachs for helpful comments and discussion. I gratefully acknowledge financial support by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF) through the project „DFINEQ – Entwicklung von Armut und Ungleichheit in Deutschland und Frankreich in den vergangenen 25 Jahren“.

1 Introduction

Between 2003 and 2005, Germany introduced a set of four comprehensive labor market reforms, the so-called Hartz reforms, to tackle its high unemployment and stagnating growth. The Hartz reforms are considered the most far-reaching reform endeavor in the history of the German welfare state (Jacobi and Kluve, 2007) and have gained great importance both nationally and internationally. They are highly controversial, though, and opinions on the reforms differ widely. For some, mostly international observers, the Hartz reforms are one if not the main reason why Germany in under a decade transformed from 'the sick man of Europe' to 'economic superstar'.[1,2] For others, the reforms are socially unfair, put an end to Germany's social market economy, and increased inequality.[3] Put simply, the debate surrounding the reforms boils down to efficiency vs. equity arguments. Considering this controversy, reliable evidence on the reform's impact on both efficiency and equity is needed. However, while a large literature exists investigating the Hartz reforms' impact on unemployment,[4] so far, little is known about the reforms' distributional consequences.

In this paper, I investigate the distributional consequences of the Hartz reforms by asking whether the reforms have increased income inequality in Germany. In order to answer this question, I follow Card (1992) and exploit the variation in the intensity German regions were affected by the reforms in a difference-in-differences (DiD) type framework where I define a region's treatment intensity according to its labor market performance before the implementation of the Hartz reforms. Conceptionally, one can consider regions with a relatively poor labor market performance prior to the reforms as the treatment group while regions with a relatively good labor market performance form the control group. For the analysis, I construct a novel panel data set of county-level income inequality measures and combine them with county-level labor market characteristics.

I find that the Hartz reforms have a small positive effect on income inequality. In the baseline specification, an increase in treatment intensity by one percentage point leads to an increase in the county's Gini coefficient by 0.11 percentage points. In relation to the sample mean, this translates into an increase in the Gini coefficient by 0.40 percent. This effect is robust to the way treatment is defined (continuously or binarily), to the labor market indicator used to measure treatment intensity (unemployment rate, long-term unemployment rate, or share of social assistance recipients), as well as to sample selection. Looking at the income

[1] The sick man of Europe is a term used since the mid-19^{th} century to label a European country experiencing a time of economic difficulty.

[2] See, for instance, Economist (2004), Economist (2013), Dustmann et al. (2014), or Neue Zürcher Zeitung (2013), among others.

[3] See, for example, Frankfurter Rundschau (2014).

[4] See, for example, Krause and Uhlig (2012), Launov and Wälde (2013), or Bradley and Kügler (2019), among others.

distribution more closely, I find that it is the second to sixth income deciles which lose in terms of income shares while the highest three deciles gain.

Testing possible transmission channels, my results suggest that the increase in income inequality is partly due to a mechanical effect of the last Hartz reform on the income of households relying on government transfers. The last reform, Hartz IV, overhauled the Germany welfare system by combining earnings dependent unemployment assistance with social assistance to a new flat-rate unemployment benefit. Estimating the effect of the Hartz reforms on income inequality excluding all households relying on government transfer payments yields estimates that are about 40 percent smaller than in the baseline specification.

Another part of the increase is due to a rise in the share of households relying on transfer payments. Performing a mediation analysis in order to evaluate whether compositional changes in the working population can generate the observed increase in income inequality reveals that the Hartz reforms have a positive and statistically significant effect on the female employment rate, the share of part-time workers, as well as on the share of households relying on government transfers. Even so, only the latter has in turn an effect on income inequality.

In addition, I find that neither changes within the household, in particular an increase in the number of earners per household, nor changes in the distribution of full-time wages can account for the increase in income inequality. While not being able to test it directly, I argue that an increase in wage inequality between non-standard workers due to increasing labor market dualization may be able to explain the remaining part of the observed increase in income inequality.

Estimating the effect of the Hartz reforms on income inequality, my paper relates to two strands of the literature. First, it relates to a well established literature in labor economics studying the rising inequality of wages and incomes in developed countries.[5] For Germany, these studies find that wage inequality has risen substantially since the 1980s and 1990s. Opinions are divided over whether this rise is attributable to increasing firm heterogeneity and assortativeness in the assignment of workers to firms (Card et al., 2013; Antonczyk et al., 2010), to trade and technological changes, or to changes in labor market institutions including labor market reforms, unionization, and wage setting institutions (Dustmann et al., 2009, 2014). In contrast to wage inequality, systematic analyses of the evolution and determinants of the income distribution in Germany are rare. A number of studies have documented an increase in income inequality since the late 1990s (e.g. Corneo, 2015; Feld and Schmidt, 2016; Battisti et al., 2016) but only a few contributions adopt a systematic approach to analyze the underlying causes (e.g. Biewen and Juhasz, 2012; Biewen et al., 2019; Peichl et al., 2012).

Second, my paper relates to studies evaluating the effect of the Hartz reforms. These studies mainly focus on the effect of the Hartz reforms on unemployment and wages and can

[5] For Germany see Dustmann et al. (2009), Antonczyk et al. (2010), Fuchs-Schündeln et al. (2010), Peichl et al. (2012), Card et al. (2013), Dustmann et al. (2014), Biewen et al. (2019), Bartels (2019), among others.

be categorized into three groups. The first group uses simulations of macroeconomic models calibrated or estimated with pre-reform data (e.g. Krause and Uhlig, 2012; Krebs and Scheffel, 2013, 2017; Launov and Wälde, 2013, 2016; Hartung et al., 2018; Hochmuth et al., 2019; Bradley and Kügler, 2019). These papers usually model specific reform features and find declines of unemployment between 0.1 (Launov and Wälde, 2013) and 2.8 (Krause and Uhlig, 2012) percentage points and mixed effects on wages. A second group of papers uses discontinuities or structural breaks to analyze specific reform policies (e.g. Fahr and Sunde, 2009; Hertweck and Sigrist, 2015; Klinger and Rothe, 2012; Price, 2016). These studies indicate small declines in unemployment in response to each of the Hartz policies. Finally, a number of descriptive studies using flow analyses show that labor market stocks and flows changed significantly after the implementation of the Hartz reforms (see Carrillo-Tudela et al., 2018; Rothe and Wälde, 2017).

I contribute to this literature in several ways. First, this study is the first to provide causal evidence on the impact of the Hartz reforms on income inequality. While most of the literature evaluating the Hartz reforms focus on unemployment and sometimes wages, I specifically look at disposable household income. Since the most controversial of the four Hartz reforms, Hartz IV, targeted government transfers at the household-level, analyzing the distributional effects on disposable household income (i.e., after government intervention) may help to better understand the heavy opposition against the reforms in the German population.

Second, I test various transmission channels linking the Hartz reforms to inequality. Testing the mechanisms behind the distributional effects of the reforms empirically adds to a literature which, by mostly relying on structural models to evaluate the effects of the reforms, predetermines the transmission channels via the assumptions made.

Third, I contribute methodologically to the literature. Since the Hartz reforms were uniformly and simultaneously introduced across the country, estimating their causal impact is challenging. By using the regional variation in treatment intensity within a DiD framework, I propose an approach which has not been used in the context of the Hartz reforms before but may prove interesting for other researchers.

Fourth, exploiting regional variation in treatment intensity to estimate the causal impact of the Hartz reforms on income inequality is only possible because I construct a new and unique panel data set of county-level income inequality measures using the German Microcensus. I am not aware of any other German data set providing income inequality measures at such a disaggregated level.

The rest of the paper is organized as follows. In Section 2, I describe the institutional background and give an overview of the four Hartz reforms. Section 3 explains the empirical strategy and describes the data. Descriptive evidence is presented in Section 4. The empirical results are presented in Section 5. I test and discuss various transmission channels in Section 6. Section 7 concludes.

2 Institutional Background

In spring 2002, after years of rising unemployment, the social-democratic-green government coalition under chancellor Gerhard Schröder appointed a commission composed of 15 experts from industry, politics, and academia to prepare a report on policy reform proposals that would lead to less unemployment. This so-called 'Hartz-commission', named after its chairman Peter Hartz, personnel director at Volkswagen, worked out a program consisting of 13 innovative modules (Hartz, 2002), serving as a blueprint for the four Hartz reform packages or 'Laws for Modern Services in the Labor Market'. The Hartz reforms were implemented in three waves. The first two Hartz Laws (Hartz I and II) were implemented on January 1^{st}, 2003. Hartz III was implemented on January 1^{st}, 2004, and Hartz IV came into effect on January 1^{st}, 2005.

In order to reach their main objectives, that is, accelerating labor market flows and reducing unemployment, the reforms included a modification of active labor market policies (Hartz I and II), the modernization and reorganization of public employment services (Hartz III), as well as a comprehensive reform of the unemployment benefit and social assistance schemes (Hartz IV). Table 1 gives an overview of the main policy changes.

In short, Hartz I facilitated temporary employment and introduced new training subsidies while Hartz II regulated marginal employment (so-called 'mini and midi jobs' (*Mini- und Midijobs*)) and sponsored business start-ups by the unemployed (so-called 'Me, Incs' (*Ich-AGs*)). Hartz III restructured the Federal Employment Agency (*Bundesagentur für Arbeit*) and its local employment offices (*Arbeitsagenturen*) with the objective of making them modern and client-oriented service providers (Weise, 2011).

Hartz IV was the most debated and controversial reform package, changing the structure and generosity of unemployment benefits in order to increase work incentives for the unemployed. Before the reform, those who became unemployed received unemployment insurance payments (*Arbeitslosengeld*) covering 60 percent of previous net wages (67 percent for unemployed workers with dependent children) for a maximum duration of 32 months. After having exhausted the short-term benefits, long-term unemployed workers were eligible to time-unlimited unemployment assistance at a replacement rate of 53 percent (57 percent for workers with dependent children) (*Arbeitslosenhilfe*). Individuals who never contributed to the unemployment insurance scheme received social assistance (*Sozialhilfe*) (Bradley and Kügler, 2019).

The Hartz IV reform shortened the period the person receives unemployment insurance payments, now called unemployment benefit I (*Arbeitslosengeld I*), to six to twelve months, depending on the employment history, and further combined unemployment assistance with social assistance. After the six to twelve months of receiving unemployment benefit I (*Arbeitslosengeld I*), the unemployed person receives a flat-rate unemployment benefit II (*Arbeitslosengeld II*) which is no longer indexed to previous earnings. Individuals deemed capable of working, but who have never contributed to social security receive unemployment benefits II from the beginning. Only those individuals unable to work receive the more generous social assistance

Table 1: The Four Hartz Laws

Hartz I and II	
Implementation:	January 1^{st}, 2003
Target:	Labor Demand
Measures:	• Foundation of 'Staff Service Agencies' (*Personal-Service-Agenturen, PSA*) acting as temporary work agencies for the unemployed • Deregulation of the temporary work sector • Raising of the threshold for incomes exempt from social security contributions ('Mini-jobs') to 400 Euros per month • Introduction of 'Midijobs' with reduced social security contributions for incomes between 400.01 and 800 Euros • Introduction of 'Me, Inc' (*Ich-AG*), a start-up subsidy for the unemployed
Hartz III	
Implementation:	January 1^{st}, 2004
Target:	Market Efficiency
Measures:	• Modernization and reorganization of the public employment agencies, establishing result-based accountability and controlling of local employment offices • Conversion of local employment offices into customer-oriented one-stop-centers • Introduction of a voucher system for placement services (*Vermittlungsgutschein*) and training measures (*Bildungsgutschein*) • Introduction of a standardized profiling process to improve targeting active measures and the allocation of measures and resources
Hartz IV	
Implementation:	January 1^{st}, 2005
Target:	Labor Supply
Measures:	• Shortening of the maximum period the unemployed receive earnings-based unemployment insurance benefits, now called unemployment benefits I (*Arbeitslosengeld I*), from 32 month to 12 month • Pooling of the unemployment assistance payments and social assistance payments into the new flat rate unemployment benefits II (*Arbeitslosengeld II*) • Introduction of one-Euro-Jobs (*Ein-Euro-Jobs*) for unemployment benefits II recipients to increase their income with workfare measures in the public sector

Notes: This is a summary of the main policy changes introduced by the Hartz reforms, I do not claim completeness. For further details see Jacobi and Kluve (2007) among others.

(Jacobi and Kluve, 2007).[6]

Given the extensive nature of the Hartz reforms, the reforms have the potential to affect the distribution of income and income inequality via various channels with the direction of the effect being a priori unclear. On the one hand, the reforms may have led to a decrease in income inequality by reducing unemployment. However, the effect also depends on the type of employment the formerly unemployed are able to find. Since the Hartz reforms deregulated non-standard work, the reduction in unemployment may mainly come from an expansion of the low wage sector and an increase in temporary, marginal, and part-time work. If this is the case, it is more probable that the Hartz reforms have led to an increase in income inequality by increasing wage inequality.

Additionally, Hartz IV affected the income of the unemployed directly by combining earnings-based unemployment assistance with social assistance. But again the effect on income inequality is a priori ambiguous as Hartz IV had heterogeneous effects on household income. Some for-

[6] For a more detailed description of the individual reform packages see Jacobi and Kluve (2007).

mer unemployment assistance recipients are no longer entitled to benefits as benefits are offset by family income. Some receive higher transfers (those with relatively low previous earnings) while others receive lower transfers (those with relatively high previous earnings). Therefore, the question of whether and how the Hartz reforms have affected income inequality can only be answered empirically.

3 Empirical Strategy

This section describes the empirical strategy to estimate the impact of the Hartz reforms on income inequality. First, I explain the strategy used to identify causal effects. Next, I discuss the empirical model and present the data used for estimation.

3.1 Identification Strategy

As the previous section has shown, the Hartz reforms were uniformly and simultaneously introduced in Germany. This absence of legislative variations makes it difficult to estimate the causal impact of the reforms in a quasi-natural experimental set-up. So far, most studies evaluating the reforms have therefore either used simulations of different variants of search and matching models[7] or have relied on reduced form approaches using discontinuities or structural breaks of specific reform policies.[8]

In contrast, I follow Card (1992) and exploit the regional variation in the intensity German counties were affected by the reforms in order to evaluate the overall impact of the four reform packages on income inequality within a DiD framework. Card (1992) was the first to exploit regional variation in a DiD-type set-up. He used regional variation in the fraction of workers affected to measure the effects of an increase in the U.S. federal minimum wage without having to rely on differences in legislation.[9] I adapt his approach to fit the setting of the German Hartz reforms and use the labor market performance of German counties prior to the introduction of the reforms to indicate treatment intensity.

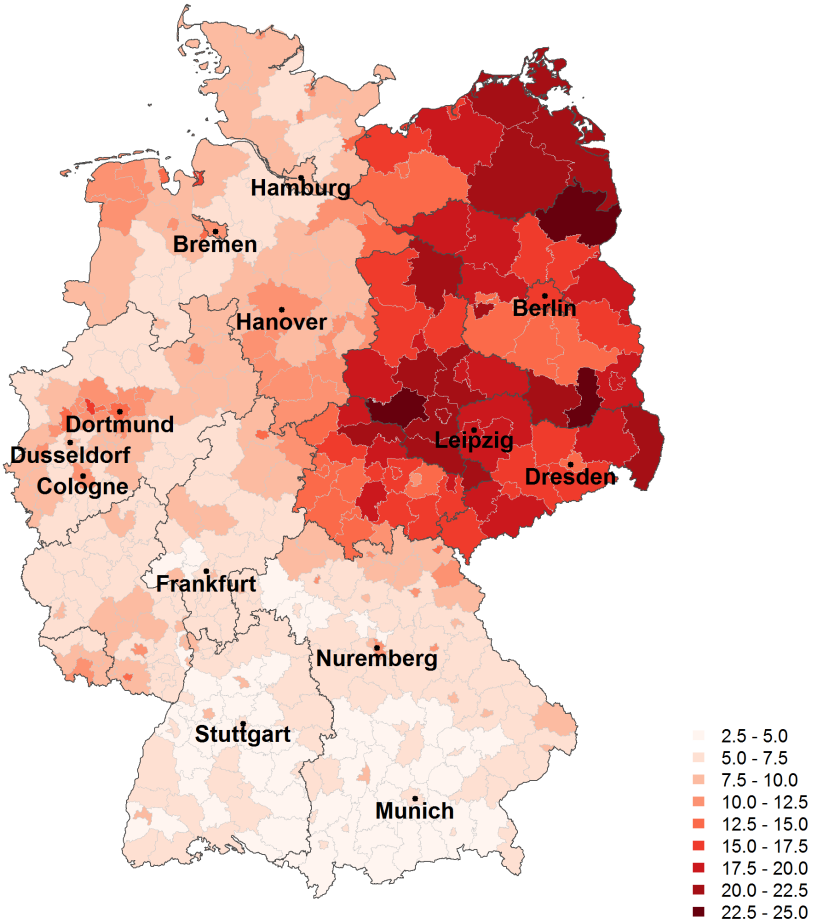
More precisely, I use variation in the county-level unemployment rate of 2002 as treatment indicator. While prior to the reforms unemployment was high in Germany as a whole, there were also considerable regional differences in unemployment rates. Figure 1 depicts the unemployment rate across German counties in 2002, i.e., one year before the first reform was introduced. Table 2 presents summary statistics.

[7] See, for instance, Krause and Uhlig (2012), Krebs and Scheffel (2013), or Launov and Wälde (2013).

[8] See Fahr and Sunde (2009), Hertweck and Sigrist (2015), or Price (2016), among others.

[9] Since Card (1992), the approach has been used in a number of empirical studies, mainly in the minimum wage literature (Caliendo et al., 2018; Dolton et al., 2010; Stewart, 2002, among others) but also to evaluate the effects of health care reforms (Cooper et al., 2011; Gaynor et al., 2013; Propper et al., 2008) or of changes in immigration policies (Clemens et al., 2018).

Figure 1: Unemployment Rates across German Counties in 2002



Notes: The figure shows the unemployment rate across German counties in 2002. Unemployment rates are measured in percent.

Table 2: County Unemployment Rates in 2002—Summary Statistics

	(1) All Counties	(2) West Germany	(3) East Germany
N	365	321	44
Mean	8.6	7.5	16.7
Sd	3.9	2.5	3.1
Min	3.3	3.3	11
P25	5.9	5.7	14.1
P50	7.6	7.2	16.8
P75	10.3	8.8	19
Max	22.7	16.9	22.7

Notes: The unemployment rate is measured in percent.

Striking differences in unemployment rates exist between counties. As Figure 1 reveals, the largest differences exist between East and West German counties, but there is also substantial variation within West and East Germany. County unemployment rates in 2002 range from 3.3 percent to 16.9 percent in West Germany and from 11.0 to 22.7 percent in East Germany, the respective standard deviations are 2.5 and 3.1.

I argue that this variation in unemployment rates implies that the Hartz reforms affected

counties with different intensities. The higher a county's unemployment rate prior to the introduction of the reforms, the stronger the county is affected. Conceptionally, within a DiD framework, one can think of counties with high unemployment rates in 2002 as 'treated' and of counties with low unemployment as 'untreated'. Since the definition of the treatment indicator is crucial to the identification strategy, I perform several robustness tests, where instead of using the unemployment rate of 2002, I use a county's long-term unemployment rate or its share of social assistance recipients (*Sozialhilfeempfänger*) to define treatment intensity. Figure C1 in the appendix shows their regional distribution in 2002, Tables B1 and B2 in the appendix present summary statistics.

Estimating the effects of the Hartz reforms on income inequality using a DiD approach gives the causal treatment effect only if the identifying assumption, i.e., the parallel trend assumption, holds. The parallel trend assumption requires that the untreated observations provide an appropriate counterfactual of the trend the treated observations would have followed in the absence of treatment. In the context of the Hartz reforms, the parallel trend assumption would be violated if trends in income inequality differed between counties with high and low unemployment rates in 2002.

One concern might be that the different trends in economic outcomes East and West Germany experienced after the German reunification may pose a threat to the validity of the identification strategy. I address this concern by including separate year fixed effects for East and West German counties in my empirical model. I also estimate heterogeneous treatment effects for East and West Germany and restrict my sample to West German counties in a robustness test. Moreover, to assess the plausibility of the parallel trend assumption, I test for differences in trends between treated and untreated counties prior to the introduction of the Hartz reforms, both descriptively (see Section 4) and formally (see Section 5).

3.2 Empirical Approach

Stated formally, I estimate the following DiD model in order to identify the impact of the Hartz reforms on income inequality:

$$Y_{ct} = \alpha_c + \beta Treatment_{ct} + \gamma' X_{ct} + \delta_{Region,t} + \epsilon_{ct} \tag{1}$$

The index c refers to the county and index t to the year. α_c denotes county fixed effects, $\delta_{Region,t}$ denotes the separate year fixed effects for East and West German counties. Y_{ct} is the outcome variable of interest, i.e., a measure of income inequality. For the main part of the analysis, I use the county's Gini coefficient but also employ income shares by deciles of the income distribution. The vector X_{ct} includes a set of demographic control variables at the county-level, namely the county's age structure, the share of females, the share of foreigners, as well as population density. $Treatment_{ct}$ is the continuous treatment indicator and is defined as

$$Treatment_{ct} = Post_t \times Treatment\ Indicator_{c,2002} \tag{2}$$

where $Treatment\ Indicator_{c,2002}$ is either the unemployment rate, the long-term unemployment rate, or the share of social assistance recipients in 2002 and $Post_t$ is a dummy variable equal to one if $t \geq 2006$ and equal to zero if $t \leq 2002$. The years 2003 to 2005 are excluded from the DiD analysis.

I am interested in the size of β, the treatment effect. As already mentioned, β only has a causal interpretation if the parallel trend assumption holds.[10] In order to asses the plausibility of the parallel trend assumption, I augment Equation 1 with leading values of the treatment indicator, i.e., $Treatment^j_{ct}$ for $j < 2003$, to test for pre-treatment trends. In addition, I also include a number of lagged treatment variables to study potentially time-varying effects, $\tilde{\beta}_j$. Following Schmidheiny and Siegloch (2019), I thus set up the following event-study model:

$$Y_{ct} = \tilde{\alpha}_c + \sum_{j=1999}^{2014} \tilde{\beta}_j Treatment^j_{ct} + \tilde{\gamma}' X_{ct} + \tilde{\delta}_{Region,t} + \tilde{\epsilon}_{ct} \tag{3}$$

where in order to standardize $\tilde{\beta}_{2003}$ to zero, I drop $Treatment^{2003}_{ct}$ from the regression.

3.3 Data Description

To estimate the models described above, I construct a panel data set which combines county-level inequality measures with labor market characteristics as well as demographic control variables. My data cover the years 1999 to 2014.[11] To construct the data, I mainly rely on information on household income from the German Microcensus. Labor market characteristics as well as further control variables come from the Statistical Offices of the Federation and the Länder (*Statistische Ämter des Bundes und der Länder*) or from the Federal Institute for Research on Building, Urban Affairs and Spatial Developments (*Bundesinstitut für Bau-, Stadt-, und Raumforschung, BBSR*).

[10] Note that Germany introduced a comprehensive tax reform in 2001. The tax reform was phased in between 2001 and 2005 and, inter alia, reduced the minimum and top income tax rate, increased the basic tax-free income, and reduced the tax burden on businesses. If the regional impact of the tax reform were correlated with both, county inequality and labor market performance in 2002, β would be biased. Since the tax reform only affected taxpayers, county unemployment rates and the regional impact of the tax reform are most likely negatively correlated. If so, I would underestimate the size of β. However, estimating the effect of the tax reform on income inequality at the county-level within the same DiD-type framework as described above where treatment intensity is either defined according to a county's share of households in the top ten, 15, or 20 percent of the national income distribution or according to its mean income shows that the tax reform did not have a statistically significant effect on income inequality (see Table B3 in the appendix). I therefore conclude that the tax reform does not pose a threat to my identification strategy.

[11] Germany introduced a federal minimum wage in 2015, mainly as a response to the Hartz reforms. I therefore stop the sample in 2014 to ensure the introduction does not bias my results.

Income Inequality

I compute county-level income inequality measures using information on monthly net household income available in the German Microcensus. The Microcensus is a representative household survey of one percent of the German population. It is carried out annually by the statistical offices of the German states (*Statistische Landesämter*) and administered by the Federal Statistical Office (*Statistisches Bundesamt*). It contains information on various demographic characteristics including the county of residence, employment status, household size, the age of all household members, and monthly net household income, among others. Its large sample size allows to compute income inequality measures at the county-level.

Since the income variable in the Microcensus is interval-censored, i.e., respondents only indicate the income class they are in rather than their precise income, I impute a continuous income figure for each household via interval regression. Including the information on a household's income class and making use of various socio-demographic characteristics, this imputation technique ensures that the empirical distribution of the continuous income variable fits the shape of the distribution of the income classes. As a result, I obtain a single income figure for each household that is consistent with the observed income limits (Royston, 2008).

Having obtained continuous household income, I compute several measures of income inequality at the county-level, namely the Gini coefficient as well as income shares per decile of the income distribution. Note that household incomes are equivalized according to the new OECD equivalence scale to account for differences in household size and adjusted for price changes using the German consumer price index.

Due to several territorial reforms where neighboring counties were merged, the number of counties in the German Microcensus varies across years, resulting in a slightly unbalanced panel data set. The most comprehensive territorial reforms were implemented in East German states where county unemployment rates and thus treatment intensities are higher (see Figure 1). Consequently, the number of 'treated' counties declines over time when using the unbalanced panel. For this reason, I restrict the sample to the 365 German counties unaffected by territorial reforms in the main analysis, ensuring that the number of 'treated' and 'untreated' counties remains constant. I use the unbalanced sample in a robustness test (see Section 5.2).

Labor Market Characteristics and Control Variables

For my main analysis, I use a county's unemployment rate of 2002 as a measure of the intensity a county was affected by the Hartz reforms. As a robustness test, I also look at a county's long-term unemployment rate, calculated as the product of the unemployment rate and the share of the unemployed which have been unemployed for more than a year, as well as a county's share of social assistance recipients in the population. Data on county-level unemployment rates are collected by the BBSR and made available via its online database. The number of social assistance recipients as well as population by county are made available by the Statistical

Offices of the Federation and the Länder (*Statistische Ämter des Bundes und der Länder*).

Furthermore, I include several control variables depicting the demographic situation in a county. I control for a county's age structure, the share of females, the share of foreigners, as well as a county's population density. County-level age shares are calculated using the German Microcensus. The share of females and the share of foreigners is made available by the Statistical Offices of the Federation and the Länder, population densities come from BBSR.

4 Descriptive Statistics

Before presenting the empirical results in the next section, I first take a closer look at the composition of the 'treatment' and 'control' group. Since I use a continuous treatment indicator, I have no treatment and control group in the traditional sense. Therefore, I sort all considered counties into one of three groups—low, medium, and high treatment intensity—according to their unemployment rate in 2002. To have three equally sized groups, I set the cutoff at the 33^{rd} and the 67^{th} percentile. Table 3 shows the distribution of covariates one year prior to the first reform, i.e., in 2002.

Table 3: The Distribution of Covariates in 2002

	Low mean	Medium mean	High mean
Treatment Indicator:			
Unemployment Rate	5.2	7.6	13.1
Inequality Measures:			
Gini	27.9	28.0	26.7
Income Share 1. Decile	3.5	3.5	3.7
Income Share 2. Decile	5.4	5.3	5.5
Income Share 3. Decile	6.5	6.4	6.6
Income Share 4. Decile	7.4	7.4	7.6
Income Share 5. Decile	8.4	8.4	8.6
Income Share 6. Decile	9.4	9.4	9.5
Income Share 7. Decile	10.6	10.6	10.7
Income Share 8. Decile	12.1	12.1	12.1
Income Share 9. Decile.	14.6	14.6	14.3
Income Share 10. Decile	22.3	22.3	21.5
Control Variables:			
Share of Females	50.8	51.2	51.2
Share of Foreigners	8.3	8.0	7.2
Age Share 25 to 34	12.4	12.4	12.1
Age Share 35 to 44	16.4	16.1	16.0
Age Share 45 to 54	13.7	13.4	14.0
Age Share 55 to 64	12.6	12.8	13.5
Age Share over 65	17.1	18.6	19.0
Population Density	313.0	501.3	817.8
East German	0.0	0.0	36.4
Observations	124	120	121

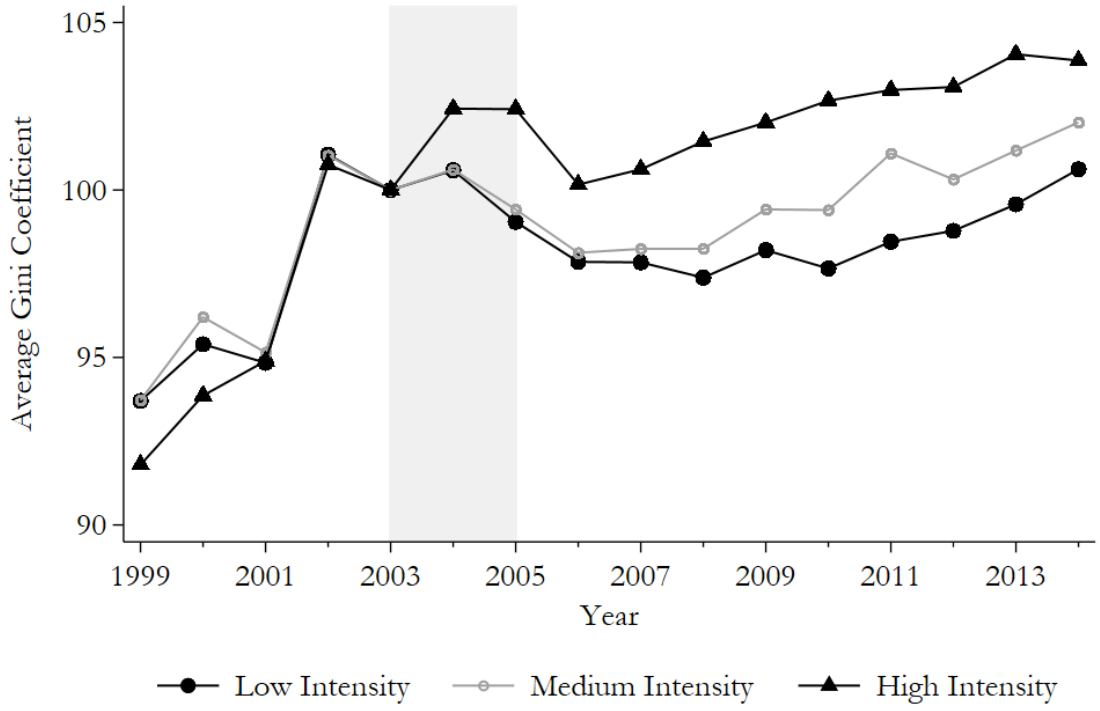
Notes: The unemployment rate, all population shares and the share of East German counties are measured in percent, the Gini coefficient lies between 0 and 100, population density is defines as population/km^2.

Counties with high treatment intensity have on average a lower level of income inequality, a lower share of foreigners, as well as a lower share of the population aged below 45 and a higher

share of the population aged above 45. These differences are due to the fact that all of the 44 East German counties in the sample are in the high treatment intensity group. Generally, incomes in East Germany are more equally distributed. Similarly, the share of foreigners is lower and the share of the older population is higher in the East. This highlights the importance to control for region-year fixed effects in the upcoming analysis as well as to analyze heterogeneous effects.

While looking at the composition of the different groups gives important insights, what matters for the validity of the DiD design is that in absence of treatment, the groups would have followed a parallel trend. To examine the parallel trend assumption visually, I plot the trend in average Gini coefficient by treatment intensity group, normalizing Gini coefficients (2003=100). Figure 2 shows that the identifying assumption is likely to hold as the average Gini coefficients of the three groups move in parallel before 2003.[12]

Figure 2: Visualization of the Parallel Trend Assumption



Notes: The figure shows the trend of the average Gini coefficients by treatment intensity group from 1999 to 2014. Gini coefficients are normalized to 2003=100.

Besides visualizing the parallel trend assumption, Figure 2 also provides first descriptive

[12] The jump in the average Gini coefficients between 2001 and 2002 is quite surprising. Potential explanations include a tax reform introduced in 2001 as well as the fact that between 2001 and 2002 income classes in the German Microcensus were adapted, due to the introduction of the Euro. I am not aware of any other structural break in the Microcensus data between 2001 and 2002. At any rate, I observe this jump in inequality for both regional and national inequality measures. Therefore, is is accounted for in the upcoming analyses via the year fixed effects.

evidence that the Hartz reforms have increased income inequality. After the introduction of the reforms, the trends in Gini coefficients of the three treatment intensity groups begin to diverge. Hereby, the increase in average Gini coefficients is largest for the high treatment intensity group. Figure C2 in the appendix visualizes the parallel trend assumption using the long-term unemployment rate and the share of social assistance recipients in 2002 as treatment indicator. The conclusions drawn remain the same.

5 Empirical Results

In this section, I present the empirical results. I start by presenting my main results, i.e., the impact of the Hartz reforms on different county-level income inequality measures, namely the Gini coefficient as well as income shares by decile of the income distribution. Next, I test the robustness of the results with respect to the definition of treatment as well as sample selection. I extend the analysis by estimating heterogeneous treatment effects between East and West German counties as well as between rural and urban counties. Besides presenting the pooled DiD treatment effects, I provide event study results for every specification in order to ascertain the plausibility of the identifying assumption.

5.1 Main Results

Table 4 presents the pooled DiD treatment effect on county Gini coefficients using the continuous unemployment rate of 2002 as treatment indicator. Columns (1) to (3) present the estimates for different model specifications. The models in columns (1) and (2) include only year or region-year fixed effects. Column (3) presents the estimates of the fully-specified model, which in addition to region-year fixed effects also controls for demographic characteristics at the county-level.

The results suggest that the reforms have increased income inequality in Germany. Independent of the specification, albeit small, the treatment effect is positive and statistically significant at the one percent level. Coefficients become smaller once I include region-year fixed effects and further decrease when adding demographic control variables to the regression. Overall, the coefficients are rather stable, however, ranging between 0.11 to 0.15. In the fully-specified model, the treatment effect is 0.11. The estimate indicates that an increase in treatment intensity by one percentage point is associated with an increase in the Gini coefficient by 0.11 percentage points. This implies an increase by 4.2 percent in terms of its standard deviation. In relation to the sample mean, the effect translate into an increase of the Gini coefficient by 0.40 percent.

Figure 3 illustrates the event study results for the same three specifications, taking 2003 as reference point. Table B4 in the appendix presents the same results in table format. The figure reveals that, as in the DiD model, coefficients become smaller once I include region-year fixed effects. Coefficients are also rather stable over time. Most importantly though, Figure 3

Table 4: The Hartz Reforms and Inequality—DiD

	(1) Gini	(2) Gini	(3) Gini
Treatment	0.148***	0.125***	0.106***
	[0.000]	[0.000]	[0.001]
Share of Females			-0.023
			[0.889]
Share of Foreigners			-0.091*
			[0.079]
Age Share 25 to 34			-0.018
			[0.591]
Age Share 35 to 44			-0.069**
			[0.034]
Age Share 45 to 54			-0.101***
			[0.000]
Age Share 55 to 64			-0.057**
			[0.036]
Age Share over 65			-0.087***
			[0.000]
Population Density			-0.001
			[0.546]
County FE	Yes	Yes	Yes
Year FE	Yes	No	No
Region-Year FE	No	Yes	Yes
Mean Dep. Variable	26.98	26.98	26.98
SD Dep. Variable	2.51	2.51	2.51
Mean Treatment Ind.	8.60	8.60	8.60
R^2	0.306	0.310	0.323
N	4745	4745	4745

Notes: p-values in brackets; * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors are clustered at the level of labor market regions.

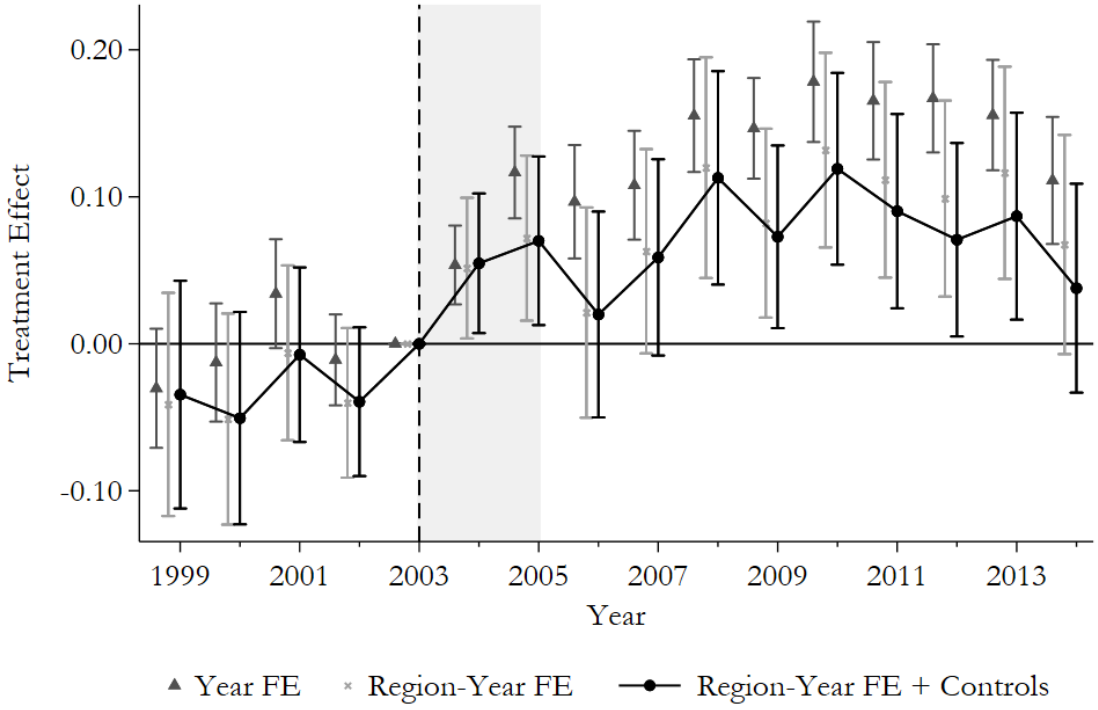
illustrates that all pre-treatment effects are statistically insignificant, independent of the model specification. Prior to the Hartz reforms, counties with low and high unemployment rates in 2002 followed the same trend. I therefore conclude that the parallel trend assumption holds and that the identification strategy and my research design are valid.

The Gini coefficient is the most common summary measure of inequality. It summarizes the whole income distribution into one single number and therefore has an intuitive interpretation. Nevertheless, it also has its drawbacks since summarizing the whole income distribution loses information. To get a more detailed picture of the effects of the Hartz reforms on the income distribution, I complement my analysis by estimating the treatment effects on income shares per decile of the income distribution. Figure 4 summarizes the results.[13]

The figure reveals that the overall effect on income inequality comes from the Hartz reforms' adverse effects on lower and middle incomes. While the treatment effect on the income share of the first decile is statistically insignificant, estimates are negative and statistically different from zero for deciles two to six. Thereby, the third decile incurs the highest loss in its income share. As a consequence of the decrease in income shares in the lower and middle part of the distribution, income shares of the upper three deciles increase.

[13] See Table B5 in the appendix presents the results in table format. Event study results are provided in Figure C3.

Figure 3: The Hartz Reforms and Inequality—Event Study



Notes: The figure plots the event study results. Circles represent point estimates, black lines represent 90 percent confidence intervals. Standard errors are clustered at the level of labor market regions.

5.2 Robustness Test

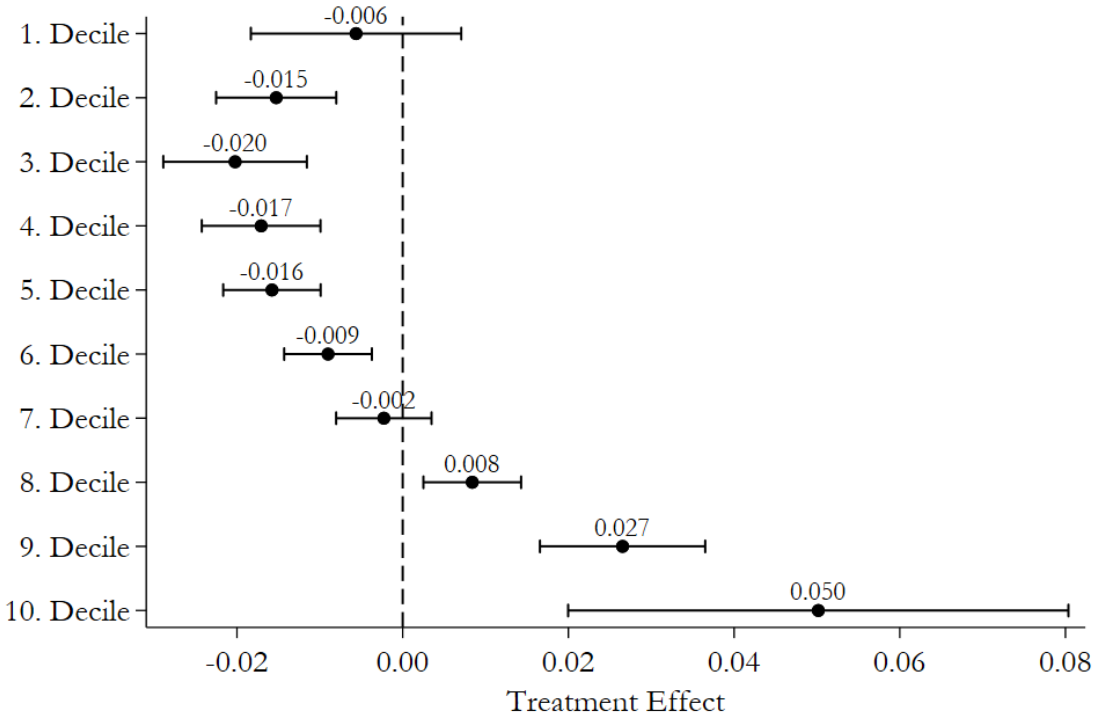
To test the robustness of my results, I undertake a wide set of additional analyses. Since the definition of treatment is crucial to my identification strategy, I modify the empirical specification by using different definitions of treatment (continuous and binary) as well as different treatment indicators (long-term unemployment rate and share of social assistance recipients). In addition, I test how results are affected by sample selection. Note, that for all robustness tests I estimate the effects of the Hartz reform on the Gini coefficient and use the fully specified model, controlling for region-year fixed effects and demographic characteristics.

Alternative Treatment Definitions

Treatment and treatment intensity of the Hartz reforms can be defined in multiple ways. So far, I have used the continuous unemployment rate of 2002 as treatment indicator. In the following, I test whether and how the results change, when I alter the definition of treatment. I do this in two ways.

First, instead of using the continuous unemployment rate of 2002, I use the winsorized continuous unemployment rate as treatment indicator to ensure that my results are not driven by outliers. In addition to winsorizing, I also derive three binary treatment indicators by sorting counties into a treatment and control group according to the distribution of the unemployment

Figure 4: Income Shares by Decile



Notes: The figure plots the DiD results for the ten income shares. Circles represent point estimates, black lines represent 90 percent confidence intervals. Standard errors are clustered at the level of labor market regions.

rate in 2002. The *Binary 50%* treatment variable uses the median of the 2002 unemployment rate distribution as the cut-off and classifies the upper half as the treatment group and the lower half as the control group. The *Binary 33%* and *Binary 25%* treatment variables change the cut-off to the upper and lower third or quarter of the distribution, respectively. Consequently, a county's exact level of the unemployment rate is no longer decisive for the estimation, but rather the distinction between low- and high unemployment counties.[14]

Second, I test the robustness of my results with respect to the choice of treatment indicator. That is, instead of the unemployment rate, I use the long-term unemployment rate as well as the share of social assistance recipients of 2002 to indicate treatment. Since the last and most controversial reform, Hartz IV, directly targeted the long-term unemployed by shortening the period the unemployed received earnings-based unemployment insurance and by combining the unemployment assistance with social assistance, it is interesting to examine whether and how the results change when either the long-term unemployment rate or the share of social assistance recipients are used as alternative treatment indicators.[15] Again, I define treatment

[14] Note, that as a consequence, the *Binary 50%* treatment variable uses all observations in the sample, while the *Binary 33%* (*Binary 25%*) uses only the upper and lower third (quarter) of the unemployment rate distribution and thus loses 33 percent (50 percent) of observations.

[15] For consistency reasons, I define the long-term unemployment rate as the share of the long-term unemployed among the working population. Alternatively, one could also look at the share of the long-term unemployed

continuously as well as binarily.

Table 5: Alternative Treatment Definitions

Panel A: Unemployment Rate					
	(1) Baseline	(2) Winsorized	(3) Binary 50%	(4) Binary 33%	(5) Binary 25%
Treatment	0.106***	0.114***	0.399***	0.694***	0.622**
	[0.001]	[0.001]	[0.007]	[0.001]	[0.015]
Demographic Controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Variable	26.98	26.98	26.98	26.78	26.59
SD Dep. Variable	2.51	2.51	2.51	2.53	2.54
Mean Treatment Ind.	8.60	8.49			
△ Mean Treatment Ind.			5.68	7.86	9.40
R^2	0.323	0.323	0.319	0.356	0.383
N	4745	4745	4745	3185	2392
Panel B: Long-term Unemployment Rate					
	(1) Continuous	(2) Winsorized	(3) Binary 50%	(4) Binary 33%	(5) Binary 25%
Treatment	0.208***	0.229***	0.346**	0.714***	0.697***
	[0.000]	[0.000]	[0.019]	[0.000]	[0.003]
Demographic Controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Variable	26.98	26.98	26.98	26.76	26.57
SD Dep. Variable	2.51	2.51	2.51	2.46	2.46
Mean Treatment Ind.	2.73				
△ Mean Treatment Ind.			2.58	3.58	4.23
R^2	0.324	0.323	0.319	0.351	0.375
N	4745	4745	4745	3159	2379
Panel B: Share of Social Assistance Recipients					
	(1) Continuous	(2) Winsorized	(3) Binary 50%	(4) Binary 33%	(5) Binary 25%
Treatment	0.101***	0.129***	0.316**	0.468***	0.421**
	[0.010]	[0.005]	[0.030]	[0.007]	[0.031]
Demographic Controls	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Variable	26.98	26.98	26.98	27.06	27.19
SD Dep. Variable	2.51	2.51	2.51	2.50	2.46
Mean Treatment Ind.	2.79				
△ Mean Treatment Ind.			2.40	3.29	3.84
R^2	0.316	0.317	0.316	0.334	0.327
N	4719	4719	4719	3146	2353

Notes: p-values in brackets; * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors are clustered at the level of labor market regions.

Table 5 presents the results. Panel A of Table 5 shows the results for the winsorized and binary treatment variables when using the unemployment rate of 2002 as treatment indicator. To facilitate comparisons, column (1) shows the baseline results from column (3) in Table 4. Column (2) presents the estimate for the winsorized treatment variable, columns (3) to (5) present the results for the binary treatment variables. Panel B and Panel C of Table 5 show the results when treatment is based on the long-term unemployment rate and the share of

among all unemployed. DiD and event study results when using the share of the long-term unemployed among all unemployed as treatment indicator are presented in Table B6 and Figure C7 in the appendix.

social assistance recipients, respectively. In both cases, column (1) presents the results for the continuous treatment variable, column (2) presents the results for the winsorized treatment variable, and columns (3) to (5) present the results for the three binary treatment variables.

The results suggest that the effect of the Hartz reforms on income inequality is robust to the way treatment is defined. In all specifications, treatment effects are statistically significant and positive. I find that winsorizing the treatment variable hardly changes the results. In contrast, point estimates become larger when treatment is defined binarily rather than continuously (see Panel A). Treatment effects for the binary treatment variables range between 0.40 and 0.62. However, one has to keep in mind, that the continuous treatment variable measures the effect of an increase in the unemployment rate by one percentage point while the differences in the average unemployment rate of 2002 between treatment and control group when treatment is defined binarily are much larger.[16] Dividing the point estimates in columns (3), (4), and (5) in Panel A by the respective differences in mean treatment indicators to roughly approximate the effect of a one percentage point increase in treatment intensity yields effects between 0.06 and 0.09. This is in the same order of magnitude as the baseline effect.

Panel B reveals that the effect of the Hartz reforms on inequality is about twice as high when using the long-term unemployment rate as treatment indicator instead of the unemployment rate. An increase in the continuous (winsorized) long-term unemployment rate by one percentage point increases the Gini coefficient by 0.21 (0.23) percentage points. When instead treatment is defined binarily, the point estimates are 0.35, 0.71, and 0.70 depending on which binary treatment variable is used. Adjusting the coefficients to take into account the differences in the average long-term unemployment rates between treatment and control group to approximate a one percentage point increase yields effects between 0.13 and 0.20. Since long-term unemployment rates are about three times smaller than unemployment rates, the higher estimates are expected.

Turning to Panel C, the results indicate that a one percentage point increase in the continuous (winsorized) share of social assistance recipients in 2002 leads to an increase in the Gini coefficient by 0.10 (0.13) percentage points. When treatment is defined binarily, the point estimates range between 0.32 and 0.47. Adjusted for the difference in the average share of social assistance recipients between treatment and control group, the effects lie between 0.11 and 0.14.

Taken together, the results demonstrate that the effect of the Hartz reforms on income inequality in Germany is robust to the definition of treatment. Moreover, the event study results show that the common trend assumption holds for all treatment indicators and all specifications (see Figures C4, C5, and C6 in the appendix).

[16] The differences between the mean treatment indicator for treated and untreated counties are 5.68, 7.86, and 9.40 for the *Binary 50%*, *Binary 33%*, and *Binary 25%* treatment indicator, respectively (cf. Table 5).

Sample Selection

In a final robustness check, I test how sample selection influences the results. One concern with using data from 1999 to 2014 in the DiD model might be that the pooled DiD treatment effect masks the fact that the financial crisis of 2008 and the subsequent economic crisis may confound the results. To the extent that the effect of the financial and economic crisis on a county's income distribution is related to its prior labor market performance, estimates will be biased. However, restricting the sample to include only the years 1999 to 2008 shows that the effects are stable over time (see columns (1) and (2) in Table 6).

Table 6: Sample Selection

	(1) Baseline	(2) 1999–2008	(3) Unbalanced	(4) West Germany
Treatment	0.106***	0.098***	0.092***	0.120***
	[0.000]	[0.003]	[0.000]	[0.001]
Demographic Controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes
Mean Dep. Variable	26.98	26.53	26.82	27.44
SD Dep. Variable	2.51	2.47	2.74	2.34
Mean Treatment Ind.	8.60	8.60	10.27	7.49
R^2	0.323	0.344	0.310	0.251
N	4745	2555	6366	5136

Notes: p-values in brackets; * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors are clustered at the level of labor market regions.

Another source of concern may be that so far I have used a balanced sample, dropping all county observations that were subject to territorial reforms between 1999 and 2014. Since almost all territorial reforms were implemented in East German states, this implies that I lose disproportionally many East German observations (see also Section 3.3). I therefore re-estimate Equation 1 using the unbalanced sample as well as a sample including only West German counties. Columns (3) and (4) of Table 6 display the results. In comparison to the baseline estimate, the estimate for the unbalanced sample is just slightly smaller, whereas the estimate for the West German sample is just slightly larger.

5.3 Heterogeneity

The descriptive results in Section 4 established considerable level-differences in the treatment indicator between East and West German counties, highlighting the importance to analyze heterogeneous treatment effects. Moreover, the previous section has already shown that the treatment effect is somewhat smaller in the unbalanced sample which includes more East German counties and somewhat larger in the West German sample, hinting at the presence of heterogeneous treatment effects. I test this formally by adjusting Equation 1 to include an interaction term between the treatment variable and two dummy variables for East and West German counties, respectively:

$$Y_{ct} = \alpha_c + \beta_E East{\times}Treatment_{ct} + \beta_W West{\times}Treatment_{ct} + \gamma' X_{ct} + \delta_{Region,t} + \epsilon_{ct} \quad (4)$$

Column (2) in Table 7 displays the results. As already suspected, treatment effects differ between East and West German counties. In fact, the effect of the Hartz reforms on income inequality seems to be entirely driven by West Germany. Here, a one percentage point increase in treatment intensity leads to a rise in the Gini coefficient by 0.12 percentage points. The effect is statistically significant at the one percent level. In contrast, the Hartz reforms do not seem to have had an impact on income inequality in East German counties. Here, coefficients are smaller and statistically insignificant.

Table 7: Heterogeneous Treatment Effects

	(1) Baseline	(2) East/ West	(3) Rural/ Urban
Treatment	0.106***		
	[0.001]		
West × Treatment		0.120***	
		[0.001]	
East × Treatment		0.051	
		[0.407]	
Rural × Treatment			0.099***
			[0.002]
Urban × Treatment			0.113***
			[0.001]
Demographic Controls	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes
Mean Dep. Variable	26.98		
Mean Treatment Ind.	8.60		
Mean Dep. Variable × East		24.56	
Mean Dep. Variable × West		27.41	
Mean Treatment Ind. × East		16.70	
Mean Treatment Ind. × West		7.49	
Mean Dep. Variable × Urban			27.78
Mean Dep. Variable × Rural			26.28
Mean Treatment Ind. × Urban			8.04
Mean Treatment Ind. × Rural			9.22
R^2	0.323	0.324	0.324
N	4745	4745	4745

Notes: p-values in brackets; * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors are clustered at the level of labor market regions.

Besides analyzing heterogeneous treatment effects for East and West German counties, I also test for heterogeneity between rural and urban counties. Column (3) of Table 7 presents the results. The difference in the treatment effect between rural and urban counties is small. An increase in treatment intensity by one percentage point increases the Gini coefficients in rural counties by 0.10 percentages points. The treatment effect on urban counties is 0.11.[17]

To summarize the results so far: I have established that the Hartz reforms had a small but positive and statistically significant effect on income inequality. In the baseline result, a one percentage point increase in treatment intensity increases a county's Gini coefficient by 0.11

[17] Event study results are presented in Figure C8 in the appendix.

percentage points. This effect is robust to the way treatment is defined, the treatment indicator used, as well as sample selection. Looking at the income distribution more closely, I find that lower and middle income deciles lose while upper income deciles gain in income shares. These estimated treatment effects have causal interpretations. Applying an event study approach, I find that the parallel trend assumption holds in all specifications. Interestingly, the analysis also reveals that the effects are driven by West German counties only. I find no statistically significant effect of the Hartz reforms in income inequality for East German counties.

6 Transmission Channels

Where does the increase in disposable household income inequality come from? Due to the extensive nature of the Hartz reforms, targeting labor demand, market efficiency, as well as labor supply, the reforms may have affected the income distribution via various channels. In the following, I will discuss and/or test these mechanisms, namely changes to the generosity of the transfer system (i.e., redistribution), changes in the composition of the working population or within households, an increase in (full-time) wage inequality, as well as increasing dualization of the labor market.

6.1 Redistribution

The last reform package, Hartz IV, overhauled the German transfer and welfare system. These changes in redistribution had heterogeneous effects on the income of transfer recipients. Using household-level microdata from the Income and Consumption Survey (*Einkommens- und Verbrauchsstichprobe*, EVS) of 2003, Blos and Rudolph (2005) simulate the effect of the introduction of Hartz IV on household income. They show that for social assistance recipients the implementation of Hartz IV had hardly any impact. In contrast, 17 percent of former unemployment assistance recipients are now no longer entitled to benefits. This is because the new unemployment benefits are offset by family income. Of the 83 percent of households still entitled to benefits, 47 percent receive higher transfer payments. These are households which before entering unemployment had relatively low earnings. The other half, i.e., households with relatively high previous earnings, receive lower transfers.

To be able to judge whether and to which extent these mechanical or redistributional effects explain the increase in income inequality, I estimate the effect of the Hartz reforms on income inequality excluding all households relying on government transfer payments. If the observed increase in inequality is partly caused by the mechanical effect on the income of households relying on government transfers, excluding these households and re-estimating Equation 1 should yield smaller treatment effects on income inequality. In order to do so, I use the Microcensus and re-calculate the county-level Gini coefficients, dropping all households which state government transfers as their main income source. Since excluding transfer recipients does not only account

for the mechanical effect but also accounts for changes in the share of households relying on transfer payments, I include the latter as an additional control variable in the regression.

The DiD treatment effects when using either the unemployment rate, the long-term unemployment rate, or the share of social assistance recipients as treatment indicators are presented in columns (2), (4), and (6) of Table 8. To facilitate the comparison, columns (1), (3), and (5) present the effect on household income inequality including all households ('Baseline').

Table 8: Inequality Excl. Transfer Recipients

	Unemployment		Long-term Unemployment		Social Assistance	
	(1) Baseline	(2) w/o Transfer	(3) Baseline	(4) w/o Transfer	(5) Baseline	(6) w/o Transfer
Treatment	0.106***	0.065**	0.208***	0.129**	0.101***	0.086**
	[0.001]	[0.032]	[0.000]	[0.021]	[0.010]	[0.021]
Transfer Recipients		0.017		0.017		0.013
		[0.449]		[0.442]		[0.578]
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Variable	26.98	25.93	26.98	25.93	26.98	25.93
SD Dep. Variable	2.51	2.55	2.51	2.55	2.51	2.55
Mean Treatment Ind.	8.60	8.60	2.73	2.73	2.79	2.79
R^2	0.323	0.279	0.324	0.279	0.316	0.277
N	4745	4745	4745	4745	4719	4719

Notes: p-values in brackets; * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors are clustered at the level of labor market regions.

For all treatment indicators, the estimated treatment effects become smaller when households relying on transfers are excluded. When the unemployment rate or long-term unemployment rate of 2002 are used as treatment indicators, estimates are about 40 percent smaller than in the baseline specifications, in the case of the share of social assistance recipients the estimate decreases by 15 percent. While this is only a rough approximation of the mechanical effect of the Hartz reforms on income inequality, the results do provide a first indication that part of the increase in income inequality is driven by a direct monetary impact on transfer reliant households.

6.2 Compositional Changes in the Working Population

Can changes in the composition of the working population explain the remainder of the increase in income inequality? Simulations of macroeconomic models calibrated to the German economy suggest that the Hartz reforms, inter alia, lowered unemployment and increased part-time work. These changes in population shares may in turn have led to the increase in income inequality.

In order to test this mechanism, I evaluate whether changes in participation and employment rates, in the share of part-time workers, as well as in the share of households relying on government transfer payments generate the increase in income inequality. I use the Microcensus and compute county-level participation rates (by gender), employment rates (by gender), the share of part-time employees in the population, as well as the share of households stating government transfers as their main income source. In a first step, I estimate the effect of the

Hartz reforms on these population shares. Next, I perform a simple mediation analysis, where I add the participation rates, employment rates, the share of part-time employees, and the share of households relying on transfers as additional controls to my baseline specification. If the Hartz reform did indeed cause compositional changes in the working population and if these changes did cause the increase in inequality, the estimates in the mediation analysis should be smaller than the baseline estimate.

Table 9 presents the DiD treatment effects on participation, employment, part-time work, and the share of households relying on transfer payments using the continuous unemployment rate of 2002 as treatment indicator. Table 10 displays the results of the mediation analysis.

Table 9: Participation, Employment, Part-time Work, and Transfers

	Participation			Employment			Part-time	Transfers
	(1) All	(2) Male	(3) Female	(4) All	(5) Male	(6) Female	(7) All	(8) HH
Treatment	-0.005	-0.070	0.060	0.075	0.017	0.133*	0.110***	0.109***
	[0.932]	[0.200]	[0.383]	[0.214]	[0.787]	[0.080]	[0.003]	[0.006]
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Variable	74.90	81.60	68.13	69.87	76.15	63.53	17.65	5.23
SD Dep. Variable	4.42	3.86	6.14	5.73	5.79	6.76	3.98	2.83
Mean Treatment Ind.	8.60	8.60	8.60	8.60	8.60	8.60	8.60	8.60
R^2	0.668	0.351	0.673	0.761	0.544	0.742	0.776	0.220
N	4745	4745	4745	4745	4745	4745	4745	4745

Notes: p-values in brackets; * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors are clustered at the level of labor market regions.

As the results in Table 9 suggest, the Hartz reforms had no effect on participation rates but increased female employment, part-time work, and the share of households relying on transfer payments.[18] These results are largely in line with the literature. Especially the increase in part-time work and female employment is well documented (Burda and Seele, 2016; Carrillo-Tudela et al., 2018).

Nevertheless, the mediation analysis in Table 10 reveals that only the increase in the share of households relying on transfer payments has a small impact on income inequality. Adding the share of households relying on transfers as an additional control variable to the baseline specification lowers the point estimate from 0.11 to 0.09. On the other hand, neither the increase in the female employment rate nor the increase in the share of part-time employees can explain the effect of the Hartz reforms on income inequality. Including participation rates, employment rates, or the share of part-time workers in the regression does not change the estimate on the Gini coefficient. Overall, I conclude that compositional changes in the working population can only explain a small part of the the increase in income inequality and only stem from an increase in the share of households relying on transfers.

[18] Note that the parallel trend assumption does not hold for the employment and male employment rate. Even study results are available on request.

Table 10: Mediation Analysis

	(1) Baseline	(2) Participation	(3) Employment	(4) Part-time	(5) Transfers	(6) All
Treatment	0.106***	0.106***	0.110***	0.106***	0.090***	0.092***
	[0.001]	[0.000]	[0.000]	[0.001]	[0.003]	[0.002]
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Participation	No	Yes	No	No	No	Yes
Employment	No	No	Yes	No	No	Yes
Part-time	No	No	No	Yes	No	Yes
Transfers	No	No	No	No	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Variable	26.98	26.98	26.98	26.98	26.98	26.98
Mean Treatment Ind.	8.60	8.60	8.60	8.60	8.60	8.60
R^2	0.323	0.328	0.332	0.323	0.338	0.343
N	4745	4745	4745	4745	4745	4745

Notes: p-values in brackets; * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors are clustered at the level of labor market regions.

6.3 Changes within the Household

In this subsection, I analyze whether changes within the household as a result of the Hartz reforms have contributed to the increase in income inequality. In the last subsection, I documented that the Hartz reforms increased the share of part-time employees as well as the female employment rate. These two developments are closely related. As Carrillo-Tudela et al. (2018) point out, a large group of low-skilled married women took up mini- and midi-jobs after the introduction of Hartz II to contribute to household income.

To test whether such an increase in the number of earners per household can explain a part of the observed increase in household income inequality, I estimate the effect of the Hartz reforms on individual net income inequality between households' main income earners only. If the effect of the Hartz reform on household income inequality does partly stem from an increase in earners per household topping up household incomes, the treatment effect of the Hartz reforms on income inequality between main income earners should be smaller than the baseline effect. Using the Microcensus, I determine a household's main income earner as the person with the highest individual net income within the household and compute county-level inequality measures using solely the individual net incomes of main income earners.[19]

Columns (2), (4), and (6) of Table 11 present the DiD treatment effects using the continuous unemployment rate, long-term unemployment rate, and share of social assistance recipients as treatment indicator. To facilitate comparisons, columns (1), (3), and (5) display the results on household income inequality ('Baseline'). The estimated treatment effects are largely in line with the results on household income inequality, even somewhat larger. This indicates that an increase in the number of earners within the household is not what drives the effect of the Hartz reforms on income inequality.[20]

[19] Note that individual net incomes are interval-censored. I therefore impute continuous income figures using interval-regressions. Moreover, I adjust individual net incomes for changes in prices.

[20] In principle, changes in household size such as an increase of single households etc. could also lead to an

Table 11: Inequality between Households' Main Income Earners

	Unemployment		Long-term Unemployment		Social Assistance	
	(1) Baseline	(2) Main Earner	(3) Baseline	(4) Main Earner	(5) Baseline	(6) Main Earner
Treatment	0.106***	0.111***	0.208***	0.233***	0.101***	0.143***
	[0.001]	[0.002]	[0.000]	[0.000]	[0.010]	[0.006]
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Variable	26.98	29.92	26.98	29.92	26.98	29.92
SD Dep. Variable	2.51	2.91	2.51	2.91	2.51	2.91
Mean Treatment Ind.	8.60	8.60	2.73	2.73	2.79	2.79
R^2	0.323	0.306	0.324	0.307	0.316	0.302
N	4745	4745	4745	4745	4719	4719

Notes: p-values in brackets; * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors are clustered at the level of labor market regions.

6.4 Wage Inequality

Wages constitute an important—for most households the most important—part of household income. One conjecture is therefore that the Hartz reforms caused an increase in wage inequality which in turn led to an increase in income inequality. In fact, a number of studies have established that wage inequality has risen substantially in Germany since the 1980s and 1990s (Dustmann et al., 2009; Card et al., 2013). Opinions are divided over whether this rise is attributable to increasing firm heterogeneity and assortativeness in the assignment of workers to firms (Card et al., 2013), to trade and technological changes, or to changes in labor market institutions including labor market reforms, unionization, and wage setting institutions (Dustmann et al., 2009, 2014). Studies explicitly analyzing the effect of the Hartz reforms on wages provide mixed evidence and in most cases do not look at the distributional effects in more detail (Launov and Wälde, 2013; Bradley and Kügler, 2019).

I analyze the effect of the Hartz reforms on the wage distribution by re-estimating Equation 1 and using wage inequality as the outcome variable. Since the Microcensus does not contain information on wages, I compute county-level wage inequality measures using information on individual daily wages from the Sample of Integrated Labor Market Biographies (SIAB) of the Institute for Employment Research (*Institut für Arbeitsmarktforschung*, IAB). The SIAB does not include any information on hours worked, I therefore restrict the sample to full-time employees.[21] Further details on the data, the sample selection, and the variables can be found in the appendix.

Table 12 presents the DiD treatment effects on the distribution of full-time wages. Column (1) displays the estimate on the Gini coefficient when using the continuous unemployment rate of 2002 as treatment indicator. Columns (2) and (3) display the results for the continuous

increase in income inequality. I discuss this possibility in Section 6.5.

[21] Note, that I have also restricted the SIAB sample to those counties included in the balanced Microcensus sample.

long-term unemployment rate and the share of social assistance recipients, respectively. Event study results can be found in Table C9 in the appendix.

Table 12: Full-time Wage Inequality

	(1) Unemployment	(2) Long-term Unemployment	(3) Social Assistance
Treatment	0.009	0.034	0.055
	[0.688]	[0.454]	[0.122]
Demographic Controls	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes
Mean Dep. Variable	29.93	29.93	29.93
SD Dep. Variable	2.34	2.34	2.34
Mean Treatment Ind.	8.60	2.73	2.79
R^2	0.644	0.644	0.643
N	4745	4745	4719

Notes: p-values in brackets; * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors are clustered at the level of labor market regions.

The results show that the Hartz reforms did not have a statistically significant impact on full-time wage inequality. While point estimates are positive, they are a order of magnitude smaller than the estimates on income inequality. Moreover, coefficients are not statistically significant, irrespective of the choice of treatment indicator. Thus, the results suggest that the effect of the Hartz reforms on income inequality did not stem from an increase in wage inequality—at least not from an increase in full-time wage inequality. Descriptive evidence suggest that following the Hartz reforms part-time work became a new and important adjustment channel in the German labor market and that wage inequality among part-time employees increased (Burda and Seele, 2016). Unfortunately, due to data restrictions, I am not able to provide evidence on the effect of the Hartz reforms on the distribution of part-time wages, but discuss this potentially important transmission channel in the next section.

6.5 Further Transmission Channels

Another possible transmission channel, which has only be hinted at in the preceding analysis, is labor market dualization. Labor market dualization refers to growing disparities between a relatively stable core labor force and a flexible margin of non-standard work (i.e., part-time, temporary, and marginal employment, fixed-term contracts etc.). Germany's path to a dual labor market began in the 1980s with the Employment Promotion Act (*Beschäftigungsförderungsgesetz*) and was reaffirmed by the Hartz reforms. The Hartz reforms deregulated the temporary work sector, introduced subsidies for marginal part-time employment exempted from or with reduced social security contributions, and further liberalized other forms of atypical employment such as lowering the age threshold for fixed-term contracts (Eichhorst and Marx, 2011).

Increasing dualization of the German labor market can lead to a rise in income inequality by increasing the share of workers in non-standard employment as well as by increasing the within-group wage inequality of non-standard workers. The growing importance of marginal

and part-time work has already been documented (see, for example, Burda and Seele (2016) as well as Section 6.2 of this paper). Nevertheless, the mediation analysis in Section 6.2 also showed that controlling for the share of part-time employees does not change the baseline estimate.

In addition to analyzing the effect of the Hartz reforms on the population share of (marginal) part-time employees, it would thus be informative to also study their effect on the distribution of (marginal) part-time wages. Unfortunately, data restrictions do not allow me to do so.[22] However, Carrillo-Tudela et al. (2018) provide first descriptive evidence that within group wage inequality increased for marginal and part-time workers after 2003/ 2005. They find that marginal workers incurred wage losses across the entire distribution but most notably above the median, for part-time employees wages below the median decreased. Far from being causal, this evidence nevertheless gives a first tentative indication that a part of the effect of the Hartz reforms on income inequality may be attributed to the increasing dualization of the German labor market.

Besides labor market dualization, another transmission channel might be changes in household structures. Peichl et al. (2012) find that changes in household structures can explain 78 percent of the rise in gross income inequality between 1991 and 2007 and 22 percent of the rise in net income inequality. Since the Hartz reforms have increased incentives for transfer recipients to live alone (benefits are now offset by household income), changes in household size may account for some of the increase in income inequality. Since the effect of the Hartz reforms on the average number of household members per county is statistically insignificant and including the average number of household members per county as an additional control variable to the baseline specification does not alter my results, I do not find this explanation very likely.[23]

Taken together, the results in this section indicate that a part of the Hartz reforms' effect on income inequality is driven by the direct effect of Hartz IV on the income of households relying on government transfers as well as by an increase in the share of transfer recipients. In contrast, neither changes within the household nor changes in the full-time wage distribution can account for the increase in disposable household inequality. Since the Hartz reforms brought about a great deal of deregulation at the margin of the labor market but apart from benefit cuts left core (i.e., full-time) workers unaffected (Eichhorst and Marx, 2011), these results are expected. What drives the remainder of the increase in inequality? I find a rise in wage inequality of non-standard workers due to the increasing dualization of the German labor market the most likely candidate. Due to data restrictions, I am not able to test this directly, though, and leave it for further research.

[22] In Germany, part-time work can mean anything from working one to 39 hours. Since the SIAB data includes no information on hours worked, looking at the distribution of part-time daily wages is not informative.

[23] Results are available on request.

7 Conclusion

The German Hartz reforms are often cited as the most far-reaching reform endeavor in the history of the German welfare state (Jacobi and Kluve, 2007). It is therefore not surprising that a huge controversy surrounds them. The public debate about the Hartz reforms boils down to efficiency vs. equity arguments, where proponents stress the reforms' apparent success in lowering unemployment and stimulating growth and opponents lament their distributional consequences. While a large literature exists investigating the Hartz reforms' impact on unemployment, so far, little is known about their impact on inequality. My paper contributes to the literature by providing first causal evidence on the effect of the Hartz reforms on income inequality.

Exploiting the regional variation in the intensity German counties were affected by the reforms in a DiD framework and using county-level data on disposable household income inequality from the German Microcensus, I find that the Hartz reforms had a small positive effect on income inequality. This effect is robust to the way treatment is defined, the choice of treatment indicator, and sample selection. Looking at the income distribution more closely, I document that it is the second to sixth income deciles which lose in terms of income shares, whereas the upper three deciles gain.

Testing for possible transmission channels, the results suggest that the increase in income inequality is partly due to a mechanical monetary effect of the last Hartz reform on the income of households relying on government transfers. Another part of the increase in income inequality is caused by a rise in the share of households relying on transfer payments. In contrast, neither an increase in the number of earners per household nor changes in the distribution of full-time wages seem to play a role. Even though I am not able to test it directly, I argue that an increase in the wage inequality of non-standard workers due to increasing labor market dualization may be able to explain the remaining part of the observed increase in income inequality.

All in all, I find that the Hartz reforms seem to have had their intended effect. Designed to increase the flexibility of the German labor market and to decrease unemployment and stimulate growth, the Hartz reforms mainly deregulated non-standard work and cut benefits for the unemployed. Distributional consequences due to changes in the transfer system and an expansion of the non-standard work sector must have been taken under consideration. However, whether the social-democratic-green government coalition under Gerhard Schröder anticipated the immense political backlash to the reforms is another question. Given the fact, that the SPD has meanwhile distanced itself from the Hartz reforms and ran their 2017 election campaign under the heading 'social equity', it stands to reason that the SPD has underestimated employees' needs for secure and stable jobs. One lesson from the German Hartz reforms may thus be that reforms aimed to tackle structural problems in the economy come at a heavy political cost when components of flexibility and security are not appropriately balanced.

References

Antonczyk, D., Fitzenberger, B. and Sommerfeld, K. (2010). Rising wage inequality, the decline of collective bargaining, and the gender wage gap, *Labour Economics* **17**(5): 835–847.

Bartels, C. (2019). Top incomes in Germany, 1871–2014, *The Journal of Economic History* **79**(3): 669–707.

Battisti, M., Felbermayr, G. and Lehwald, S. (2016). *Entwicklung der Einkommensungleichheit: Daten, Fakten und Wahrnehmungen*, Stiftung Familienunternehmen.

Biewen, M. and Juhasz, A. (2012). Understanding rising income inequality in Germany, 1999/2000–2005/2006, *Review of Income and Wealth* **58**(4): 622–647.

Biewen, M., Ungerer, M. and Löffler, M. (2019). Why did income inequality in Germany not increase further after 2005?, *German Economic Review* **20**(4): 471–504.

Blos, K. and Rudolph, H. (2005). Simulationsrechnungen zum Arbeitslosengeld II: Verlierer, aber auch Gewinner, *IAB Kurzbericht 17*, Institut für Arbeitsmarkt-und Berufsforschung (IAB).

Bradley, J. and Kügler, A. (2019). Labor market reforms: An evaluation of the Hartz policies in Germany, *European Economic Review* **113**: 108–135.

Burda, M. C. and Seele, S. (2016). No role for the Hartz reforms? Demand and supply factors in the German labor market, 1993–2014, *SFB 649 Discussion Paper 10*, Humboldt-Universität Berlin.

Caliendo, M., Fedorets, A., Preuss, M., Schröder, C. and Wittbrodt, L. (2018). The short-run employment effects of the German minimum wage reform, *Labour Economics* **53**: 46–62.

Card, D. (1992). Using regional variation in wages to measure the effects of the federal minimum wage, *ILR Review* **46**(1): 22–37.

Card, D., Heining, J. and Kline, P. (2013). Workplace heterogeneity and the rise of West German wage inequality, *The Quarterly Journal of Economics* **128**(3): 967–1015.

Carrillo-Tudela, C., Launov, A. and Robin, J.-M. (2018). The fall in German unemployment: A flow analysis, *IZA Discussion Papers 11442*, Institute of Labor Economics (IZA).

Clemens, M. A., Lewis, E. G. and Postel, H. M. (2018). Immigration restrictions as active labor market policy: Evidence from the Mexican bracero exclusion, *American Economic Review* **108**(6): 1468–87.

Cooper, Z., Gibbons, S., Jones, S. and McGuire, A. (2011). Does hospital competition save lives? Evidence from the English NHS patient choice reforms, *The Economic Journal* **121**(554): F228–F260.

Corneo, G. (2015). Kreuz und quer durch die deutsche Einkommensverteilung, *Perspektiven der Wirtschaftspolitik* **16**(2): 109.

Dolton, P., Bondibene, C. R. and Wadsworth, J. (2010). The UK national minimum wage in retrospect, *Fiscal Studies* **31**(4): 509–534.

Dustmann, C., Fitzenberger, B., Schönberg, U. and Spitz-Oener, A. (2014). From sick man of Europe to economic superstar: Germany's resurgent economy, *Journal of Economic Perspectives* **28**(1): 167–88.

Dustmann, C., Ludsteck, J. and Schönberg, U. (2009). Revisiting the German wage structure, *The Quarterly Journal of Economics* **124**(2): 843–881.

Eberle, J. and Schmucker, A. (2019). Creating cross-sectional data and biographical variables with the Sample of Integrated Labour Market Biographies 1975–2017: Programming examples for Stata, *FDZ-Methodenreport 4*, Institut für Arbeitsmarkt-und Berufsforschung (IAB).

Economist (2004). Germany on the Mend., https://www.economist.com/news/2004/11/17/germany-on-the-mend. Accessed: 2020-07-31.

Economist (2013). Wunderreform, https://www.economist.com/europe/2013/03/16/wunderreform. Accessed: 2020-08-02.

Eichhorst, W. and Marx, P. (2011). Reforming German labour market institutions: A dual path to flexibility, *Journal of European Social Policy* **21**(1): 73–87.

Fahr, R. and Sunde, U. (2009). Did the Hartz reforms speed-up the matching process? A macro-evaluation using empirical matching functions, *German Economic Review* **10**(3): 284–316.

Feld, L. P. and Schmidt, C. M. (2016). Jenseits der schrillen Töne, *Perspektiven der Wirtschaftspolitik* **17**(2): 188.

Frankfurter Rundschau (2014). Zehn Jahre Hartz IV und die Folgen, https://www.fr.de/meinung/zehn-jahre-hartz-folgen-11188166.html. Accessed: 2020-08-02.

Fuchs-Schündeln, N., Krüger, D. and Sommer, M. (2010). Inequality trends for Germany in the last two decades: A tale of two countries, *Review of Economic Dynamics* **13**(1): 103–132.

Gaynor, M., Moreno-Serra, R. and Propper, C. (2013). Death by market power: Reform, competition, and patient outcomes in the National Health Service, *American Economic Journal: Economic Policy* **5**(4): 134–66.

Hartung, B., Jung, P. and Kuhn, M. (2018). What hides behind the German labor market miracle? Unemployment insurance reforms and labor market dynamics, *CESifo Working Paper 7379*, CESifo Munich.

Hartz, P. (2002). *Moderne Dienstleistungen am Arbeitsmarkt: Vorschläge der Kommission zum Abbau der Arbeitslosigkeit und zur Umstrukturierung der Bundesanstalt für Arbeit; Bericht der Kommission*, AS-Druck.

Hertweck, M. S. and Sigrist, O. (2015). The ins and outs of German unemployment: A transatlantic perspective, *Oxford Economic Papers* **67**(4): 1078–1095.

Hochmuth, B., Kohlbrecher, B., Merkl, C. and Gartner, H. (2019). Hartz IV and the decline of German unemployment: A macroeconomic evaluation, *IZA Discussion Paper 12260*, Institute of Labor Economics (IZA).

Jacobi, L. and Kluve, J. (2007). Before and after the Hartz reforms: The performance of active labour market policy in Germany, *Zeitschrift für ArbeitsmarktForschung-Journal for Labour Market Research* **40**(1): 45–64.

Klinger, S. and Rothe, T. (2012). The impact of labour market reforms and economic performance on the matching of the short-term and the long-term unemployed, *Scottish Journal of Political Economy* **59**(1): 90–114.

Krause, M. U. and Uhlig, H. (2012). Transitions in the German labor market: Structure and crisis, *Journal of Monetary Economics* **59**(1): 64–79.

Krebs, T. and Scheffel, M. (2013). Macroeconomic evaluation of labor market reform in Germany, *IMF Economic Review* **61**(4): 664–701.

Krebs, T. and Scheffel, M. (2017). Labor market institutions and the cost of recessions, *Working paper*.

Launov, A. and Wälde, K. (2013). Estimating incentive and welfare effects of nonstationary unemployment benefits, *International Economic Review* **54**(4): 1159–1198.

Launov, A. and Wälde, K. (2016). The employment effect of reforming a public employment agency, *European Economic Review* **84**: 140–164.

Neue Zürcher Zeitung (2013). Als Deutschland noch Reformen wagte, https://www.nzz.ch/als-deutschland-noch-reformen-wagte-1.18044860. Accessed: 2020-08-02.

Peichl, A., Pestel, N. and Schneider, H. (2012). Does size matter? The impact of changes in household structure on income distribution in Germany, *Review of Income and Wealth* **58**(1): 118–141.

Price, B. (2016). The duration and wage effects of long-term unemployment benefits: Evidence from Germany's Hartz IV reform, *Working paper*.

Propper, C., Burgess, S. and Gossage, D. (2008). Competition and quality: Evidence from the NHS internal market 1991–9, *The Economic Journal* **118**(525): 138–170.

Rothe, T. and Wälde, K. (2017). Where did all the unemployed go? Non-standard work in Germany after the Hartz reforms, *IAB-Discussion Paper 18*, Institut für Arbeitsmarkt-und Berufsforschung (IAB).

Royston, P. (2008). Multiple imputation of missing values: Further update of ice, with an emphasis on interval censoring, *The Stata Journal* **7**(4): 445–464.

Schmidheiny, K. and Siegloch, S. (2019). On event study designs and distributed-lag models: Equivalence, generalization and practical implications, *CESifo Working Paper 7481*, CESifo Munich.

Stewart, M. B. (2002). Estimating the impact of the minimum wage using geographical wage variation, *Oxford Bulletin of Economics and Statistics* **64**: 583–605.

Weise, F.-J. (2011). Der Umbau der Bundesanstalt/Bundesagentur für Arbeit zum modernen Dienstleister, *Vierteljahrshefte zur Wirtschaftsforschung* **80**(1): 67–78.

Appendix

A Sample of Integrated Labor Market Biographies (SIAB)

The SIAB is a two percent random sample drawn from administrative social security records in Germany and made available by the Institute for Employment Research (IAB). It is representative of all individuals covered by the social security system (i.e. employees, benefit recipients, individuals officially registered as job-seeking or participating in programs of active labor market policies) and thus covers about 80 percent of the German workforce. The self-employed, civil servants, and individuals currently doing their military service are not included in the sample.

In the weakly anonymous version of the SIAB, which can be accessed via a research visit at the IAB or via remote data access only, the individuals' county of residence (since 1999) and place of work (since 1975) are available to the researcher. Therefore the data are well suited to compute wage inequality measures at the regional level. Note, that the SIAB data is structured in spells. In order to facilitate the analysis, I follow Eberle and Schmucker (2019) and create a cross-sectional data set, using June, 30^{th} as reference date.

Furthermore, daily wages in the SIAB are right-censored at the highest level of earnings subject to social security contributions. In order to impute the right-tail of the wage distribution, I follow Dustmann et al. (2009) and impute censored wages under the assumption that the error term in the wage regression is normally distributed with different variances for each age group, education group, and year (for more information on the imputation technique and assumptions made see Dustmann et al. (2009)).

Since the SIAB contains no information on hours worked, I restrict the sample to full-time employees. Moreover, I drop unrealistically low and high wages (i.e. daily wages below ten Euros or above the social security contribution limit). Adjusting for price changes, I again compute county-level inequality measures, using the information on the employee's county of residence.

B Additional Tables

Table B1: County Long-term Unemployment Rates in 2002—Summary Statistics

	(1) All	(2) West Germany	(3) East Germany
N	365	321	44
Mean	2.7	2.3	6
Sd	1.8	1.3	1.9
Min	.3	.3	2
P25	1.5	1.4	4.3
P50	2.2	2	5.6
P75	3.5	2.9	7.6
Max	9.9	7.5	9.9
Observations	6935	6935	6935

Notes: The long-term unemployment rate is measured in percent.

Table B2: County Shares of Social Assistance Recipients in 2002—Summary Statistics

	(1) All	(2) West Germany	(3) East Germany
N	363	319	44
Mean	2.8	2.8	2.8
Sd	1.6	1.6	1.4
Min	.4	.4	.9
P25	1.6	1.5	2
P50	2.5	2.6	2.4
P75	3.5	3.6	3.2
Max	10.1	10.1	7.8
Observations	6935	6935	6935

Notes: The share of social assistance recipients is measured in percent.

Table B3: Tax Reform 2001 and Income Inequality—DiD

	(1) Top 20%	(2) Top 15%	(3) Top 10%	(4) Mean Income
Treatment	-0.020	-0.031	-0.053	-0.001
	[0.315]	[0.210]	[0.108]	[0.310]
Demographic Controls	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes
Mean Dep. Variable	27.03	27.03	27.03	27.03
SD Dep. Variable	2.51	2.51	2.51	2.51
Mean Treatment Ind.	0.25	0.25	0.25	0.25
R^2	0.315	0.316	0.317	0.315
N	4015	4015	4015	4015

Notes: p-values in brackets; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; standard errors are clustered at the level of labor market regions.

Table B4: Hartz Reforms and Inequality—Event Study

	(1) Gini	(2) Gini	(3) Gini
Treatment × 1999	-0.030	-0.041	-0.035
	[0.219]	[0.371]	[0.462]
Treatment × 2000	-0.013	-0.051	-0.051
	[0.601]	[0.240]	[0.249]
Treatment × 2001	0.034	-0.006	-0.007
	[0.130]	[0.864]	[0.837]
Treatment × 2002	-0.011	-0.040	-0.039
	[0.560]	[0.194]	[0.200]
Treatment × 2004	0.054***	0.051*	0.055*
	[0.001]	[0.077]	[0.058]
Treatment × 2005	0.117***	0.072**	0.070**
	[0.000]	[0.036]	[0.045]
Treatment × 2006	0.097***	0.021	0.020
	[0.000]	[0.624]	[0.639]
Treatment × 2007	0.108***	0.063	0.059
	[0.000]	[0.136]	[0.147]
Treatment × 2008	0.155***	0.120***	0.113**
	[0.000]	[0.009]	[0.011]
Treatment × 2009	0.147***	0.082**	0.073*
	[0.000]	[0.036]	[0.054]
Treatment × 2010	0.178***	0.132***	0.119***
	[0.000]	[0.001]	[0.003]
Treatment × 2011	0.165***	0.112***	0.090**
	[0.000]	[0.006]	[0.025]
Treatment × 2012	0.167***	0.099**	0.071*
	[0.000]	[0.015]	[0.077]
Treatment × 2013	0.156***	0.116***	0.087**
	[0.000]	[0.008]	[0.043]
Treatment × 2014	0.111***	0.068	0.038
	[0.000]	[0.136]	[0.380]
Demographic Controls	No	No	Yes
County FE	Yes	Yes	Yes
Year FE	Yes	No	No
Region-Year FE	No	Yes	Yes
Mean Dep. Variable	27.06	27.06	27.06
SD Dep. Variable	2.50	2.50	2.50
Mean Treatment Ind.	8.60	8.60	8.60
R^2	0.275	0.278	0.290
N	5840	5840	5840

Notes: p-values in brackets; * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors are clustered at the level of labor market regions.

Table B5: Income Shares—DiD

	(1) 1. Decile	(2) 2. Decile	(3) 3. Decile	(4) 4. Decile	(5) 5. Decile	(6) 6. Decile	(7) 7. Decile	(8) 8. Decile	(9) 9. Decile	(10) 10. Decile
Treatment	-0.006	-0.015***	-0.020***	-0.017***	-0.016***	-0.009***	-0.002	0.008**	0.027***	0.050***
	[0.465]	[0.001]	[0.000]	[0.000]	[0.000]	[0.005]	[0.515]	[0.019]	[0.000]	[0.007]
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Variable	3.66	5.40	6.53	7.53	8.50	9.52	10.69	12.21	14.53	21.43
SD Dep. Variable	0.44	0.43	0.42	0.36	0.30	0.26	0.25	0.33	0.56	1.45
Mean Treatment Ind.	8.60	8.60	8.60	8.60	8.60	8.60	8.60	8.60	8.60	8.60
R^2	0.079	0.312	0.356	0.333	0.245	0.132	0.075	0.106	0.183	0.244
N	4745	4745	4745	4745	4745	4745	4745	4745	4745	4745

Notes: p-values in brackets; * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors are clustered at the level of labor market regions.

Table B6: Alternative Treatment Indicator: Share of Long-term Unemployed among all Unemployed

	(1) Continuous	(2) Binary 50%	(3) Binary 33%	(4) Binary 25%
Treatment	0.034***	0.244*	0.540***	0.666***
	[0.002]	[0.091]	[0.003]	[0.001]
Demographic Controls	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes
Mean Dep. Variable	26.98	26.98	26.88	26.85
SD Dep. Variable	2.51	2.51	2.37	2.28
Mean Treatment Ind.	29.50			
△ Mean Treatment Ind.		11.32	15.53	18.05
R^2	0.322	0.317	0.327	0.333
N	4745	4745	3146	2392

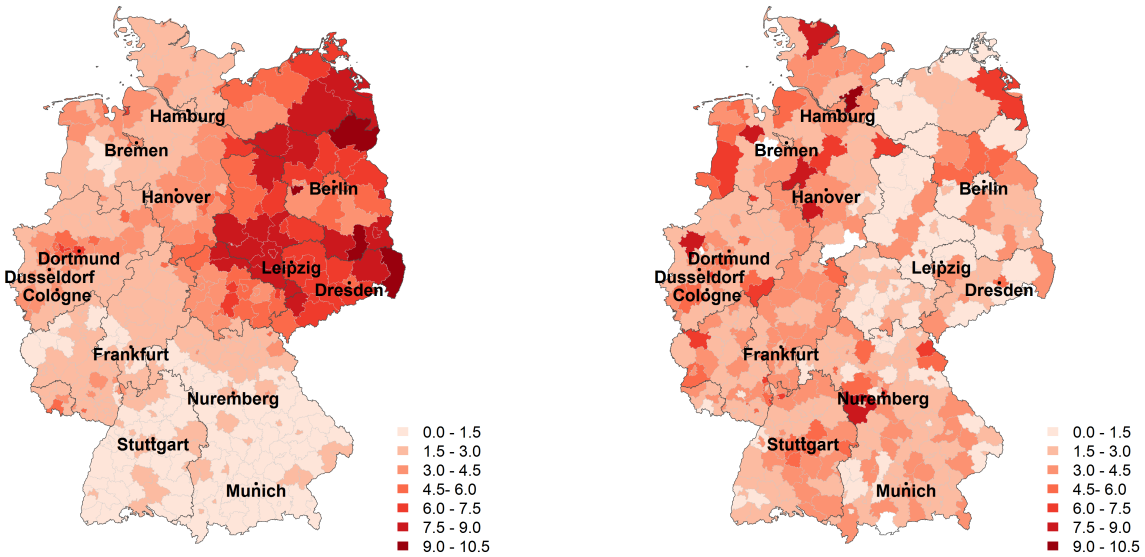
Notes: p-values in brackets; * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors are clustered at the level of labor market regions.

C Additional Figures

Figure C1: Alternative Treatment Indicators across German Counties in 2002

(a) Long-term Unemployment Rate across Counties in 2002

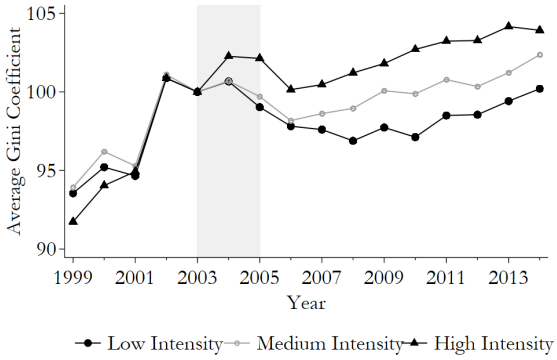
(b) Share of Social Assistance Recipients across Counties in 2002



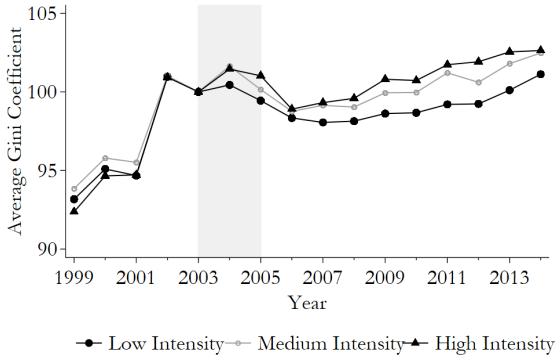
Notes: The figure shows long-term unemployment rates and share of social assistance recipients across German counties in 2002. Long-term unemployment rates and shares of social assistance recipients are measured in percent.

Figure C2: Alternative Treatment Indicators: Visualization of the Parallel Trend Assumption

(a) Long-term Unemployment Rate



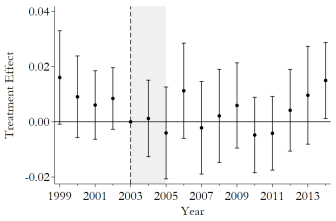
(b) Share of Social Assistance Recipients



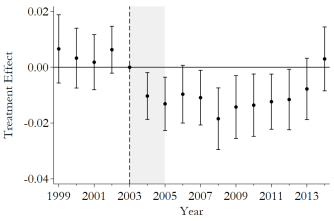
Notes: The figure shows the trend of the average Gini coefficients by treatment intensity group from 1999 to 2014. Gini coefficients are normalized to 2003=100.

Figure C3: Income Shares—Event Study

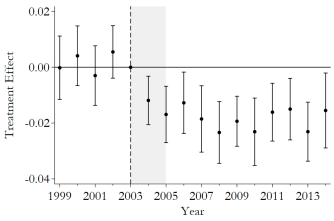
(a) 1. Decile



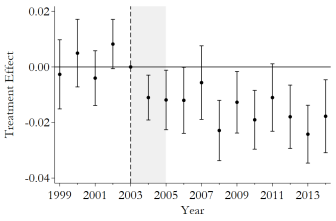
(b) 2. Decile



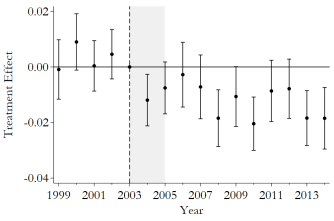
(c) 3.Decile



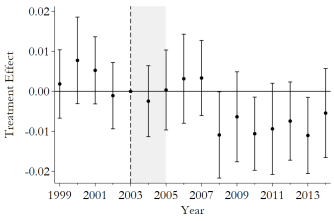
(d) 4. Decile



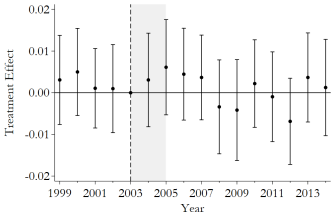
(e) 5. Decile



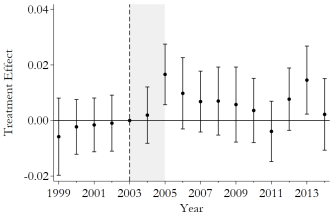
(f) 6. Decile



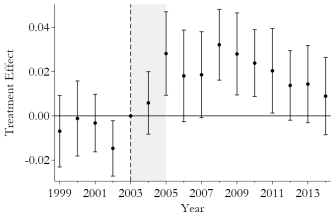
(g) 7. Decile



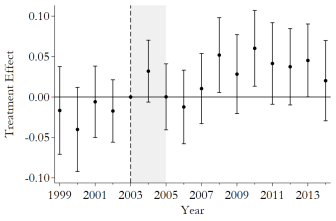
(h) 8. Decile



(i) 9. Decile

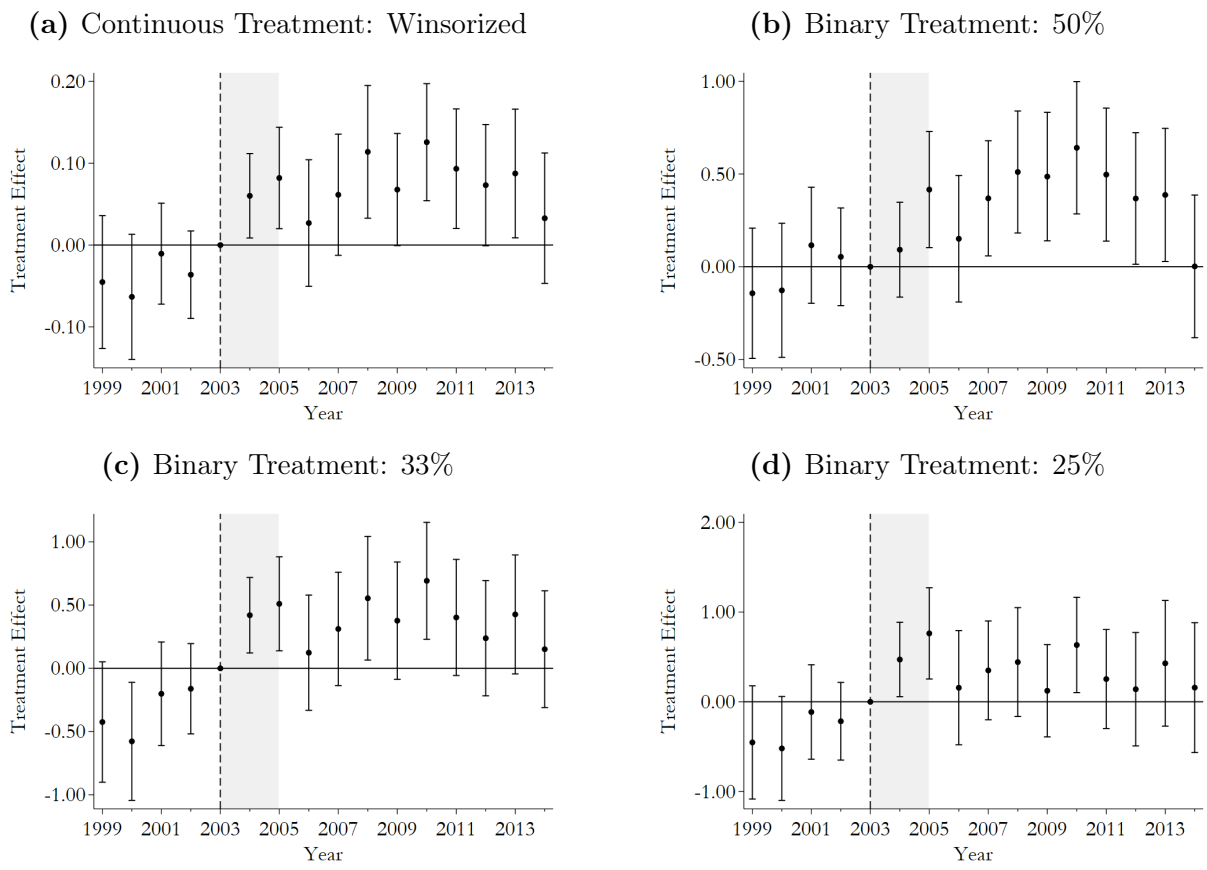


(j) 10. Decile



Notes: The figure plots the event study results. Circles represent point estimates, black lines represent 90 percent confidence intervals. Standard errors are clustered at the level of labor market regions.

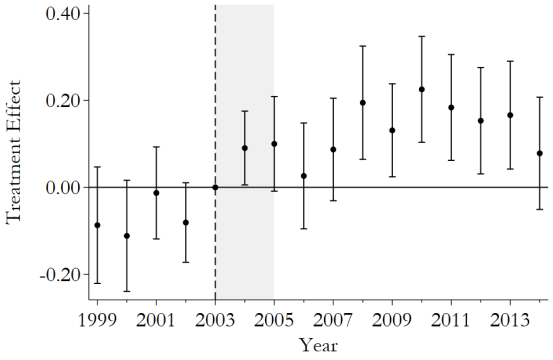
Figure C4: Alternative Treatment Definitions—Event Study



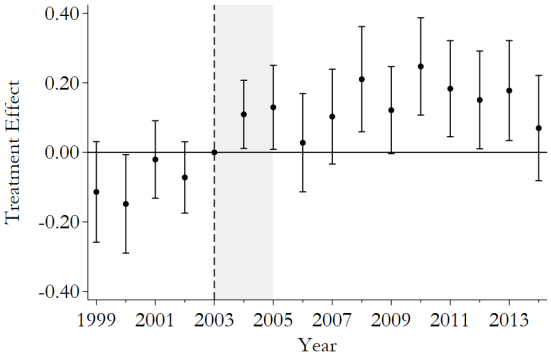
Notes: The figure plots the event study results. Circles represent point estimates, black lines represent 90 percent confidence intervals. Standard errors are clustered at the level of labor market regions.

Figure C5: Alternative Treatment Indicators: Long-term Unemployment Rate—Event Study

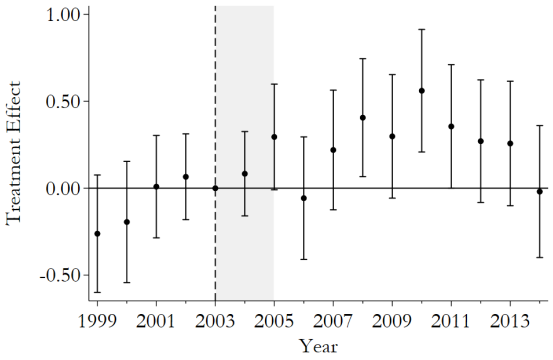
(a) Continuous Treatment



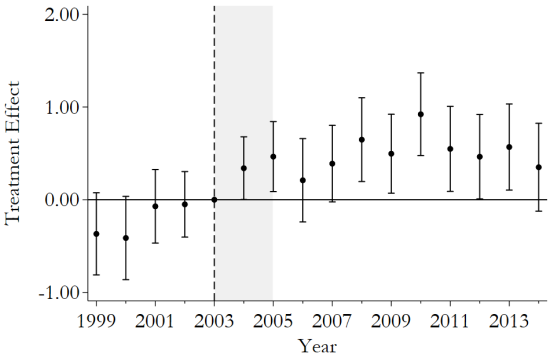
(b) Winsorized



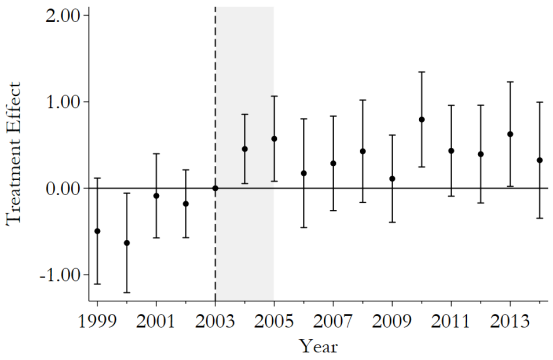
(c) Binary Treatment: 50%



(d) Binary Treatment: 33%



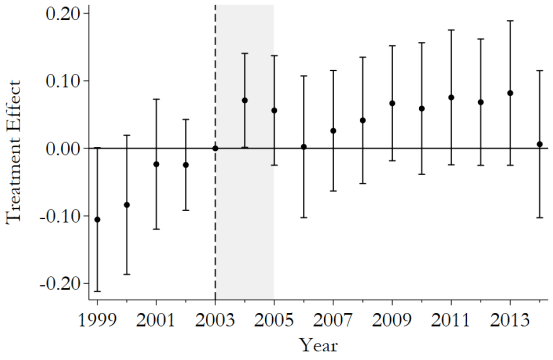
(e) Binary Treatment: 25%



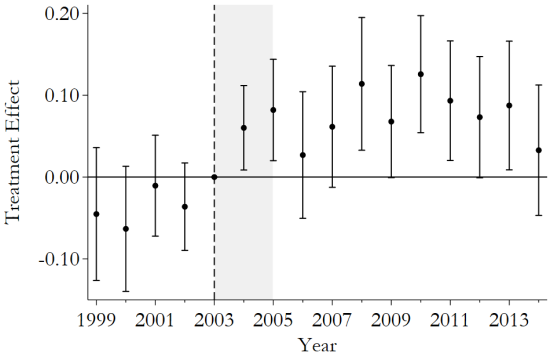
Notes: The figure plots the event study results. Circles represent point estimates, black lines represent 90 percent confidence intervals. Standard errors are clustered at the level of labor market regions.

Figure C6: Alternative Treatment Indicators: Social Assistance Recipients—Event Study

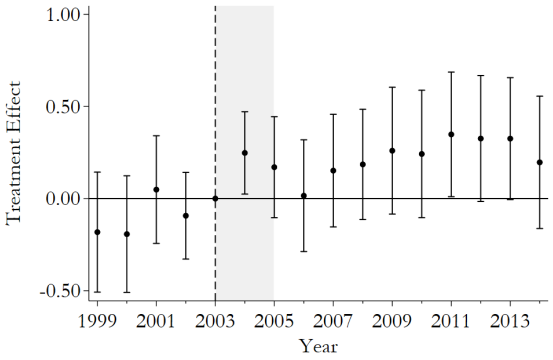
(a) Continuous Treatment



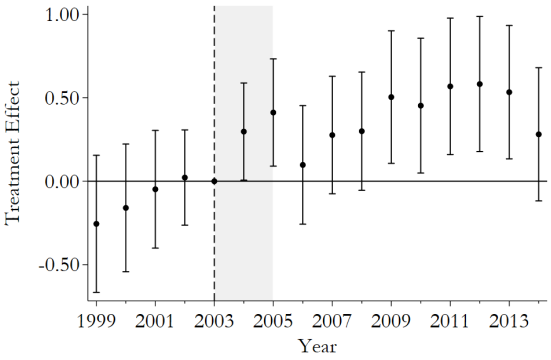
(b) Winsorized



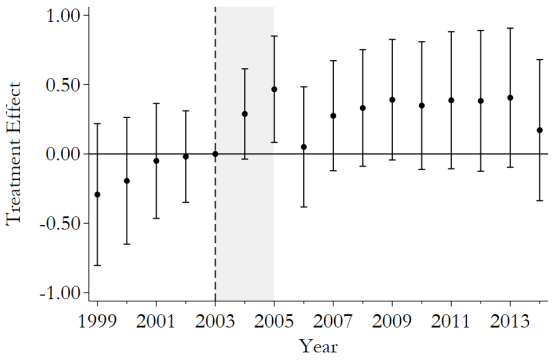
(c) Binary Treatment: 50%



(d) Binary Treatment: 33%

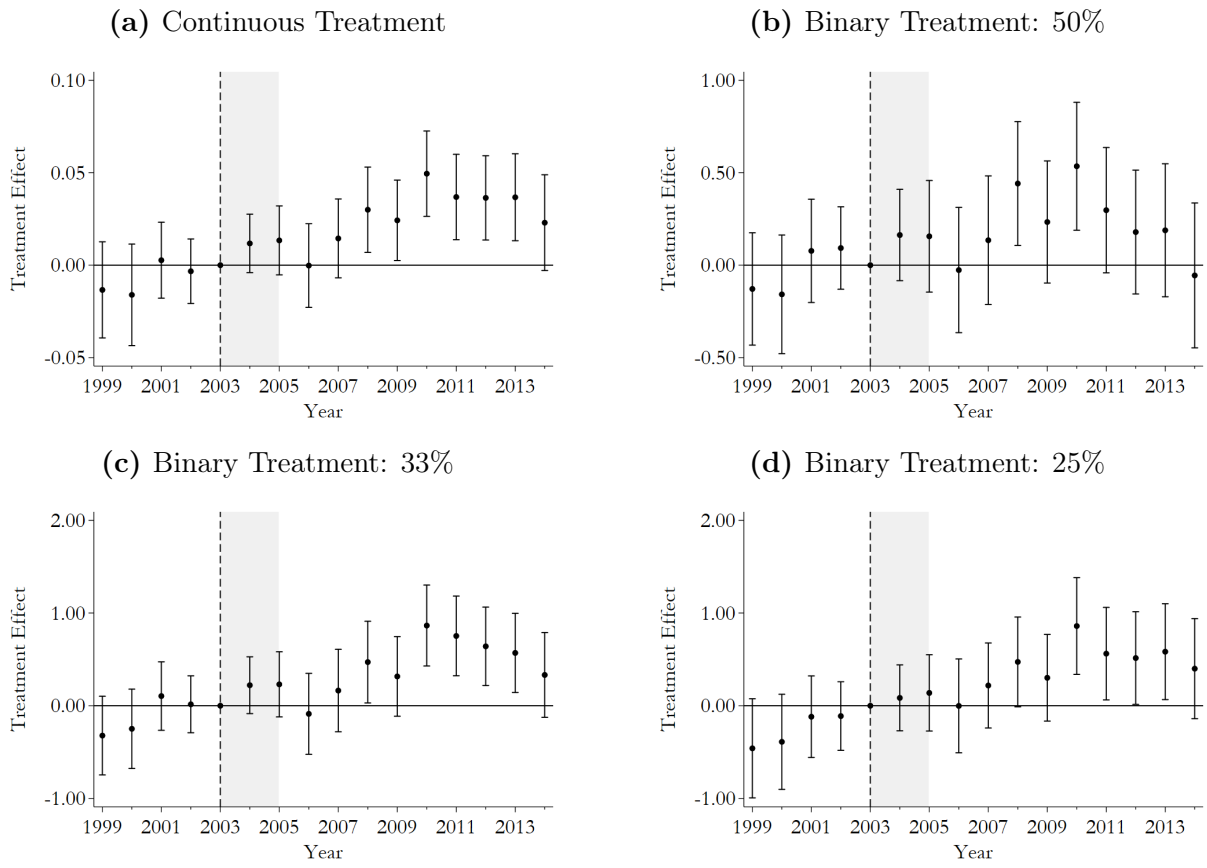


(e) Binary Treatment: 25%



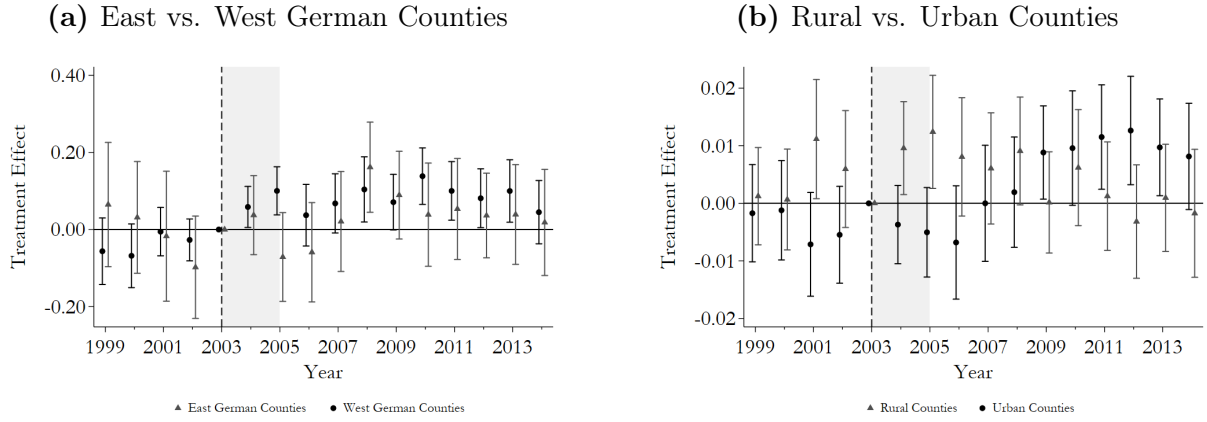
Notes: The figure plots the event study results. Circles represent point estimates, black lines represent 90 percent confidence intervals. Standard errors are clustered at the level of labor market regions.

Figure C7: Alternative Treatment Indicators: Share of Long-term Unemployed among all Unemployed—Event Study



Notes: The figure plots the event study results. Circles represent point estimates, black lines represent 90 percent confidence intervals. Standard errors are clustered at the level of labor market regions.

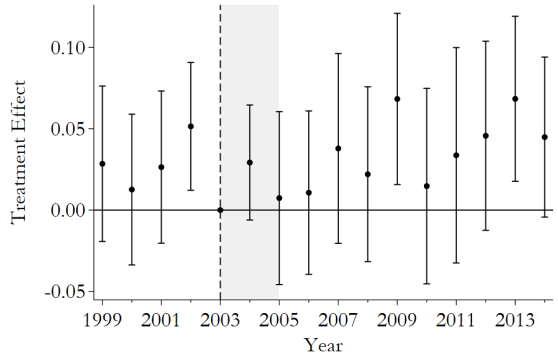
Figure C8: Heterogeneous Effects—Event Study



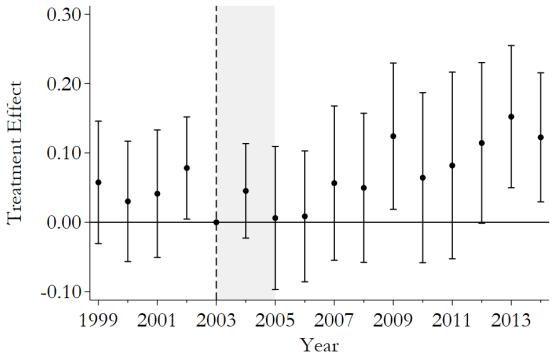
Notes: The figure plots the event study results. Circles represent point estimates, black lines represent 90 percent confidence intervals. Standard errors are clustered at the level of labor market regions.

Figure C9: Full-time Wage Inequality—Event Study

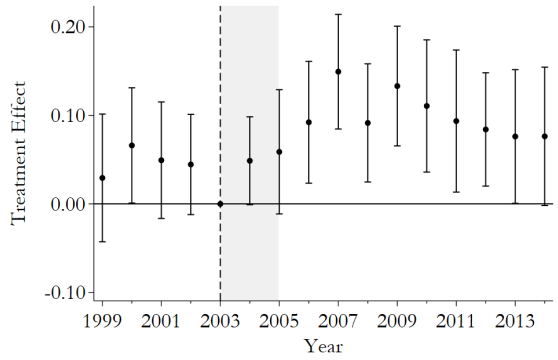
(a) Continuous Treatment: Unemployment



(b) Continuous Treatment: Long-term Unemployment



(c) Continuous Treatment: Social Assistance



Notes: The figure plots the event study results. Circles represent point estimates, black lines represent 90 percent confidence intervals. Standard errors are clustered at the level of labor market regions.