

# The economic incidence of social security contributions

A discontinuity approach with linked employer-employee data

Kai-Uwe Müller<sup>\*</sup>, Michael Neumann<sup>‡</sup>

February 14, 2018

## Abstract

We estimate economic incidence of social security contributions (SSC) on the basis of cross-sectional earnings distributions. The approach exploits discontinuities in earnings distributions at kinks in the budget set which are informative about tax incidence. Contrary to most research on SSC incidence, it does not rely on policy reforms, panel data, or hours information. When the location of kinks does not change significantly, estimates represent equilibrium incidence and are less affected by short-run adjustment frictions than results based on policy reforms. We thoroughly discuss the framework proposed by Alvaredo and Saez (2007) including identifying assumptions and related problems in empirical applications. We also suggest parametric and non-parametric estimators. The approach is applied to earnings caps of SSC in Germany where the marginal SSC rate drops to zero. We use linked employer-employee data providing precise measures of gross earnings and actually paid contributions. Therefore we can analyze gross as well as net earnings which as is argued improves identification. We find substantial negative discontinuities at most earnings caps of SSC in the distribution of net earnings. Together with smooth gross earnings distributions around the caps this provides consistent empirical evidence that legal and economic incidence of SSC coincide.

Keywords: incidence; social security contributions; discontinuities; linked employer-employee data

JEL classification: H22; J38; H55; J20

<sup>\*</sup> Corresponding author, German Institute for Economic Research Berlin (DIW), Mohrenstr. 58, D-10117 Berlin, Tel.: +49 30 89789 154, Fax: +49 30 89789 114, E-mail: kmueller@diw.de

<sup>‡</sup> German Institute for Economic Research Berlin (DIW), Mohrenstr. 58, D-10117 Berlin, Tel.: +49 30 89789 353, Fax: +49 30 89789 114, E-mail: mneumann@diw.de

# 1 Introduction

In many countries payroll taxes or social security contributions (SSC) account for a material part of total taxation. Their average contribution to government funding among OECD countries is similar to personal income taxation (around 9% of total tax revenues, OECD, 2016). With earnings from employment usually being the sole tax base, the potential of SSC to distort efficient allocations on the labor market is large. Who bears the burden of SSC is thus a crucial question for distributional or welfare analysis.

By contrast to income taxation SSC are formally shared between employers and employees. According to standard economic theory legal incidence does not matter for the effective sharing of the tax burden though (invariance of incidence proposition). The SSC burden is primarily borne by employees because labor demand is considered to be more elastic than supply (Fullerton and Metcalf, 2002). More recent empirical research challenges the standard view and finds economic incidence to be identical with formal incidence (Saez et al., 2012). Overall the evidence on the burden-sharing of SSC is inconclusive (Melguizo and Gonzalez-Paramo, 2013). This might be partly explained by the variety of estimation approaches and their underlying assumptions.

The majority of the empirical literature on the economic incidence of SSC relies on longitudinal variation. Most earlier observational regressions lack clean identification, though. More credible approaches exploit exogenous variation generated by policy reforms (Gruber, 1994, 1997; Benmarker et al., 2009; Saez et al., 2012). Suitable SSC reforms are infrequent, however. Only one policy change occurred, for instance, in Germany which created sufficient institutional variation and provided a setting with a valid control group (Neumann, 2015). Except for Saez et al. (2012) estimates from these studies are based on a difference-in-differences framework. They rely on the common-trends assumption, depict short-term responses and are biased towards zero under optimization frictions. Data restrictions exacerbate these limitations. Should hours of work not be observed – which is common in administrative data sets – behavioral responses (of labor supply and demand) to SSC have to be assumed away to interpret estimates in terms of incidence. Given the incomplete information, this presumption cannot be tested.

In this paper we therefore follow a different approach suggested by Alvaredo and Saez (2007). The alternative framework is based on cross-sectional data and utilizes earnings caps in a given SSC schedule. The discontinuous drop of the marginal SSC rate generates a downward kink in the average tax schedule. Depending on how the SSC burden is shared, a positive or negative discontinuity emerges in the distribution of gross earnings at the cap. Only when statutory equals economic incidence, the gross earnings distribution is smooth around the cap. This can be exploited to estimate economic incidence. The approach circumvents some of the aforementioned methodological issues. It neither requires exogenous policy reforms nor panel data or information on hours of work. Estimates represent long-term incidence effects as long as earnings caps do not change markedly. Adjustment frictions are thus less of a problem here.

Our paper is the first application of the approach for non-Spanish data. The main contribution is thus to provide fresh evidence on economic incidence of SSC based on an approach which facilitates estimating long-term incidence. A second contribution is that we thereby elaborate on the identifying assumptions and discuss the empirical implementation using parametric as well as non-parametric estimators. A third contribution of this paper is the application of linked employer-employee data (the German Structure of Earnings Survey, GSES) allowing for a more robust analysis. Discontinuity estimates from a single earnings distribution are hard to interpret under measurement error. A roughly smooth gross earnings distribution might represent a substantial effect or result from noisy data. The GSES includes a separate measure of SSC paid by employees.<sup>1</sup> We argue that this allows for distinguishing between the two competing interpretations.

We find at large very small discontinuities in empirical gross earnings distributions. Together with substantial negative discontinuities in net earnings this provides consistent empirical evidence that legal and economic incidence of SSC coincide in Germany. We demonstrate that the findings are robust with respect to methodological decisions, different caps and sample periods.

This finding is in line with recent evidence. Saez et al. (2012) evaluate a unique policy reform in Greece where a different SSC regime was implemented for employ-

---

<sup>1</sup>Saez et al. (2012) also observe the exact amount of SSC in their data for Greece. They use a different framework based on a cohort-specific policy reform.

ees who started working from 1993 onwards. They do not find gross earnings of otherwise similar workers, i.e. of those who entered the labor market shortly before and after the reference date, to differ systematically. This implies that legal and economic incidence coincide. Skedinger (2014) finds small effects of payroll tax cuts for young Swedish workers. In Finland a similar reform targeted at older low-wage workers did not trigger any wage effects (Huttunen et al., 2013). These results contradict earlier studies for non-European countries which found economic incidence to be entirely with workers (Gruber, 1994, 1997). Our study provides further evidence that employers are hardly able to shift their SSC burden to employees under a more or less centralized wage setting regime as it exists in Germany and other European countries.

The remainder of this paper is organized as follows. Section 2 presents the methodological framework, discusses the underlying assumptions for identification and the empirical implementation of the approach. Section 3 applies the approach to German linked employer-employee data. Section 4 discusses the results in the light of previous findings and concludes.

## **2 Methodology**

In this section we, first, introduce the analytical framework developed by Alvaredo and Saez (2007). Second, we spell out the identifying assumptions. Third, we present parametric and non-parametric estimators to implement the discontinuity approach empirically.

### **2.1 Analytical framework**

We start with a narrow framework centered around a general notion of economic incidence captured by a change in the hourly gross wage rate. The model does not describe any incidence mechanisms. Underlying labor supply and demand elasticities which drive economic incidence in the standard model are, for example, not explicitly included. Employees' and employers' taxes besides SSC are also not covered. The latter does not (qualitatively) affect the analysis as long as additional taxes do not vary systematically around the SSC caps.

Let  $t$  denote the symmetric SSC rate which has to be paid by employers ( $r$ ) and employees ( $e$ ), i.e.  $t = t_r = t_e$ . Assume ability or preference for work  $n$  to follow a smooth cumulative (marginal) distribution function denoted by  $P(n)$  ( $p(n)$ ). For simplicity let realized gross earnings  $y$  equal ability, i.e.  $y = n$ , if either the SSC rate  $t$  is zero or legal and economic incidence coincide. Should legal and economic incidence differ,  $y$  is distorted such that

$$y = \frac{n}{1 + ts} \quad (1)$$

with  $s$  being a homogeneous shifting parameter which is positive (negative) if employers (employees) shift some burden to employees (employers). This framework can be generalized along several dimensions. Instead of using a symmetric SSC rate  $t$ , the rates might differ between employers and employees, i.e.  $t_r \neq t_e$  (Appendix A.1). Likewise shifting  $s$  can be asymmetric between employers and employees (Appendix A.2), or  $s$  can be heterogeneous across different employees and employers (Appendix A.3). As shown in Appendix A neither of those extensions eliminates a discontinuity at an earnings cap.

We therefore stick to the parsimonious specification in equation (1) which is not restrictive for SSC institutions in most countries, such as Germany. First, statutory SSC rates for employers and employees are often identical. This holds for the period under observation (1995-2010) in this paper, too. Second, earnings caps apply to employers as well as employees. Separate identification of  $s_r$  and  $s_e$  with varying SSC rates requires differential variation in employers' and employees' SSC rates.

An earnings cap for SSC at  $\bar{y}$  where  $t$  drops to 0 yields the following relationships between the distributions of ability and realized earnings ( $F(y), f(y)$ ):

$$F(y) = P(n) = \begin{cases} P(y(1 + ts)) & \forall y \leq \bar{y} \\ P(y + \bar{y}ts) & \forall y > \bar{y} \end{cases} \quad (2)$$

$$f(y) = \begin{cases} p(y(1 + ts))(1 + ts) & \forall y \leq \bar{y} \\ p(y + \bar{y}ts) & \forall y > \bar{y} \end{cases} \quad (3)$$

Considering the left and right limit of the density of realized gross earnings when  $y$

approaches  $\bar{y}$ :

$$f(\bar{y})_- = p(\bar{y}(1 + ts))(1 + ts) = p(\bar{y} + \bar{y}ts)(1 + ts) = p(\bar{n})_-(1 + ts) \quad (4)$$

$$f(\bar{y})_+ = p(\bar{y} + \bar{y}ts) = p(\bar{n})_+ \quad (5)$$

As  $p(n)$  is assumed to be smooth, there will be some discontinuity in the density  $f(y)$  as long as  $s \neq 0$ , i.e. economic differs from legal incidence. The density drops (jumps), if  $s > 0$  ( $s < 0$ ), that is if employers (employees) are able to shift some of their legal burden. Full shifting to employees (employers) implies  $s = +(-)1$ . The size of the discontinuity at the cap is equivalent to the share of the SSC rate shifted ( $ts$ ) in relative terms to (i.e. in per cent of) the density directly above:

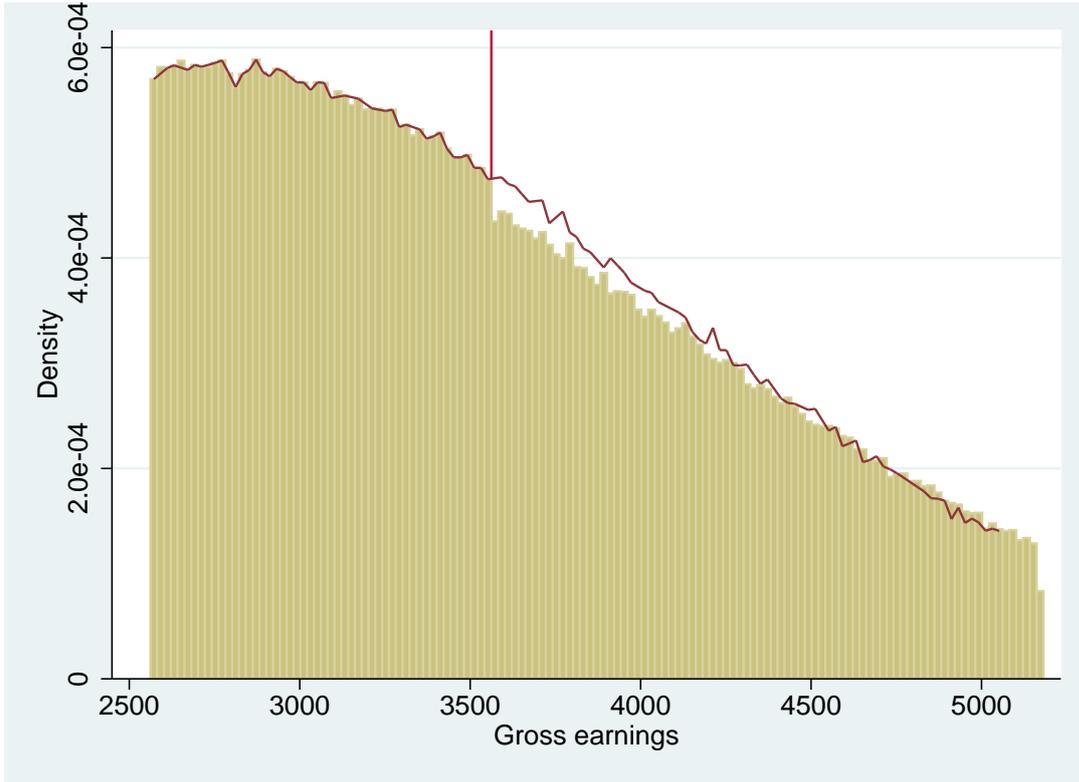
$$\frac{f(\bar{y})_- - f(\bar{y})_+}{f(\bar{y})_+} = ts \quad (6)$$

Therefore  $s$  can be expressed as a function of observed measures. To put it more intuitively: Depending on their bargaining power and other factors like wage rigidities, employers or employees will shift some of their SSC burden. This manifests itself in a negotiated wage which is influenced by the average SSC rate. At an earnings cap the previously flat average SSC schedule features a downward kink. From that point the SSC burden decreases relative to earnings. Under constant shifting the wage rate is adjusted to a lesser degree with increasing distance of earnings to the cap. This implies that the gross earnings distribution expands (compresses) to the right of the cap when incidence is rather with employees (employers). A negative (positive) discontinuity emerges at the cap.

To demonstrate this, the gross earnings distribution  $f(y)$  is simulated on the basis of equation (3) under the assumption of full shifting of employers' SSC to employees ( $s = 1$ , see Appendix B for more details). Given our assumptions the simulated earnings densities feature a negative discontinuity at  $\bar{y}$  in the magnitude of the drop in the marginal SSC rate of roughly 7.6% (bars in Fig. 1).

The deviation between the actual and a counterfactual density without earnings cap (additional curve in Fig. 1) is not uniform over the area above the cap. For a negative (positive) slope the gap diminishes (increases) slowly with the distance to  $\bar{y}$ . To see the reason, consider both densities at the earnings level  $\tilde{y} > \bar{y}$  (equations

**Fig. 1:** Distributions of simulated gross earnings with negative discontinuity at  $\bar{y}$



Source: Own simulation.

(7) and (8)).

$$f(\tilde{y})_{\text{Cap}} = p(\tilde{y} + \bar{y}ts) \quad (7)$$

$$f(\tilde{y})_{\text{No cap}} = p(\tilde{y}(1 + ts))(1 + ts) = p(\tilde{y} + \tilde{y}ts)(1 + ts) \quad (8)$$

For a monotonously negative (positive) slope of the ability density above  $n = \bar{y}$  it holds that  $p(\tilde{y} + \bar{y}ts) > (<) p(\tilde{y} + \tilde{y}ts)$ . The inequality becomes larger with the distance to  $\bar{y}$ . At the same time, for  $s \neq 0$ , the factor  $(1 + ts)$  constantly increases the density in the case of no earnings cap. Assuming a negative slope, this implies that the deviation between both densities is at its maximum directly at the cap and diminishes slowly until it even might reverse (Fig. 1). For a positive slope, the deviation increases with the distance to  $\bar{y}$ .

A change in the marginal SSC rate might induce behavioral reactions, for instance an adjustment of working hours. This is particularly relevant as in most models of tax incidence, labour supply and demand elasticities are crucial determinants of economic incidence. The fact that we do not explicitly model hours responses implies that we are agnostic about potential mechanisms of economic in-

vidence. While this is not problematic per se, hours responses could have effects on the shape of the distribution around an earnings cap (Saez, 2010). Identification of a discontinuity induced by economic incidence then becomes more difficult albeit in many cases not impossible (Appendix A.4). Yet, there is no evidence for behavioral adjustments at earnings caps in the literature (Liebman and Saez, 2006; Alvaredo and Saez, 2007; Saez, 2010). Neither does the observed distribution of gross earnings suggest any hours reactions in our application (see below).

## 2.2 Identification

The emergence of a discontinuity in the density of gross earnings (net earnings or labor costs) which is proportional to the drop in the marginal SSC rate at an earnings cap requires a number of conditions to be met. A closer look at those assumptions is helpful to sort out different identification problems when we apply the framework to actual data. Assuming smoothly distributed abilities or preferences for work is a technical requirement. This is not restrictive in applications and will thus not be discussed further.

A first set of assumptions concerns the *shifting mechanism*. We assume that the burden of SSC is shifted at the individual level through changes in the wage rate. This implies, first, that incidence effects at a more aggregate level - e.g. firms' shifting the burden to their whole workforces or shifting at the sectoral level - would invalidate the approach as individual earnings below and above the cap would not be differentially affected by a cap and no discontinuity would emerge. Second, the burden of SSC must not be shifted through other margins of employees' compensation (e.g. premia, non-pecuniary benefits, or paid/unpaid overtime).

We further need to assume that SSC rates and caps are perfectly salient. If this is not the case, (some) employers and employees might either ignore SSC when bargaining over wages, or perceived rates and earnings caps vary over individuals. In both instances a discontinuity could be blurred or even eliminated. A similar effect have optimization frictions preventing wages to adjust according to SSC liabilities. This also holds for re-negotiations of wages, e.g. after an institutional change in the earnings cap or an employee's promotion which implies a pay increase above an earnings cap. While virtually all microeconomic evidence on economic incidence

relies on these assumptions, some arguments can be brought forward that certain assumptions are less critical for our approach and setting.

As the German earnings caps are located relatively high in the earnings distribution individual wage negotiations are more likely. The salience assumption is underpinned by the fact that earnings caps are relevant for employers as well as employees. Employers should be particularly aware of their SSC liabilities when calculating labor costs, even more so when they have several employees with earnings above the cap. Salience is plausible for the German setting where earnings caps have been in existence for a long time and have rarely been changed significantly (sub-section 3.1). The persistence of SSC institutions also facilitates the no-frictions assumption. If internalizing the reduced SSC liability at the cap is costly, it takes only place when benefits exceed those costs. The drop in the average SSC rate slightly above the cap and hence the benefits of a reaction are indeed relatively small. On the other hand, when SSC institutions are persistent and salient wage negotiations – which for example lead to the crossing of the cap – can directly incorporate the drop of the marginal SSC rate. In that sense the framework where identification does not rely on policy reforms captures long- and not merely short-term effects and is thus less affected by optimization frictions<sup>2</sup>.

## 2.3 Estimation

It is straightforward to bring this framework to the data: The size of a potential discontinuity has to be determined at the cap of a given earnings distribution. We follow two estimation strategies, a non-parametric sorting test proposed by McCrary (2008) and a parametric approach where a polynomial is fitted to the distribution. The magnitude of the shifting  $ts$  is calculated by normalizing the discontinuity estimate by the density directly right of the cap.

In the first step of the McCrary test a histogram is estimated where no bin contains observations below and above a potential discontinuity. Then the density to the left and right of the cap is estimated by local linear regressions to avoid bias

---

<sup>2</sup>Caps are adjusted regularly according to the change in the average gross wage bill in the preceding year. These minor changes might make a complete internalization of legal SSC liabilities impossible but carry minor weight as they go along with the overall shift in the earnings distribution.

at the boundaries. The test statistic consists of the log difference in the prediction slightly left and right of the cap. The procedure is subject to the choice of the bin size for the underlying histogram estimates, the selection of the bandwidth for the local linear smoother as well as the earnings interval underlying estimation.

In an alternative parametric approach we also estimate initially a histogram of the earnings distribution. Then a polynomial of degree five is fitted to the density values at the center of each histogram bin (equation (9)). The discontinuity is measured by an indicator variable differentiating between bins above and below the cap. We restrict the coefficients to be identical on both sides of the cap to increase the precision of the fit for values close to the cap. This is justifiable as the potential discontinuity stretch across the entire distribution above the cap<sup>3</sup>.

$$d_j = \sum_{p=1}^9 \beta_p y_j^p + \delta \mathbb{1}(y_j > \bar{y}) + \epsilon_j \quad (9)$$

Here  $d_j$  denotes the height of bin  $j$ ,  $y_j$  are earnings in the middle of bin  $j$ ,  $\mathbb{1}$  is an indicator function,  $\beta$  and  $\delta$  are parameters, and  $\epsilon$  is an error term. While the non-parametric test is not subject to functional form assumptions about the earnings distribution, it might be more sensitive to outliers close to the discontinuity. By allocating the same weights to all bins in the estimation window the parametric approach is more robust in that respect. The fact that identification is less dependent on individuals very close to the cap also decreases the diluting effect of optimization frictions (sub-section 2.2). In the basic specification we use an estimation interval of 2000 € which is split into 100 bins on either side of the cap. This implies a bin size of 10 €.

We provide some Monte-Carlo evidence on the performance of these estimators (Appendix B). Gross earnings distributions  $f(y)$  are simulated on the basis of equation (3) under the assumption of full shifting of employers' SSC to employees ( $s = 1$ ). Given our assumptions the simulated earnings densities feature a negative discontinuity at  $\bar{y}$  in the magnitude of the drop in the marginal SSC rate of roughly 7.6% (Fig. 1 in the Appendix). The McCrary test and the parametric test identify a discontinuity of -7.2% and -7.6%, respectively, which matches the underlying

---

<sup>3</sup>There is a small change in the curvature, though (section 2.1). We therefore estimate a more flexible polynomial as a robustness check.

parameters  $s$  and  $t$  (Fig. B1 in the Appendix). The confidence interval of the McCrary estimate translates to an interval of  $(-1.24, -0.64)$  for the shifting parameter  $s$ . The corresponding interval for  $s$  based on the parametric test is slightly narrower:  $(-1.21, -0.79)$ .

## 3 Application to German data

### 3.1 Institutions

The German social security system consists of pension, unemployment, health, and long-term care insurance. The contribution rates are flat with daily gross earnings as tax base. SSC have formally been shared equally between employees and employers until the end of 2004. Since then a share of 0.9 pp. are paid exclusively by employees.<sup>4</sup> Most SSC rates have been quite constant. The total SSC rate varied around 40% with pension (around 20%) and health insurance (around 14%) as the most important branches (Fig. C1 in the Appendix). Marginal SSC rates only apply up to earnings caps. There is one threshold for pension and unemployment insurance, another one for health and long-term care insurance. For the sake of readability we do not separately refer to unemployment and long-term care insurance throughout the paper. Both caps differed between East and West Germany until 2001 when the health insurance cap in East Germany was adjusted to the level of West Germany (Fig. C2 in the Appendix).

SSC rates as well as earnings caps are subject to yearly gradual changes which are difficult to exploit for the identification of economic incidence. Considerable discontinuous changes are rare with the strong increases of the earnings cap of health insurance in East Germany in 2001 and pension insurance in 2003 as notable exceptions. While the former is evaluated by Neumann (2015), the latter is difficult to analyze because suitable German panel data is right-censored at the pension earnings cap. The cross-sectional approach outlined above does not rely on policy reforms and panel data. We are therefore able to draw on uncensored earnings data here. The caps are at different positions in the respective earnings distributions. The threshold for health insurance in West Germany is around the 75th quantile

---

<sup>4</sup>This share increases by 0.25 pp. for childless employees.

and by far the lowest (Tab. 1). The health cap in East Germany and the pension cap in West Germany come in second at around the 90th quantile. The pension cap in East Germany is up high in the earnings distribution above the 95th quantile.

**Tab. 1:** Quantiles of earnings caps

Wave	West Germany		East Germany	
	Pension / Unemployment	Health / Long-term care	Pension / Unemployment	Health / Long-term care
1995	.91	.76	.97	.89
2001	.90	.76	.96	.93
2006	.92	.75	.96	.91
2010	.90	.72	.95	.89

*Source:* GSES 1995, 2001, 2006, 2010; own calculations.

One peculiarity of the German social security system which could affect our empirical analysis is the possibility to substitute public for private health insurance. Eligibility for private health insurance depends on exceeding the income threshold for compulsory insurance (called *Versicherungspflichtgrenze*), except for the self-employed and civil servants who are excluded here. This threshold had been equal to the earnings cap of health insurance until both caps were decoupled in 2003. Since then the threshold for compulsory insurance lies between the earnings caps of health/care and pension/unemployment insurance. The employees' share of SSC does not depend on earnings but on personal characteristics under private insurance. Employers have to pay half of private contributions but only up to the maximum value of employers' SSC under public insurance.

Private health insurance interferes with our analysis of gross earnings when it affects the change in the marginal SSC rate at the analyzed earnings cap. This is only the case for the earnings cap of health insurance before 2003.<sup>5</sup> Being slightly below or above this earnings cap might change the insurance system applicable to the employee. Although we do not observe whether an employee is privately health insured, it is not obvious how severely this distorts a potential discontinuity at that cap. First, the cap for employers' SSC is in most cases identical under public and private insurance. Their incentives below vs. above the threshold are comparable

<sup>5</sup>As contributions to private health insurance are included in the GSES' measure of employees' SSC, the impact on the analysis of observed net earnings is more severe. See sub-section 3.2 for a detailed discussion.

as long as private is at least as expensive as public health insurance which normally holds (Neumann, 2015). Second, the switch to private health insurance usually occurs earliest in the first year after earnings increased above the threshold. A large fraction of employees with earnings very closely above the earnings cap, i.e. a large fraction of our estimation sample, are still publicly health insured in a given year.

## 3.2 Data

We use the German Structure of Earnings Survey (GSES, *Verdienststrukturerhebung*) which consists of repeated cross-sections of employees. It is representative for individuals employed in October in firms with more than ten employees<sup>6</sup>. The information is provided by employers and is part of the official labor cost statistics of the German Statistical Office. Firms are therefore obliged to cooperate and provide the information. We primarily use monthly gross earnings information that also refers to the month of October. The GSES separately includes all earnings components which are subject to SSC (like regular earnings and compensation for overtime hours) as well as some of those which are not (like tax-free premia for shift work, working on Saturdays/Sundays/holidays, or night employment) which allows for exactly locating the kink in an individual's budget set induced by earnings caps.

We further use information on actually paid employees' SSC including an employee's contributions to social insurance (i.e. the employee's compulsory and voluntary share of contributions to the pension, unemployment, health and care insurance). Together with information on gross earnings  $y$  we can calculate actual individual net earnings  $c$ . There is no independent information on employers' SSC; labor costs  $z$  are thus not considered separately here. The measure of actually paid employees' SSC also entails contributions to private health insurance (section 3.1) as well as occupational insurance schemes. As we do not observe whether an employee is privately health insured a potential discontinuity in the net earnings distribution at the earnings cap might be attenuated<sup>7</sup>.

---

<sup>6</sup>Few branches are also missing for certain waves. Neither of these omissions is systematically related to SSC and earnings caps and affecting our estimates.

<sup>7</sup>The threshold for compulsory public insurance exceeds the earnings cap of health insurance since 2003, though (see section 3.2). The estimates based on the health insurance cap in 2006 and 2010 are therefore not distorted by private health insurance.

We use four waves of data: 1995, 2001, 2006, and 2010. We exclude civil servants and self-employed as they do not contribute to the general SSC scheme. Home workers are also not included in the sample. We also exclude part time and marginally employed persons to limit behavioral adjustments. The majority of these employees has earnings below the caps, anyway. The final sample size is comparably large: For the wave 1995 more than 700,000 observations are available, about 640,000 for the wave 2001, more than 1.2 million for 2006, and more than 1 million employees for the wave 2010. Further, as the GSES is an official statistic and the information is provided by employers reliability of individual earnings information is high (Fitzenberger et al., 2013; Antonczyk et al., 2010). In order to increase the power of the estimation East and West Germany are pooled throughout the analysis. Waves are either analyzed pooled or separately. Pooling data across regions and time requires normalizing the earnings data such that it is measured in deviations from the relevant cap.

### **3.3 Empirical results**

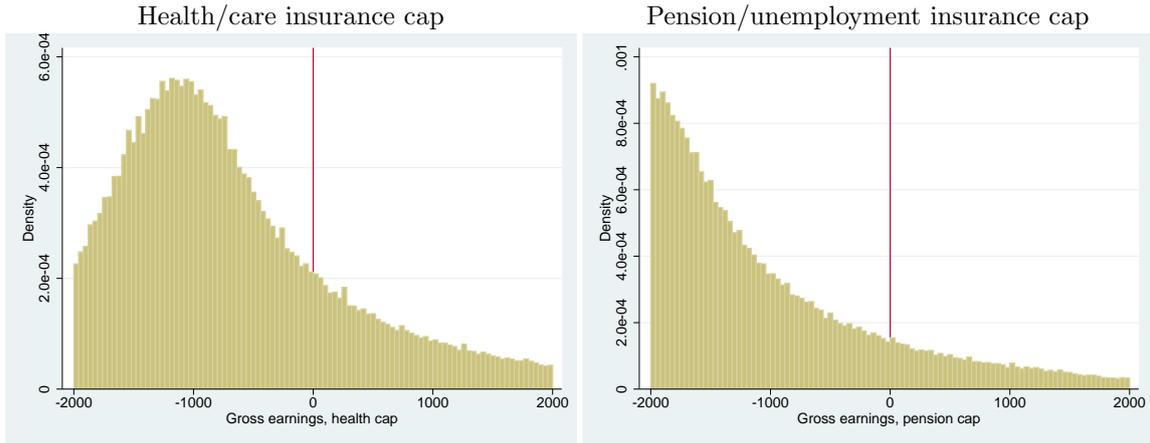
We first present results for the empirical gross earnings distribution including sensitivity analyses supporting their robustness. Results from gross earnings are in principle conclusive in terms of incidence. Yet, the subsequent analysis of net earnings allows for additional plausibility and robustness checks.

#### **Gross earnings**

There is no clear-cut discontinuity in histograms of the monthly gross earnings distribution at the pension or the health cap based on pooled data for all years (Fig. 2). Although being indexed at earnings caps and based on a sizeable data set, the histograms are not completely smooth. There are some minor spikes at various points of the distribution. One of those visible spikes is located directly above the pension insurance cap. To a lesser degree this is also true for the health insurance cap which becomes more visible in histograms for single waves (Fig. C3 in the Appendix).

The majority of those spikes can be explained by round number bunching as documented exemplarily in the non-normalized earnings distribution for West Ger-

**Fig. 2:** Distribution of monthly gross earnings, all years pooled



Source: GSES 1995, 2001, 2006, 2010; own calculations.

many in 2006 (Fig. C4 in the Appendix). In addition standard wages defined by collective agreements or sampling bias might be responsible. In similar fashion it is conceivable that contracted earnings are oriented towards prominent numbers like an earnings threshold. In addition, in some years the earnings threshold was a round number. A spike at an earnings cap result in an obviously biased estimate of the discontinuity (Fig. C5 in the Appendix). We control for these non-substantive factors, i.e. things not related to economic incentives, by including a dummy for the first bin to the right of the cap in equation (9)<sup>8</sup>. Recall that we assume that the underlying ability distribution is smooth translating into a smooth earnings distribution to the left and right of the cap. As a potential discontinuity stretches across the entire distribution of earnings above the cap (see also Fig. 1), a single spike at the cap can not be confused with a discontinuity induced by burden shifting.

Applying the non-/parametric estimators to the pooled sample we find very small negative discontinuities at both earnings caps (Tab. 2, Fig. 3). We get the more negative point estimates with the parametric approach with  $-.016$  for the health and  $-.012$  for the health cap. In the pooled sample the average drop in the marginal

<sup>8</sup>For the non-parametric approach we predict the height of the first bin to the right of the cap using the parametric model with the impact of the dummy being deducted. Kleven and Waseem (2013) controls for round number bunching by adding a joint control variable for all bins containing a round number. This is not appropriate here as, first, the estimation windows usually do not contain more than one or two round numbers and, second, as argued the spike at an earnings cap might be induced by other mechanisms.

SSC rate amounts to 8.1 pp. at the health and about 12.4 pp. at the pension insurance earnings cap for employees and employers, respectively (Tab. 2). The shifting parameters implied by the parametric discontinuity estimates, therefore, are  $s = -0.198$  for the health and  $s = -0.097$  for the pension cap. For the non-parametric approach the corresponding values are virtually zero.

**Tab. 2:** Discontinuity estimates – gross earnings

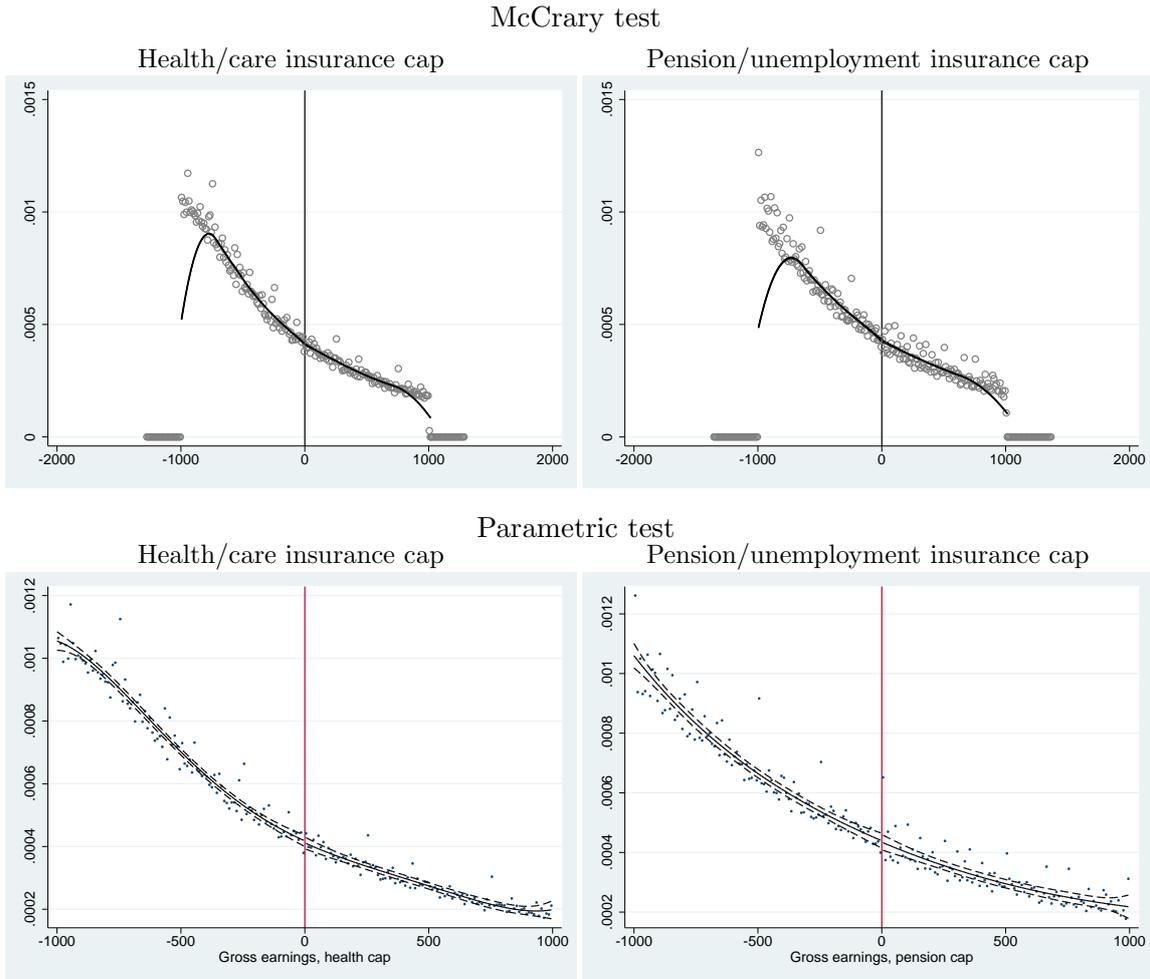
	Pension/unemployment cap					Health/care cap				
	$\Delta t$	McCrary		Polynomials		$\Delta t$	McCrary		Polynomials	
		p.e.	t	p.e.	t		p.e.	t	p.e.	t
Pooled sample	-.124	-.002	-.18	-.012	-1.85	-.081	-.002	-.24	-.016	-2.69
Single waves										
1995	-.126	.051	2.25	.001	.06	-.071	.002	.14	.016	1.29
2001	-.128	-.043	-1.76	-.043	-2.29	-.076	-.045	-2.65	-.051	-3.46
2006	-.130	.024	1.07	.004	.28	-.084	-.021	-1.81	-.035	-3.39
2010	-.114	-.042	-2.36	-.017	-1.81	-.088	.014	.97	.003	-.30

*Notes:*  $\Delta t$  – average drop in marginal SSC rate, p.e. – point estimate, t – t-value ( $\alpha=0.05$ ).  
*Source:* GSES 1995, 2001, 2006, 2010; own calculations.

The fact that point estimates are close to zero implies that there is no evidence that either employers or employees shift (some of) a significant part of their SSC burden. It is nevertheless instructive to look at the confidence intervals for  $s$  to see the maximum amount of shifting supported by these estimates. The left (right) limit of the confidence interval for the parametric discontinuity estimate at the health cap translates to  $s = -0.341$  ( $s = -0.054$ ), i.e. employers shift up to 34% of their contributions to health insurance to their employees. The confidence interval for the non-parametric estimate ranges between  $s = -0.26$  and  $s = 0.09$  and thus includes scenarios where incidence is slightly more on either employers or employees. The confidence interval for the most distinct estimate at the pension cap implies that employers shift at maximum roughly 20% of their contributions to pension insurance. Altogether, the results based on the pooled sample imply that formal and economic incidence of SSC to pension and health insurance (almost) coincide.

This is also supported by the estimates for the separate waves. Although there are some economically and statistically significant deviations from zero, they are usually not verified by the respective alternative method. A notable exemption is the year 2001 where all estimates (across insurance schemes and empirical approaches) are significantly negative (Tab. 2). It would, however, be highly doubtful to interpret

**Fig. 3:** Discontinuity tests, monthly gross earnings, all years pooled



Source: GSES 1995, 2001, 2006, 2010; own calculations.

these estimates as substantive findings. Noise in the sub-samples of single waves is notably higher as illustrated by more spiky histograms (Fig. C3 in the Appendix).

We conduct a number of sensitivity analyses related to choices made in the estimation. First, we vary the amount of bins: the estimation interval of 2000 €/month is split into between 10 and 50 bins on either side of the cap. The baseline estimates are based on 100 bins implying a bin size of 10 €/month. Second, we vary the estimation interval between 1000 and 3000 €/month holding the amount of bins at the level of the baseline specification. The degree of the polynomial is varied between 3 and 9. Finally, we relax the assumption of equal coefficients on either side of the cap in equation (9)<sup>9</sup>. Almost none of the alternative discontinuity estimates are

<sup>9</sup>Note that the specification of the polynomial also impacts non-parametric estimates as the height of the first bin to the right of the cap is predicted based on the parametric model (see above in this section).

qualitatively different from the baseline specification (Tab. 3). The estimates are very robust in terms of estimation decisions confirming the result that statutory equals economic incidence.

**Tab. 3:** Discontinuity estimates – gross earnings, sensitivity

	<b>Pension/unemployment cap</b>				<b>Health/care cap</b>			
	<b>McCrary</b>		<b>Polynomials</b>		<b>McCrary</b>		<b>Polynomials</b>	
	p.e.	t	p.e.	t	p.e.	t	p.e.	t
<i>Baseline</i>	-0.002	-.18	-.012	-1.85	-0.002	-.24	-.016	-2.69
<i>Number of bins</i>								
10	.016	1.54	.002	.14	-.028	-4.36	-.035	-3.93
20	.015	1.32	-.006	.32	-.001	-.12	-.021	-3.15
50	.026	2.19	.002	.2	.004	.56	-.010	-1.49
<i>Estimation interval</i>								
1000	.015	.79	.024	2.45	.004	.43	-.015	-1.88
3000	-.006	-.63	-.005	-.91	-.001	-.12	.054	14.44
<i>Degree of polynomial</i>								
3	-.004	-.35	-.018	-3.15	0.01	1.28	.051	11.31
7	-.005	-.4	-.022	-2.8	-.000	-.03	-.016	-2.39
9	-.001	-.13	.004	.48	.002	.2	-.007	-1
<i>Interactions</i>								
	-.008	-.74	.019	2.79	-.001	-.17	.008	1.35

*Notes:* p.e. – point estimate, t – t-value ( $\alpha=0.05$ ).; robustness tested given the baseline parameter; baseline specification with 100 bins, estimation interval of 2000, polynomial of degree 5 and coefficients being restricted to coincide on either side of the cap.

*Source:* GSES 1995, 2001, 2006, 2010; own calculations.

## Net earnings

In theory analyzing gross earnings densities suffices to determine economic incidence. Similar to the literature on behavioral responses at kinks or notches of tax schedules (Kleven et al., 2011; Kleven and Waseem, 2013), identifying a discontinuity at an earnings cap can be hard in empirical practice, however. Take measurement error as an example. When earnings are biased by a random disturbance term the implicit location of an earnings cap varies across individuals in an unknown way. As a result, a discontinuity estimate might be biased towards zero. Further, distortions like round number bunching, standard wages defined by collective agreements or sampling bias might result in imprecise discontinuity measures. Although we use high quality data with a comparatively large sample size (section 3.2), potential distortions cannot be entirely excluded which is also supported by visible albeit

small noise in the pooled gross earnings distributions (Fig. 2).

We argue here that the additional investigation of the net earnings distribution<sup>10</sup> can enhance the power of our estimation approach. Getting consistent estimates from two earnings distributions provides a broader foundation for conclusions about economic incidence. Saez et al. (2012) show this for a different identification strategy.

The behavior of gross earnings around a cap has implications for net earnings (c):

$$c = \begin{cases} y(1-t) = \frac{n}{1-ts}(1-t) & \forall y \leq \bar{y} \\ y - \bar{y}t = n - \bar{y}ts - \bar{y}t & \forall y > \bar{y} \end{cases} \quad (10)$$

It holds that  $c = n$  if  $s = -1$ , that is if the burden is shifted entirely to employers. By similar arguments as in section 2 for gross earnings this implies that for  $s = -1$  the density of net earnings is smooth, whereas the densities of gross earnings are discontinuous. Should economic and legal incidence coincide, the gross earnings density is continuous while the density functions of net earnings is not.

In the previous section the gross earnings distribution is found to be smooth around the cap(s). This could thus be explained by no shifting or distortions like measurement error. Analyzing net earnings allows for two robustness checks which help to differentiate between the two competing interpretations. First, net earnings can be calculated manually by equation (10). According to theory the difference between the discontinuity estimates of calculated net earnings and gross earnings is exactly equal to the size of the drop in the marginal tax rate. This not being the case suggests that the estimation is distorted and/or lacks the power to precisely identify a discontinuity of the relevant size.

A major advantage of our data is that actually paid employees' SSC are available. Note that using this information is different from calculating net earnings mechanically from observed gross earnings. When measurement error is present the distribution of calculated net earnings would still feature a discontinuity as biased gross earnings serve as the basis for calculating net earnings. By contrast, actually paid employees' SSC are based on true earnings and are, thus, independent from

---

<sup>10</sup>Net earnings are here defined as gross earnings minus employees' SSC. The argument is also true for labor costs defined as gross earnings plus employers' SSC. As we observe actually paid employees' SSC but not actually paid employers' SSC we however only focus on net earnings here.

potential measurement error in gross earnings. When measurement error is present the discontinuities in the distributions of both gross earnings as well as observed net earnings are biased towards zero. Analyzing observed net earnings thus serves as a test whether a smooth gross earnings distribution is conclusive in terms of economic incidence or results from distortions.

Turning to our statistical tests (see Appendix C for the corresponding graphs), we find substantial and significantly negative point estimates for calculated as well as observed net earnings in the pooled sample (Tab. 4). Both the non-parametric and the parametric test yield similar estimates. As expected from the change in the marginal SSC rates discontinuities at the pension cap are larger in magnitude than those at the health cap.

Discontinuity estimates for calculated net earnings coincide with or are slightly more negative than the change in the marginal SSC rate at the respective cap. Given the slightly negative discontinuities for gross earnings the differences between both estimates are consistent with the theoretical prediction. This implies that both approaches are – albeit noise – able to identify the correct size of a discontinuity.

**Tab. 4:** Discontinuity estimates – net earnings

	Pension/unemployment cap					Health/care cap				
	$\Delta t$	McCrary		Polynomials		$\Delta t$	McCrary		Polynomials	
		p.e.	t	p.e.	t		p.e.	t	p.e.	t
Calculated	-.124	-.123	-10.41	-.139	-22	-.081	-.096	-13.17	-.093	-18.88
<i>Placebo</i>										
y+-100	.	.061	4.84	-.008	-.91	.	-.017	-2.58	-.014	-2.39
150	.	.035	2.81	.005	.58	.	.018	2.75	.031	4.97
200	.	.042	3.25	.037	4.19	.	.03	4.93	.053	8.07
Observed	-.124	-.069	-5.58	-.094	-10.08	-.081	-.057	-7.26	-.066	-11.68
<i>Placebo</i>										
y+-100	.	.004	.3	-.035	-3.69	.	.01	1.36	-.008	-1.31
150	.	.006	.47	-.02	-2.57	.	.01	1.34	.017	2.76
200	.	.012	.82	.021	2.52	.	.007	1.01	.032	4.8

*Notes:*  $\Delta t$  – average drop in marginal SSC rate, p.e. – point estimate, t – t-value ( $\alpha=0.05$ ).

*Source:* GSES 1995, 2001, 2006, 2010; own calculations.

For observed net earnings the magnitude of the majority of the discontinuities is somewhat smaller than the drop in the marginal SSC rate (Tab. 4). As argued above this is expected. We do not observe whether an individual is privately health

insured resulting in discontinuity estimates being biased towards zero (section 3.2). The discrepancy between the size of the discontinuity and the drop in marginal SSC rates should thus not be over-emphasized. The fact that we nevertheless find a substantial and significant negative discontinuity for observed net earnings excludes measurement error as an alternative interpretation of a(n) (almost) smooth gross earnings distribution. It rather supports the interpretation of no or small burden shifting.

As an additional robustness check we conduct a sensitivity analysis and look at ‘placebo thresholds’ below the health and above the pension insurance cap for the pooled sample (Tab. 4). The area between both earnings caps is not used for this robustness check to avoid potential interferences among them. Although we find a few sizeable discontinuities, the placebo estimates are considerably less negative than the estimates at the true cap (or even positive). Moreover, found discontinuities at placebo thresholds are not robust across approaches.

## 4 Discussion and conclusion

In this paper we analyze economic incidence of social security contributions (SSC) with cross-sectional estimators. The main advantage of this approach is that it does not rely on reform-induced, exogenous changes in SSC institutions over time. Panel data which often suffer from insufficient sample sizes or imprecise or missing information (e.g. hours of work) are therefore not needed. Moreover, identification does not depend on the validity of a control group and is not solely based on short-term responses.

We thoroughly present the original framework of Alvaredo and Saez (2007) and discuss the identifying assumptions more explicitly. An empirical implementation of the framework is sketched out and it is shown that non-parametric as well as parametric estimators capture the effects of economic incidence on different earnings densities at earnings caps of SSC.

This paper applies the framework to employer-employee data. The German Structure of Earnings Survey (GSES) used here satisfies the necessary requirements in terms of data quality: We pool four waves of data and generate a sample size sufficiently large to reduce random noise in earnings distributions. Being part of

the official labor cost statistics the amount of earnings subject to SSC is measured with little error. Moreover, it contains a direct measure of employees' SSC which allows us to analyze discontinuities at earnings caps for gross as well as net earnings distributions which we argue to improve identification.

The overall picture of the empirical evidence is conclusive. We see no or very small significant discontinuities in the gross earnings distribution which means that legal and economic incidence coincide. This interpretation is confirmed by the identification of statistically significant negative discontinuities in the distribution of observed net earnings. The result is robust for the parametric and non-parametric estimator. Differences in magnitude of the point estimates are qualitatively consistent for varying SSC rates between the pension/unemployment and the health/long-term care insurance cap. Estimates from observed net earnings are slightly downward-biased because of measurement error for privately insured individuals. We thus do not over-emphasize the fact that discontinuities are somewhat below the drop in the marginal SSC rates at the cap leaving room for partial incidence on employees.

Our reading of the overall evidence is that economic and legal incidence are more or less identical for SSC in Germany. Employers and employees thus share the burden of SSC for health and pension insurance. Even in the extreme cases supported by our estimates neither employees nor employers shift a substantial share of their SSC burden to the respective other side of the market. Based on actually paid employees' SSC we argue that measurement issues do not drive this result. Our local estimates at earnings caps would not be informative about underlying incidence when the drop in average SSC rates above the cap was not or not yet internalized into bargained earnings during wage negotiations. Saliency or adjustment frictions however are less convincing in our setting with long-standing earnings caps and our empirical framework which is based on equilibrium incidence.

The result of non-standard economic incidence is in line with recent studies for different countries which are mostly based on quasi-experimental identification (Saez et al., 2012; Skedinger, 2014; Benmarker et al., 2009; Korkeamäki and Uusitalo, 2009; Huttunen et al., 2013). It also corroborates previous evidence for Germany which exploits a reform of the earnings cap of the health insurance in East Germany (Neumann, 2015). There are, however, studies which provide evidence for complete shifting of SSC to employees (Gruber, 1994, 1997). A potential explanation could be

the different institutional settings, in particular divergent wage setting mechanisms (Alesina and Perotti, 1997; Daveri and Tabellini, 2000).

## References

- Alesina, Alberto and Roberto Perotti**, “The Welfare State and Competitiveness,” *The American Economic Review*, December 1997, *87* (5), 921–939.
- Alvaredo, Facundo and Emmanuel Saez**, “The Effects of Payroll Taxes on Earnings: Evidence from Spanish Administrative Data,” Working Paper 2007.
- Antonczyk, Dirk, Bernd Fitzenberger, and Katrin Sommerfeld**, “Rising wage inequality, the decline of collective bargaining, and the gender wage gap,” *Labour Economics*, 2010, *17* (5), 835–847.
- Benmarker, Helge, Erik Mellander, and Björn Öckert**, “Do regional payroll tax reductions boost employment?,” *Labour Economics*, October 2009, *16* (5), 480 – 489.
- Daveri, Francesco and Guido Tabellini**, “Unemployment, growth and taxation in industrial countries,” *Economic Policy*, April 2000, *15* (30), 47–104.
- Fitzenberger, Bernd, Karsten Kohn, and Alexander C. Lembcke**, “Union Density and Varieties of Coverage: The Anatomy of Union Wage Effects in Germany,” *Industrial and Labor Relations Review*, January 2013, *66* (1), 169–197.
- Fullerton, Don and Gilbert E. Metcalf**, “Tax incidence,” in A. J. Auerbach and M. Feldstein, eds., *A. J. Auerbach and M. Feldstein, eds.*, Vol. 4 of *Handbook of Public Economics*, Elsevier, 2002, chapter 26, pp. 1787–1872.
- Grabka, Markus**, “Alternative Finanzierungsmodelle einer sozialen Krankenversicherung in Deutschland – Methodische Grundlagen und exemplarische Durchführung einer Mikrosimulationsstudie.,” Dissertation, Technical University Berlin, 2004.
- Gruber, Jonathan**, “The Incidence of Mandated Maternity Benefits,” *American Economic Review*, 1994, *84* (3), 622–41.
- , “The Incidence of Payroll Taxation: Evidence from Chile,” *Journal of Labor Economics*, July 1997, *15* (3), 72–101.

- Huttunen, Kristiina, Jukka Pirttilä, and Roope Uusitalo**, “The employment effects of low-wage subsidies,” *Journal of Public Economics*, 2013, *97*, 49 – 60.
- Kleven, Henrik J. and Mazhar Waseem**, “Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan,” *The Quarterly Journal of Economics*, 2013, *128* (2), 669–723.
- Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez**, “Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark,” *Econometrica*, 2011, *79* (3), 651–692.
- Korkeamäki, Ossi and Roope Uusitalo**, “Employment and wage effects of a payroll-tax cut—evidence from a regional experiment,” *International Tax and Public Finance*, December 2009, *16* (6), 753–772.
- Liebman, Jeffrey and Emmanuel Saez**, “Earnings Responses to Increases in Payroll Taxes,” Technical Report, Harvard University, University of California, Berkeley, and NBER, 2006.
- McCrary, Justin**, “Manipulation of the running variable in the regression discontinuity design: A density test,” *Journal of Econometrics*, 2008, *142* (2), 698 – 714.
- Melguizo, Angel and Jose Gonzalez-Paramo**, “Who bears labour taxes and social contributions? A meta-analysis approach,” *SERIEs*, 2013, *4* (3), 247–271.
- Neumann, Michael**, “Earnings Responses to Social Security Contributions: Evidence from two German Administrative Data Sets,” Technical Report, DIW Discussion Paper No. 1489, German Institute for Economic Research (DIW Berlin) 2015.
- OECD**, “OECD revenue statistics,” 2016.  
<https://stats.oecd.org/Index.aspx?DataSetCode=REV>; accessed 2015.
- Saez, Emmanuel**, “Do Taxpayers Bunch at Kink Points?,” *American Economic Journal: Economic Policy*, 2010, *2* (3), 180–212.

– , **Manos Matsaganis, and Panos Tsakloglou**, “Earnings Determination and Taxes: Evidence From a Cohort-Based Payroll Tax Reform in Greece,” *The Quarterly Journal of Economics*, 2012, 127 (1), 493–533.

**Skedinger, Per**, “Effects of Payroll Tax Cuts for Young Workers,” *Nordic Economic Policy Review*, 2014, 5 (1), 125–169.

## A Model extensions

### A.1 Varying SSC rates

We assume identical statutory SSC rates for employers ( $r$ ) and employees ( $e$ ) in our model. Relaxing this assumption yields

$$y = \frac{n}{1 + t_r \max(0, s) + t_e \min(s, 0)}. \quad (11)$$

As in our baseline model,  $s > 0$  implies that employers are able to shift part of their burden to their employees. In that case, the distortion of  $n$  depends on the SSC rate of the employer,  $t_r$ . If  $s < 0$ , employees are able to shift part of their burden to their employers and the distortion of  $n$  depends on the SSC rate of the employee,  $t_e$ .

Considering the left and right limit of the density of realized gross earnings when  $y$  approaches  $\bar{y}$ :

$$\begin{aligned} f(\bar{y})_- &= p\{\bar{y}[1 + t_r \max(0, s) + t_e \min(s, 0)]\} [1 + t_r \max(0, s) + t_e \min(s, 0)] \\ &= p\{\bar{y} + \bar{y}t_r \max(0, s) + \bar{y}t_e \min(s, 0)\} [1 + t_r \max(0, s) + t_e \min(s, 0)] \\ &= p(\bar{n}) [1 + t_r \max(0, s) + t_e \min(s, 0)] \end{aligned} \quad (12)$$

$$\begin{aligned} f(\bar{y})_+ &= p\{\bar{y} + \bar{y}t_r \max(0, s) + \bar{y}t_e \min(s, 0)\} \\ &= p(\bar{n}) \end{aligned} \quad (13)$$

The interpretation of the size of the discontinuity depends on its sign. If the distribution exhibits a drop (jump), full shifting is implied by a change of  $t_r\%$  ( $t_e\%$ ).

### A.2 Varying SSC rates with asymmetric incidence

When, in addition, employers' and employees' contributions are allowed to be shared independently, gross earnings in terms of ability become

$$y = \frac{n}{1 + t_r s_r - t_e s_e} \quad (14)$$

As before, the potential discontinuity in the earnings distribution at the threshold can be characterized as follows:

$$f(\bar{y})_- = p(\bar{n}) [1 + t_r s_r - t_e s_e] \quad (15)$$

$$f(\bar{y})_+ = p(\bar{n}) \quad (16)$$

The discontinuity, thus, identifies an average of the SSC rates of employers and employees weighted by the corresponding shifting parameters. Separate identification of  $s_r$  and  $s_e$  would require differential changes of employer and employee SSC rates over time.

### A.3 Heterogeneous shifting

The analysis in the main text is based on the assumption of  $s$  to be homogeneous. Allowing for heterogeneity in  $s$  raises the question whether a potential discontinuity in a given distribution at an earnings cap is informative about average economic incidence in the population. We thus use the simulation from the estimation section 2.3 (for details on the simulation see Appendix B) to check how representative this local estimate is. Homogeneity from the main simulation is relaxed and  $s$  is allowed to vary according to alternative distributional assumptions.

The first and second moments are held constant throughout this exercise: the shifting parameter  $s$  has a mean of 0.5 and a standard deviation of .239 for each distribution. This means that employers are on average able to shift half of their SSC burden to employees. For an assumed drop in the SSC rate of 7.6% at the cap, this yields a ‘true’ drop in the gross earnings density of .038 for the average individual. We start with a *uniform* distribution where  $s$  varies between 0 and 1. Alternatively we estimate discontinuities for underlying *normal* and *log-normal* distributions of  $s$ . Finally we allow for a correlation between the the shifting parameter and earnings. We use a *bivariate normal* distribution (with first and second moments for  $y$  and  $s$  as above) and a correlation coefficient of .5. This means that employers’ ability to shift their SSC burden to employees is increasing in earnings.

For homogenous  $s$  we find a discontinuity of  $-0.037$  with the non-parametric and  $-0.038$  with the parametric approach (Tab. A1). Both estimates match the ‘true’ discontinuity of  $-0.038$ . Allowing for heterogeneity in  $s$  we see that both, the parametric and the non-parametric test employed in this paper yield discontinuity estimates of between  $-0.036$  and  $-0.037$ . The estimates also closely resemble the true mean incidence in these scenarios regardless of the distribution of  $s$  with confidence intervals increasing slightly. As long as the distribution of shifting parameters is not related to earnings, the locally identified estimate is representative for mean incidence in the sample.

**Tab. A1:** Discontinuity estimates – simulation of heterogeneous shifting parameter

	McCrary		Polynomials	
	p.e.	CI	p.e.	CI
<i>Homogeneous s</i>	-0.037	[-.057, -.019]	-0.038	[-.054, -.024]
<i>Heterogeneous s</i>				
Uniform	-0.036	[-.063, -.010]	-0.037	[-.052, -.021]
Normal	-0.036	[-.058, -.014]	-0.037	[-.052, -.022]
Log-normal	-0.037	[-.060, -.013]	-0.037	[-.052, -.023]
<i>Heterogeneous s, correlation with y</i>				
Bivariate normal	-0.046	[-.065, -.027]	-0.044	[-.058, -.030]

*Notes:* p.e. – point estimate, CI – 95% confidence intervall.

A mean of the shifting parameter of  $s = 0.5$  and a drop in the marginal SSC rate of 7.6% yield a ‘true’ drop in density of .038 for the average in individual in the population.

*Source:* Own simulation.

The picture changes in a situation where the employers’ ability of shifting their SSC burden is correlated with earnings. For a positive correlation coefficient of .5 we find a discontinuity estimate of  $-.046$  (Tab. A1). The upward bias in the estimate of mean incidence can be explained by selection: The cap is located in the upper part of the earnings distribution. Given the assumed positive combination with shifting this results in an above-average  $s$  for individuals around the cap. The local estimator is no longer able to capture mean incidence in the sample under these circumstances.

## A.4 Behavioral responses

Alternative margins of adjustment, e.g. quantity adjustments on the labor market, are not modelled in our framework (sub-section 2.1). A change of the marginal SSC rate at an earnings cap, however, also entails labor supply and demand incentives. This is particularly relevant as underlying labor supply and demand elasticities might drive both behavioral responses as well as economic incidence. Alvaredo and Saez (2007) provide a model extension with variable labor supply. Assume individuals with skill (or preference for work)  $n$  to maximize quasi-linear utility  $u^n(c, z)$  which is a function of net earnings  $c$  and labor costs  $z$  (with  $c = z - T(z) = \frac{1-t}{1+t}z$  where  $T(z)$  is total tax liability). It holds that  $\partial u/\partial c > 0$  and  $\partial u/\partial z < 0$  (costs of labor supply). Quasi-linear preferences rule out income effects. This is not restrictive in our framework, though. As the discontinuity at the cap is driven by

employees located close to the cap, the change in average tax rate is limited and so is the incentive for income effects.

Under neoclassical labor supply a positive labor supply elasticity makes it sub-optimal to locate close to the earnings cap (Saez, 2010). That is, if the labor supply elasticity is homogeneous and strictly positive, the distribution of earnings will feature a gap around the cap. Although a potential discontinuity generated by economic incidence would theoretically be present at the borders of the gap, it would be impossible to identify it empirically as the probability mass directly around the cap is zero.

If the labor supply elasticity is heterogeneous and zero for at least some individuals, the earnings distribution would feature a dip instead of a gap (Saez, 2010; Neumann, 2015). Depending on the size of the elasticities, it is theoretically still possible to identify a discontinuity in the density of gross earnings. The discontinuity is in principle not affected by the labor supply behavior. As discussed above a potential discontinuity is informative about economic incidence for individuals who locate close to the earnings cap. Those people have by definition small labor supply elasticities, otherwise they would not sit there. When labor supply behavior and economic incidence are correlated, the incidence estimate is highly selective.

In the case of a dip it is a practical empirical question whether an estimator picks up the curvature of such a dip and is still able to identify a discontinuity within this area. We conducted a number of Monte-Carlo simulations varying the size of the dip due to behavioral responses in order to study the sensitivity of the estimators in that regard (for details on the simulation see Appendix B). The non-parametric estimator performs in general better than the parametric estimator. It is able to pick up the curvature generated by behavioral responses and in most cases still identifies a discontinuity at the earnings cap (Tab. A2). By contrast, estimates based on the parametric approach are completely out of place. The parametric estimator is not able to identify discontinuities in the presence of sizeable dips in earnings densities.

We start with a scenario where roughly two thirds of all employees have a positive labor supply elasticity  $e$  which is simulated to be uniformly distributed in the interval  $(0, .7)$ . Remaining employees are assigned an elasticity of zero. Economic incidence is assumed to be homogeneous, fully with employees ( $s = 1$ ), and independent from  $e$ . The distribution thus features a dip around the earnings cap and, in addition,

**Tab. A2:** Discontinuity estimates – simulation of behavioral responses

Simulation parameters	Actual dis-continuity	McCrary		Polynomials	
		p.e.	CI	p.e.	CI
$P(e = 0) = 1; s = 1$	-.075	-.072	[ -.094;-.049 ]	-.076	[ -.092;-.06 ]
$P(e \sim U(0, .7)) = .67;$ $P(e = 0) = .33; s = 1$	-.075	-.051	[ -.100;0 ]	.141	[ .117;.165 ]
$P(e \sim U(0, .7)) = .67;$ $P(e = 0) = .33; s = 1 - e$	-.054	-.048	[ -.100;.003 ]	.089	[ .065;.113 ]
$P(e \sim U(0, .7)) = .9;$ $P(e = 0) = .1; s = 1 - e$	-.050	.033	[ -.116; .196 ]	.251	[ .219;.284 ]

*Notes:* p.e. – point estimate, CI – 95% confidence interval.

*Source:* Own simulations.

exhibits a discontinuity at  $\bar{y}$ . The non-parametric estimate of  $-.051$  is smaller than the true value of  $-.075$  but remains significant, although the behavioral distortion in general reduces precision. (Tab. A2).

According to the neo-classical model incidence is determined by the labor supply elasticity ( $s = 1 - e$ ): Highly elastic employees bear less of the SSC burden than individuals with lower elasticities. We model this relationship between  $e$  and  $s$  in the last two simulations. First, the elasticity is assumed to be uniformly distributed between zero and  $.7$  for two thirds of employees while all others are inelastic as above. The incidence parameter is thus  $s = .77$  on average resulting in a true discontinuity of  $-.054$ . The McCrary test almost exactly delivers that value (Tab. A2). The better performance in comparison to the first simulation is a combination of a selective local estimate with the downward bias under behavioral adjustments.

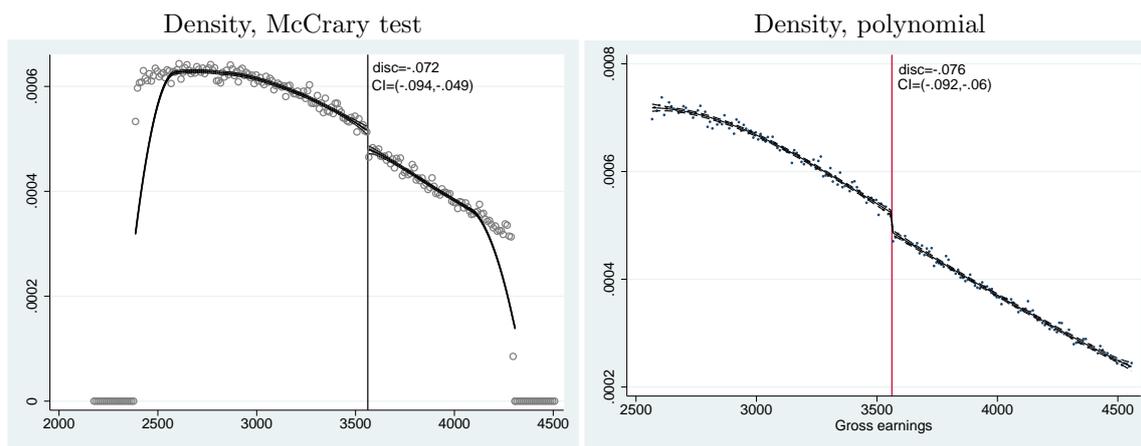
Second, we increase the share of people with labor supply elasticities in the interval  $(0, .7)$  to  $.9$  which means that the the overall fraction with  $e = 0$  is reduced to  $.1$ . The dip around the earnings cap is increased and the related average discontinuity decreases to  $-0.050$ . The non-parametric point estimate of  $.33$  no longer represents the correct direction of incidence (Tab. A2). It seems that the combination of a large dip, a small amount of probability mass at  $\bar{y}$ , and a relatively small actual discontinuity also disables identification based on the non-parametric approach. Taken altogether behavioral responses introduce bias to incidence estimates. When incidence is correlated with elasticities, the local estimator is selective. Under large behavioral responses identification of a discontinuity is lost even with a non-parametric estimator.

Empirical studies have not been able to identify a gap or a dip around earnings caps (Liebman and Saez, 2006; Neumann, 2015). We do not find evidence for behavioral responses either (section 3.3). People with high earnings may indeed have small labor supply elasticities. Alternatively adjustment costs could be too high to locate optimally given the relatively moderate increase in utility (Neumann, 2015).

## B Monte-Carlo evidence

This section provides some Monte-Carlo evidence on the performance of the parametric and non-parametric estimator of a discontinuity in an earnings distribution. Gross earnings distributions  $f(y)$  are simulated on the basis of equation (3) under the assumption of full shifting of employers' SSC to employees ( $s = 1$ ). SSC rates for employees and employers are identical. We assume the underlying ability  $n$  to be normally distributed with  $p(n) \sim N(2,913; 4,481)$  and take 1000 random draws of 2,000,000 observations from this distribution. We assume a drop in the marginal SSC rate of roughly 7.6% at 3,562 €. The choice of distributional and tax parameters draws on our empirical application to the health insurance cap in 2006. The resulting density features a clear negative discontinuity (Fig. 1 in the main text). The McCrary test and the parametric test identify a discontinuity of -7.2% and 7.6%, respectively, which matches the underlying parameters  $s$  and  $t$  (Fig. B1). The confidence interval of the McCrary estimate translates to an interval of  $(-1.24, -0.64)$  for the shifting parameter  $s$ . The corresponding interval for  $s$  based on the parametric test is slightly narrower:  $(-1.21, -0.79)$ .

**Fig. B1:** Discontinuity tests, simulation

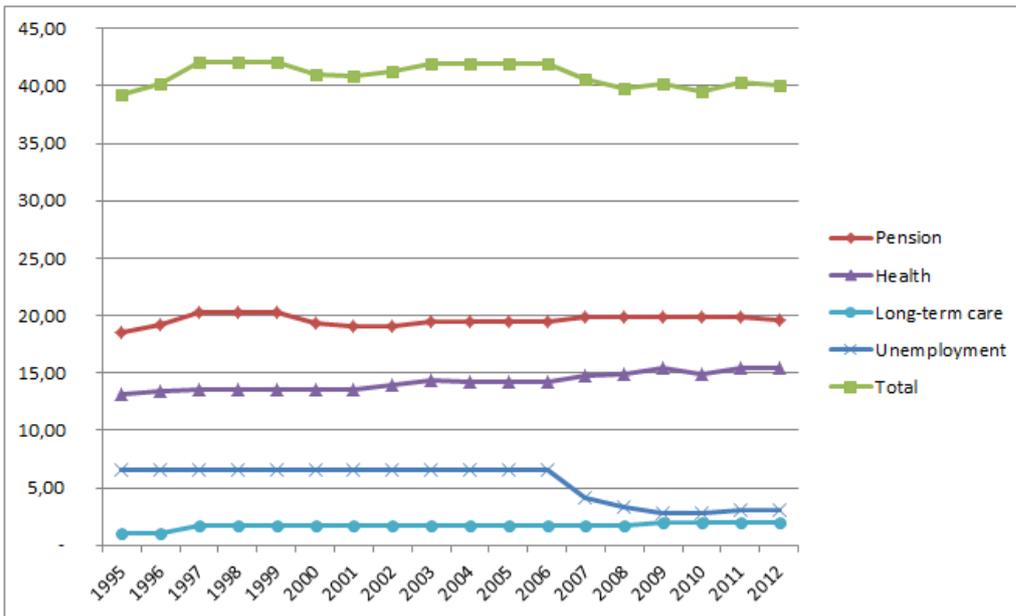


*Notes:* disc – p.e. of discontinuity, CI – confidence interval ( $\alpha=0.05$ ).

*Source:* Own simulation.

## C Additional figures

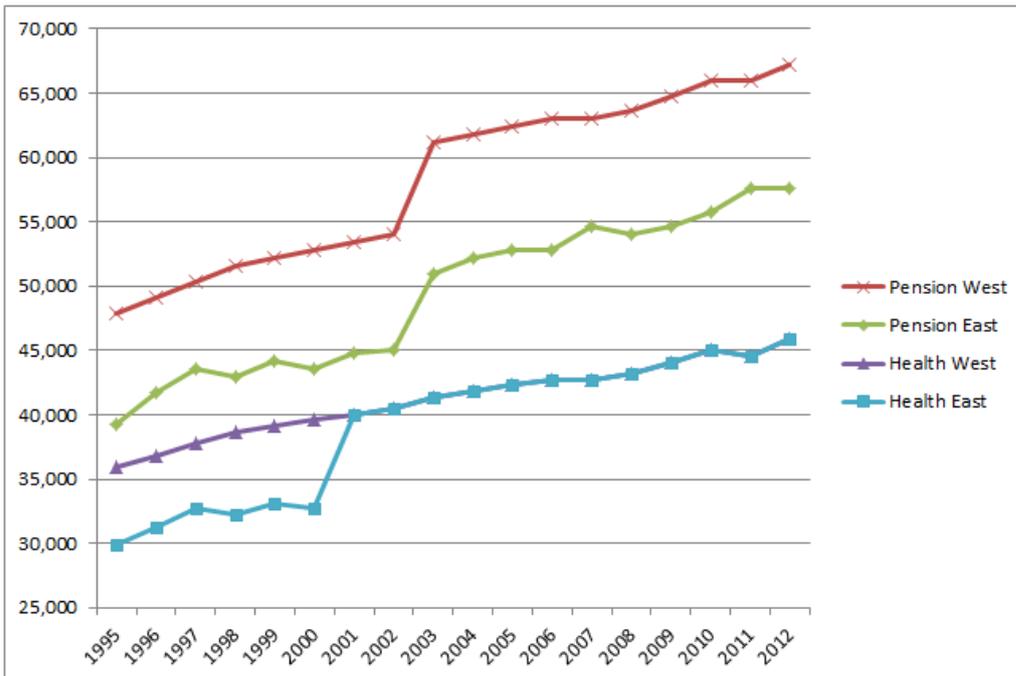
**Fig. C1:** Development of SSC rates over time



*Notes:* The additional fee for childless employees, introduced in 2005, is omitted; The change of SSCs which came into effect in July 2005, are considered as of 2006. Until 2006, SSC rates for health insurance varied between health insurance companies and the given numbers are averages. In 2001 for example, it varied between 11.0% and 14.9% Grabka (2004).

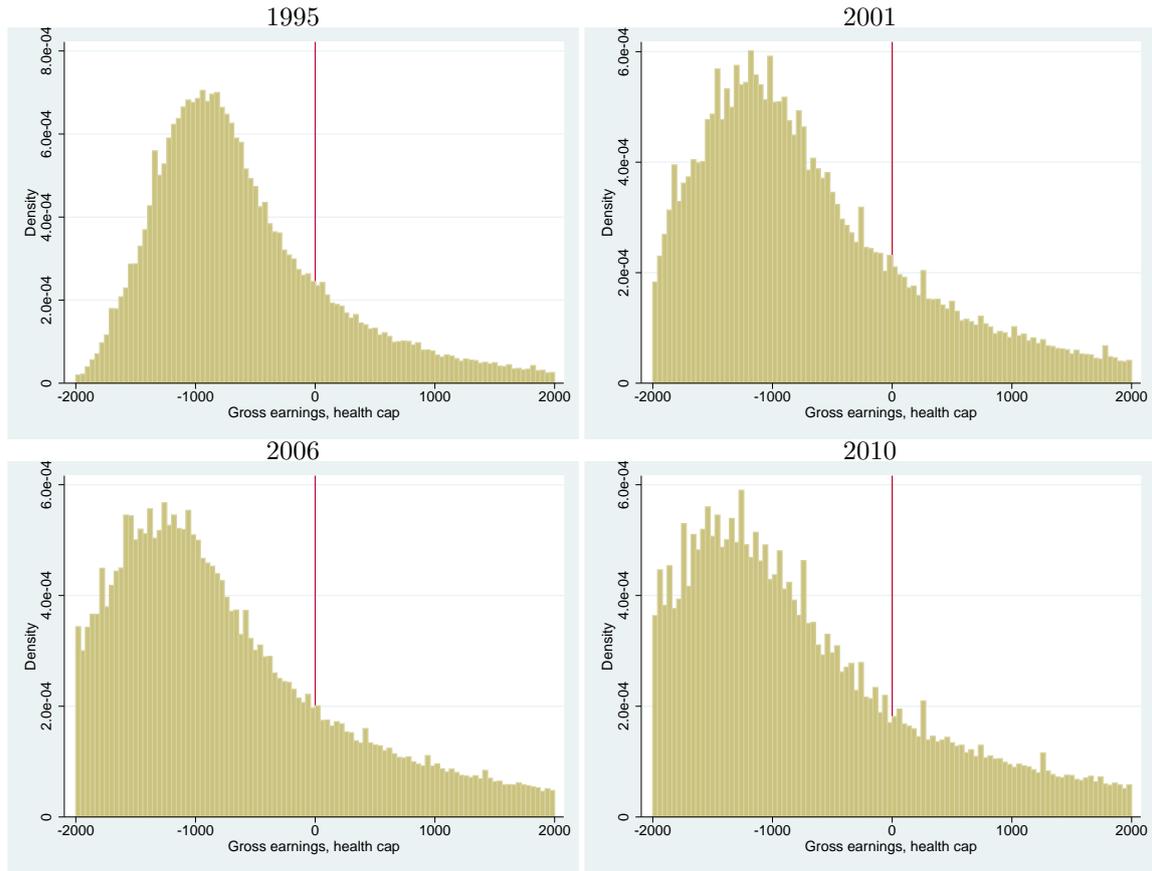
*Source:* German Statistical Office.

**Fig. C2:** Development of earnings caps over time

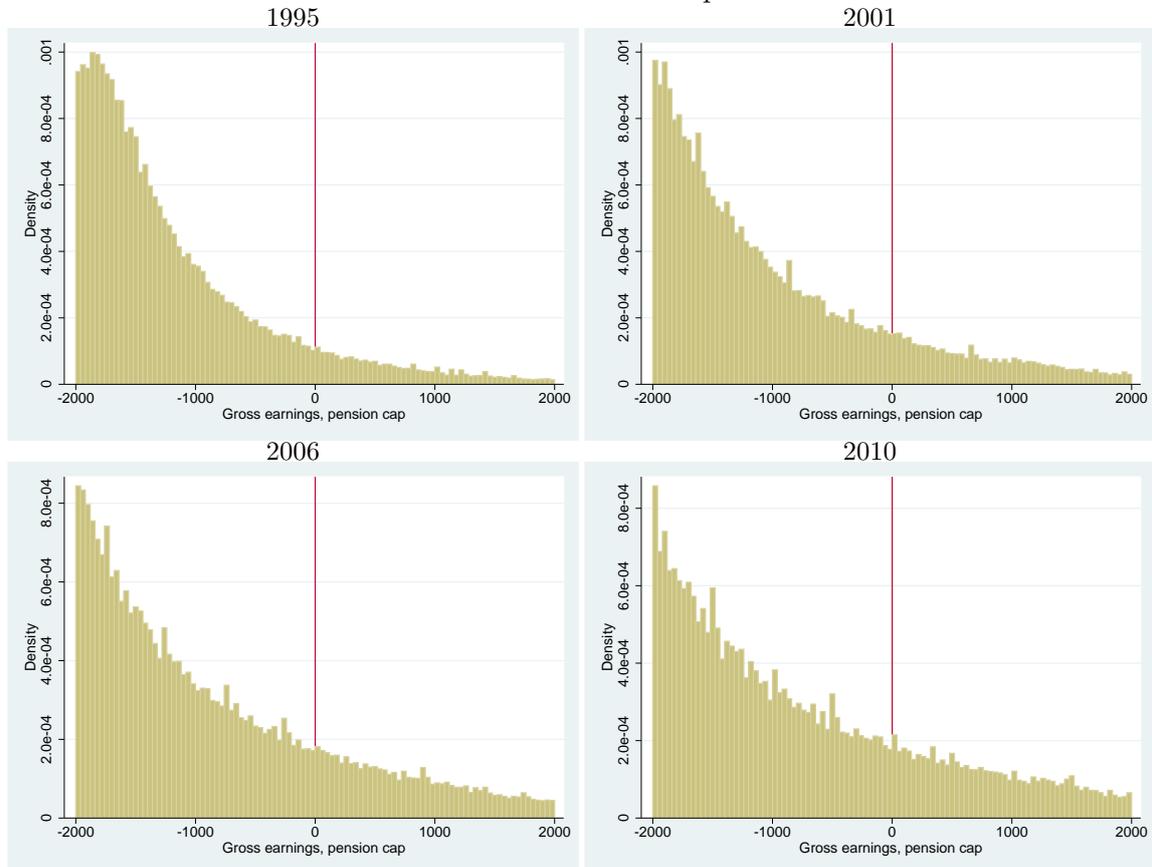


*Source:* German Statistical Office.

**Fig. C3:** Distribution of gross earnings, single years  
Health insurance cap

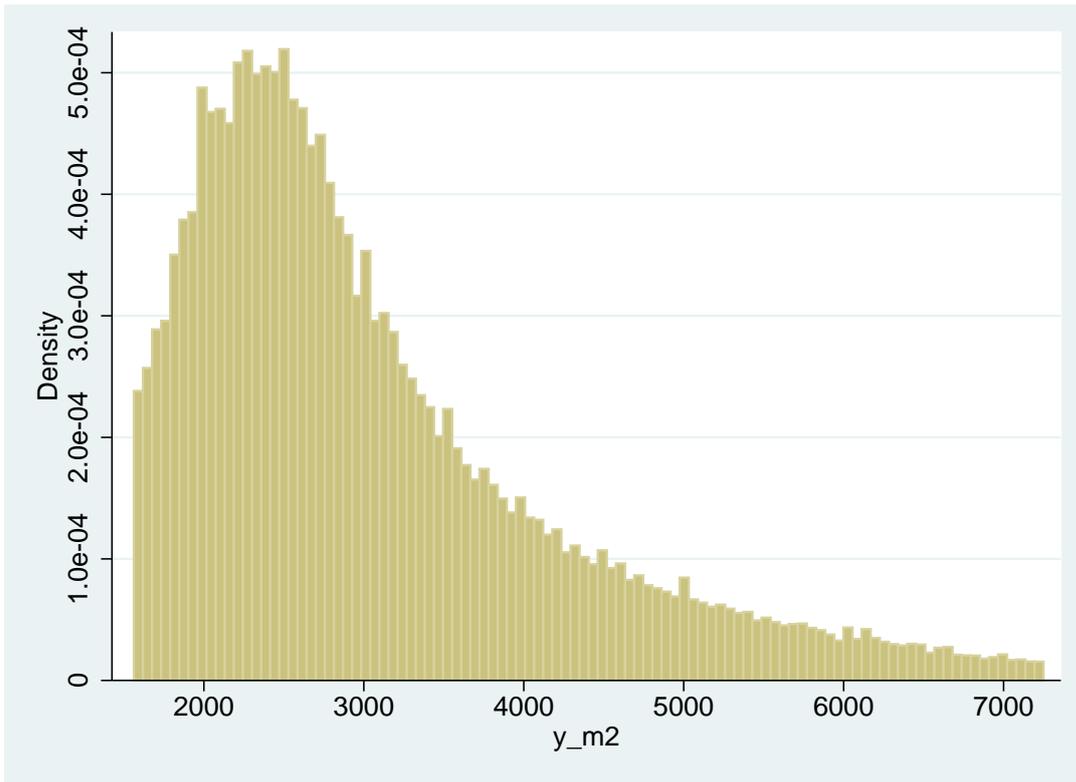


Pension insurance cap



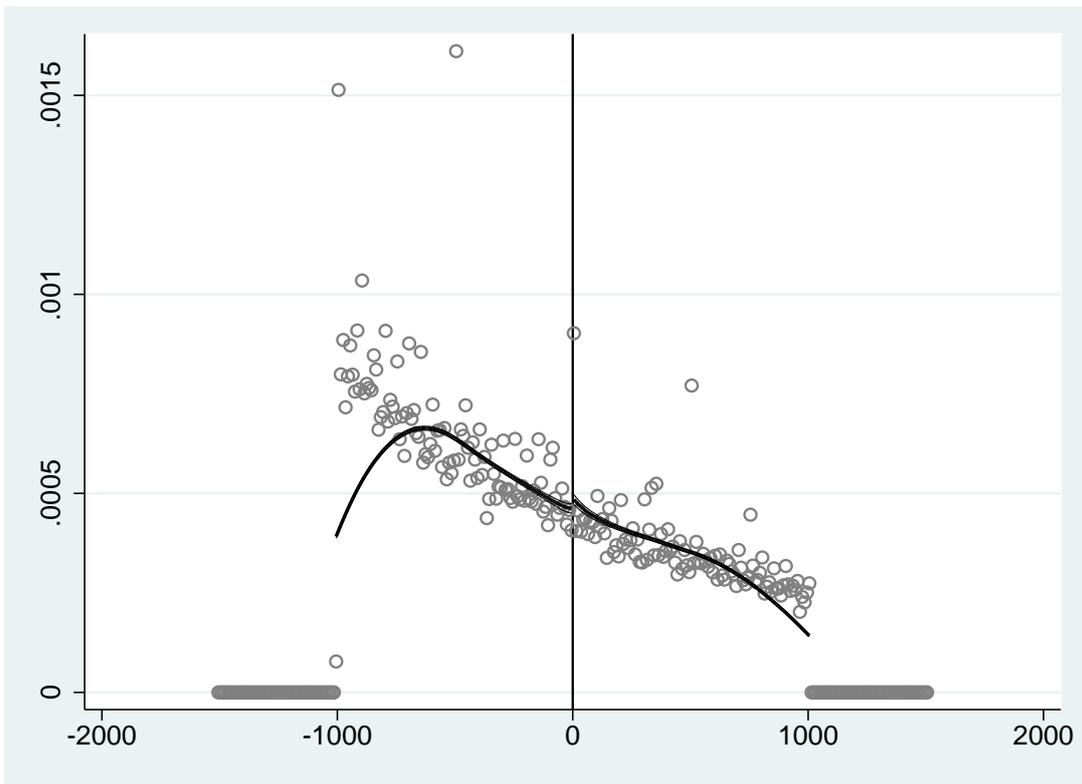
Source: GSES 1995, 2001, 2006, 2010; own calculations.

**Fig. C4:** Distribution of non-normalized gross earnings, West Germany, 2006



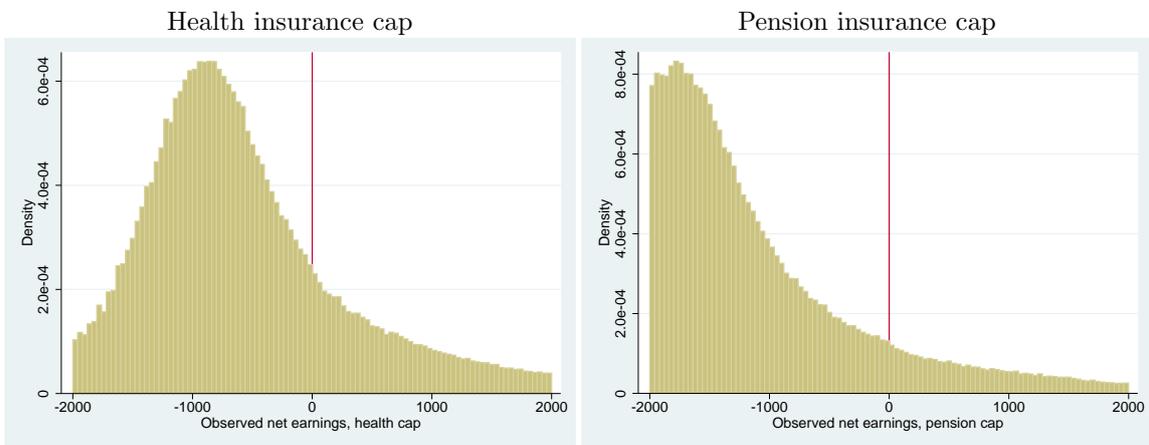
Source: GSES 2006.

**Fig. C5:** Non-parametric discontinuity test, gross earnings, 2010, no correction



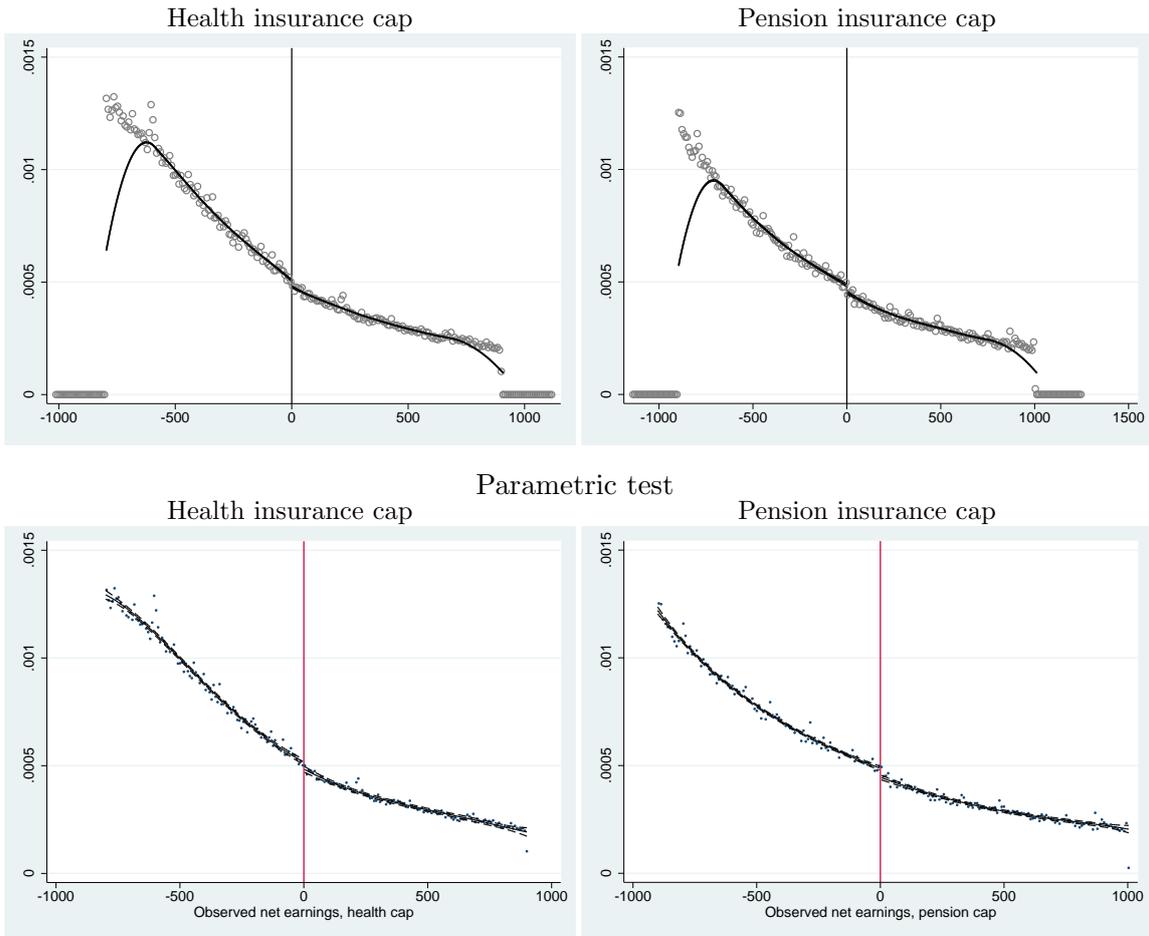
Source: GSES 2010.

**Fig. C6:** Distribution of observed monthly net earnings, all years pooled



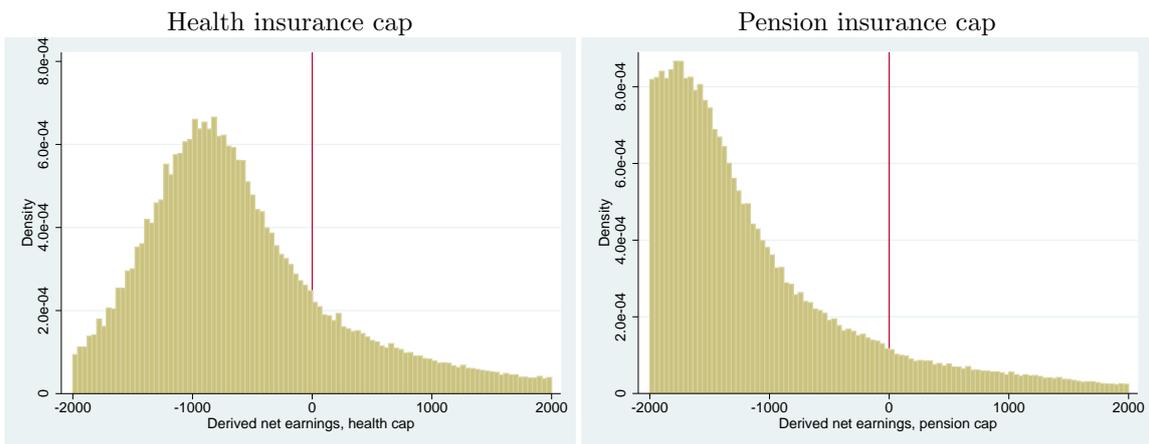
Source: GSES 1995, 2001, 2006, 2010; own calculations.

**Fig. C7:** Discontinuity tests, observed monthly net earnings, all years pooled  
McCrary test



Source: GSES 1995, 2001, 2006, 2010; own calculations.

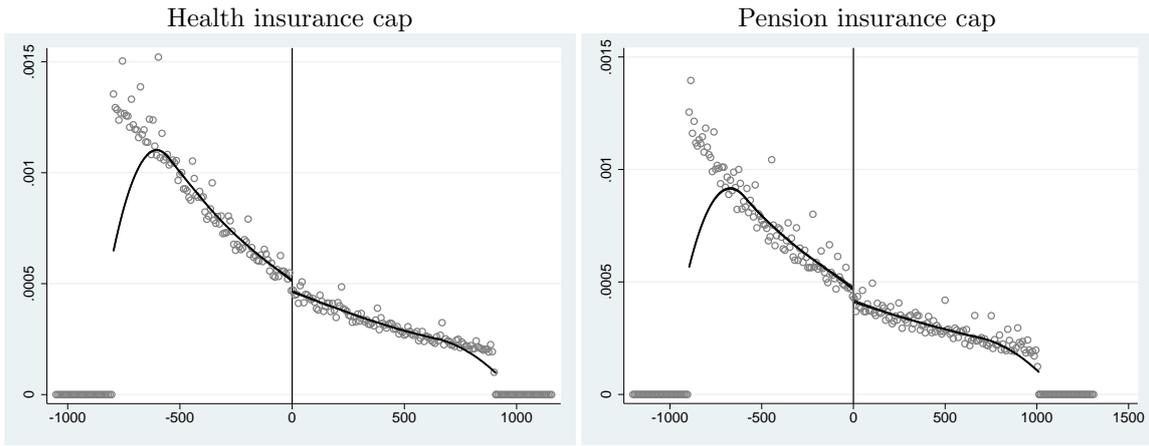
**Fig. C8:** Distribution of calculated monthly net earnings, all years pooled



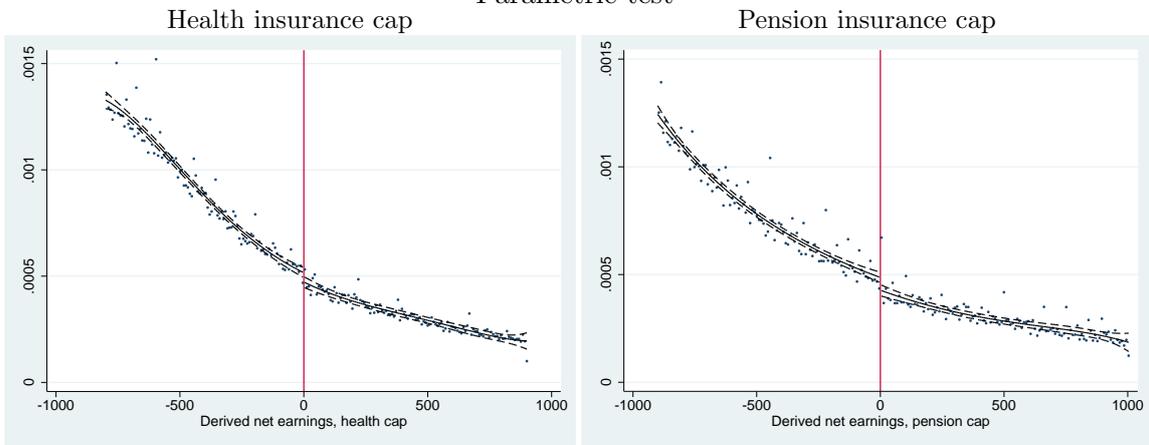
Source: GSES 1995, 2001, 2006, 2010; own calculations.

**Fig. C9:** Discontinuity tests, calculated monthly net earnings, all years pooled

McCrary test



Parametric test



Source: GSES 1995, 2001, 2006, 2010; own calculations.