

Müller, Kai-Uwe; Neumann, Michael

Conference Paper

How reliable are incidence estimates based on cross-sectional distributions? Evidence from simulations and linked employer-employee data

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Empirical Public Economics, No. A14-V2

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Müller, Kai-Uwe; Neumann, Michael (2015) : How reliable are incidence estimates based on cross-sectional distributions? Evidence from simulations and linked employer-employee data, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Empirical Public Economics, No. A14-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

This Version is available at:

<https://hdl.handle.net/10419/112920>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

How reliable are incidence estimates
based on cross-sectional distributions?
Evidence from simulations and linked employer-employee data

Kai-Uwe Müller*, Michael Neumann

March 1, 2015

Preliminary and incomplete –
please do not cite without permission of the authors!

Abstract

The paper validates an empirical approach developed by Alvaredo and Saez (2007) which estimates the economic incidence of social security contributions (SSC) on the basis of cross-sectional earnings distributions. The method utilizes discontinuities at earnings caps where the marginal SSC rate drops. It does not rely on policy reforms, panel data, or hours information. We demonstrate on the basis of simulations that this comes at the cost of ruling out optimization frictions and measurement error destroying identification. Behavioral responses and incidence heterogeneity are secondary problems. Exploiting German linked employer-employee data that provide separate measures of gross and net earnings we find substantial negative discontinuities in net earnings. Together with small, erratic discontinuities in gross earnings this provides consistent empirical evidence that legal and economic incidence of SSC coincide in Germany.

Keywords: incidence; social security contributions; Monte Carlo simulation; linked employer-employee data

JEL classification: H22; J38; H55; J20

* 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

1 Background and questions

The majority of the empirical literature on the economic incidence of payroll taxes or social security contributions (SSC) exploits longitudinal variation (see Melguizo and Gonzalez-Paramo (2013) for a meta-analysis). Observational regressions usually lack clean identification, though. More credible approaches exploit exogenous variation generated by policy reforms (Gruber, 1994, 1997; Benmarker et al., 2009). Suitable SSC reforms are infrequent, however. Only one policy change occurred, for instance, in Germany which generated sufficient institutional variation and provided a setting with a valid control group (Neumann, 2014). Besides relying on the common-trends assumption, these difference-in-difference type estimates only depict short-term responses. Data restrictions exacerbate these limitations. Should hours of work not be observed – which is the case in German administrative data – 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.

Alvaredo and Saez (2007) suggest an alternative framework to circumvent some of the aforementioned methodological issues. Their approach utilizes earnings cap(s) defined within a given SSC schedule. Economic incidence can then be analyzed on the basis of cross-sectional earnings distributions. The discontinuous drop of the marginal SSC rate at an earnings cap reduces the tax wedge for earnings beyond the cap. In a system with shared contributions, employees and employers benefit according to their formal share in contributions as long as wage rates remain unaltered. When economic differs from statutory incidence, however, a drop in the marginal tax rate above the cap will lead to a change of the hourly wage related to earnings beyond the cap. Alvaredo and Saez show that this change results in a discontinuity in the gross earnings distribution at the cap which can be exploited to estimate economic incidence: A decrease (increase) of the wage rate compresses (expands) the density above the earnings cap and results in a positive (negative) discontinuity. Under the assumption that the beneficiary of the reduced SSC liability has borne the burden in the first place, this implies that economic incidence is more on the side of the employer (employee). When statutory equals economic incidence, the wage does not change and the gross earnings distribution is smooth at the cap.

An obvious advantage of this approach is that it neither depends on an exogenous policy reform. Nor does it require panel data or information on hours of work. Estimates derived from this framework might represent long-term incidence effects when the earnings cap has not changed for a certain period of time. A first contribution of this paper is to refine the approach theoretically and show that these advantages come at the cost of restrictive assumptions in other respects. We show that a potential discontinuity at an earnings cap is generated by wage rate responses

of employees locating slightly above the threshold. Drawing conclusions about economic incidence presupposes that the reduction of the SSC liability at the earnings cap is immediately shared after crossing the threshold. In empirical practice the approach is prone to be distorted by measurement and sampling error as well as imperfect observation of the parameters of the social security system. We assess the sensitivity of the method by way of Monte Carlo-like simulations.

We also argue that looking at measures of net earnings and labour costs in addition to an analysis of the gross earnings distribution increases the robustness of the approach in the presence of such distortions. However, most available data sets only include gross earnings. Net earnings and labour costs have to be deducted by a deterministic rule which perpetuates most of the potential distortions and renders robustness tests useless. The second contribution of this paper is to apply the method to a linked employer-employee data set (the German Structure of Earnings Survey, GSES). This data set includes directly observed and precise measures of gross as well as net earnings. The robustness of the incidence estimates can for the first time be tested on these two different distributions.

Our simulation exercises confirm theoretical reservations related to measurement error and optimization frictions. Both potentially destroy the identification of incidence parameters at the earnings cap. On the contrary, behavioral responses of labor supply and demand as well as heterogeneity in shifting of the SSC burden are a lesser problem. We find substantial and significant discontinuities in ‘observed’ net earnings. Together with at large small and often insignificant discontinuities in gross earnings this provides consistent empirical evidence that legal and economic incidence of SSC coincide in Germany.

The remainder of this paper is organized as follows. Section 2 presents and critically discusses the methodology. Section 3 conducts Monte Carlo-like simulations to assess the impact of data limitations as well as the violation of assumptions. Section 4 applies the approach to a German data set. Section 5 concludes.

2 Methodology

Assume ability n to follow the smooth cumulative (marginal) distribution function $P(n)$ ($p(n)$). Let n be defined such that realized gross earnings equal ability $y = n$, if either the SSC rate t is zero or legal and economic incidence coincide. The latter implicitly assumes that there are no behavioural responses to SSC in terms of labor supply or demand; this will be relaxed below. Should legal and economic incidence differ, y is distorted such that $y = \frac{n}{1+t*s}$ with s being a homogeneous shifting parameter which is positive (negative) if employers (employees) shift some burden to employees (employers). We show by simulation in section 3 that relaxing the assumption and allowing for heterogeneity in shifting does not distort identification within our framework.¹

To simplify the exposition of the approach, we assume identical statutory SSC rates for both sides of the labor market. Relaxing this assumption would yield $y = \frac{n}{1+t_r*s - t_e*(1-s)}$ with r (e) denoting employers (employees). When employers' and employees' contributions are allowed to be shared independently gross earnings become $y = \frac{n}{1+t_r*s_r - t_e*s_e}$. The simplification is not restrictive in the German case, however, for two reasons: First, statutory SSC rates are indeed equal in the analyzed years. Second, the earnings caps apply for employers as well as employees which is why s_r and s_e could not be identified separately anyway.

Introducing 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 ($F(y)$):

$$F(y) = P(n) = \begin{cases} P(y(1+t*s)) & \forall y \leq \bar{y} \\ P(y + \bar{y}t*s) & \forall y > \bar{y} \end{cases}$$

$$f(y) = \begin{cases} p(y(1+t*s))(1+t*s) & \forall y \leq \bar{y} \\ p(y + \bar{y}t*s) & \forall y > \bar{y} \end{cases}$$

Equations (1) and (2) state the left and right limit of the density of realized gross earnings when $y \rightarrow \bar{y}$:

$$f(\bar{y})_- = p(\bar{y}(1+t*s))(1+t*s) = p(\bar{y} + \bar{y}t*s)(1+t*s) = p(\bar{n})_- (1+t*s) \quad (1)$$

$$f(\bar{y})_+ = p(\bar{y} + \bar{y}t*s) = p(\bar{n})_+ \quad (2)$$

As $p(n)$ is assumed to be smooth, there will be a discontinuity in the density of gross earnings $f(y)$ as long as $s \neq 0$, which means that some part of the burden is shifted and 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 to employees (employers). Full shifting would be implied by $s = +(-)1$ and result in a discontinuity of $+(-)t\%$ relative to the density directly below the cap.

¹This will be shown analytically in the next version of the paper.

To put it more intuitively: For employees' earnings exceeding the earnings cap, an additional Euro of gross earnings translates into relatively more net earnings and less labor costs. If as a result hourly gross wages increase faster (slower) than below the cap, the distribution of gross earnings becomes less (more) concentrated which generates a discontinuous drop (jump) at the earnings cap. Faster (slower) earnings increases mean that a share of the saved labor costs (additional net earnings) benefits the employee (employer)².

Besides the technical but hardly restrictive assumption of smoothly distributed preferences, the outlined methodology is subject to several, arguably more far-reaching, assumptions:

- (1) There are no optimization frictions. Wages can be adjusted immediately after crossing the threshold according to the drop in the marginal SSC rate. The beneficiary of the reduced SSC liability must have borne the burden in the first place. If those two conditions are not met, the estimate is very selective and cannot be generalized. A potential discontinuity at an earnings cap is only informative about the sharing of the additional rent for employees with earnings closely above the cap. The incidence estimate is a local one in this framework. The potential impact of optimization frictions is simulated in section 3.
- (2) SSC rates and caps are perfectly salient. If this is not the case, perceived rates and caps might vary over individuals which might blur the discontinuity.
- (3) The burden of SSC is shifted at the individual level. An alternative channel of incidence might be shifting to the whole workforce which would invalidate the outlined method.
- (4) The object of negotiations about the sharing of the reduced SSC is restricted to labor earnings subject to SSC only. That is, there must not be adjustments at other margins of employees' compensation (premia, non-pecuniary benefits, paid/unpaid overtime).³

While assumptions (3) and (4) underlie most microeconomic evidence on economic incidence, assumptions (1) and (2) are particularly strong for the approach outlined here. On the other hand they are supported by several arguments: In a system with shared contributions an earnings cap is relevant for employers and employees. In particular employers arguably have good knowledge about the calculation of labor costs. This holds all the more in Germany where earnings caps have been in existence for a long time and have rarely been changed significantly

²Following this intuition a direct analysis of earnings *increases* would seem natural. One problem, however, is how to deal with new employment relationships. Second, such an analysis requires panel data – earnings in suitable administrative data sets in Germany are right-censored, though.

³If these margins are separately observed, they could be accounted for in the analysis.

(section 4.1). The caps are, however, subject to minor changes on a yearly basis to account for inflation. This adjustment might complicate optimization, if wages are not adapted accordingly.

The same arguments are also crucial for the question whether the estimates should be interpreted as long- or short-term effects. Assuming the earnings cap is perfectly 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 discontinuity rather estimates long-term effects. In contrast evaluations of policy reforms with longitudinal data analyze agents' short-run reactions to a change in incentives.

As shown by Alvaredo and Saez (2007), the analysis above has interesting implications for labor costs (z) and net earnings (c). Their deterministic relationships with gross earnings are described as follows:

$$z = \begin{cases} y(1+t) = \frac{n}{1+t*s}(1+t) & \forall y \leq \bar{y} \\ y + \bar{y}t = n - \bar{y}t * s + \bar{y}t & \forall y > \bar{y} \end{cases} \quad (3)$$

$$c = \begin{cases} y(1-t) = \frac{n}{1-t*s}(1-t) & \forall y \leq \bar{y} \\ y - \bar{y}t = n - \bar{y}t * s - \bar{y}t & \forall y > \bar{y} \end{cases} \quad (4)$$

It holds that $z = n$ ($c = n$) if $s = 1$ ($s = -1$), that is if the burden is shifted entirely to employees (employers). By similar arguments as above this implies that if $s = 1$ ($s = -1$), the distribution of labor costs (net earnings) is smooth and the distributions of gross earnings as well as net earnings (labor costs) are discontinuous. If economic and legal incidence coincide, the distribution of gross earnings is smooth and the distributions of net earnings as well as labor costs are discontinuous. In theory, this deterministic relationship between the three wage concepts makes it superfluous to analyze more than one distribution. In practice, e.g. under frictions or imperfect salience, however, analyzing all three can serve as a test of the power of the estimation.

2.1 Behavioral Responses

In a neoclassical labor supply model a positive labor supply elasticity makes it suboptimal 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 in practice.

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. Section 3 shows by means of simulations that it will then be possible to identify a potential discontinuity. This is true, even if economic incidence and behavioral responses are tightly

related. As argued above, a potential discontinuity identifies economic incidence only for individuals who locate close to the earnings cap. Those people should have small labor supply elasticities (otherwise they would not be there). Given that elasticity and economic incidence are correlated, this implies that the incidence estimate is highly selective.

A gap or dip around earnings caps has, however, neither been found in previous empirical work nor in this study. People with high earnings may have small labor supply elasticities or adjustment costs could be too high to locate optimally given the moderate increase in utility (Neumann, 2014).

2.2 Empirical Implementation

The empirical implementation of this framework is straightforward. We estimate the size of the discontinuities in the earnings distributions at the earnings cap. We follow two alternative strategies. First, in a non-parametric sorting test proposed by McCrary (2008) the density to the left and right of the cap is estimated by local linear regressions. The bandwidth is calculated by cross-validation. Second, in a parametric approach a high-order polynomial is fitted to the earnings distributions. The discontinuity is identified by a dummy variable discriminating between earnings below and above the cap. $t*s$ is calculated by normalizing the discontinuity estimate by the density directly left of the cap.

Choosing a bin width of the empirical earnings distribution is crucial for both approaches and directly linked to the identifying assumptions stated above. A larger bin size might, for example, attenuate the assumption on wage adjustments directly above the cap. Given a larger size, the bin directly above the cap contains a larger range of earnings. At the same time this also increases the natural discontinuity between two bins in a non-uniform distribution. This might in turn confound the actual discontinuity generated by economic incidence.

Similar to the more widespread bunching methods (Kleven et al., 2011; Kleven and Waseem, 2013), identifying a discontinuity at \bar{y} is demanding in terms of data quality. A large sample size is essential to reduce the noise in the distribution. In order to identify the location of \bar{y} , the exact amount of labor earnings subject to SSC has to be observed. Several problems might arise in that respect:

- (1) **Non-systematic measurement error** in gross earnings y could arise as a result of imprecisely/erroneously reported information. Let p_i be a random disturbance term for individual i , the observed gross earnings are $y_i^{obs} = y_i + p_i$. As we do not observe p_i , the assumed location of the earnings cap is $\bar{y} - p_i$ which varies across individuals in an unknown way. Thus, the discontinuity estimate might be biased towards zero.
- (2) **Systematic measurement error** is present in gross earnings y because certain

components that are not subject to SSC (premia for shift work, night employment, holidays) are included in the data, but cannot be observed separately. Moreover, Systematic measurement error appears if components that are subject to SSC (overtime hours) are not included. The consequences are similar to non-systematic measurement error.

- (3) **Unobserved variation in social security schemes** (private health insurance, private or occupational pension schemes) influences labor taxes and individual SSC thresholds. Again, the consequences are similar to non-systematic measurement error.

Since the consequences are very similar, we take the example of non-systematic measurement error and analyze the extent of the noise necessary to distort identification. We run several simulations on the basis of parameters of the observed earnings distribution (section 3). In empirical practice, however, the extent of these potential errors usually remains unobserved.

It might prove helpful to observe more than one earnings measure in the presence of non-systematic measurement error. Under the assumption that potential errors in different observed wage concepts are independent, using incidence estimates from different earnings distributions generates a more robust empirical foundation to draw policy conclusions. Consider the following example. The gross earnings distribution is found to be smooth around the cap. This could be explained by no shifting or measurement error. According to the first explanation the distributions of labor costs or net earnings should feature a discontinuity consistent with the drop of the marginal SSC rate. Under the alternative scenario of measurement error all three distributions would be found smooth around the cap; the results would be inconclusive as far as incidence is concerned. We exploit the separate measures of gross and net earnings in the GSES data to generate more robust empirical results in this regard (section 4).

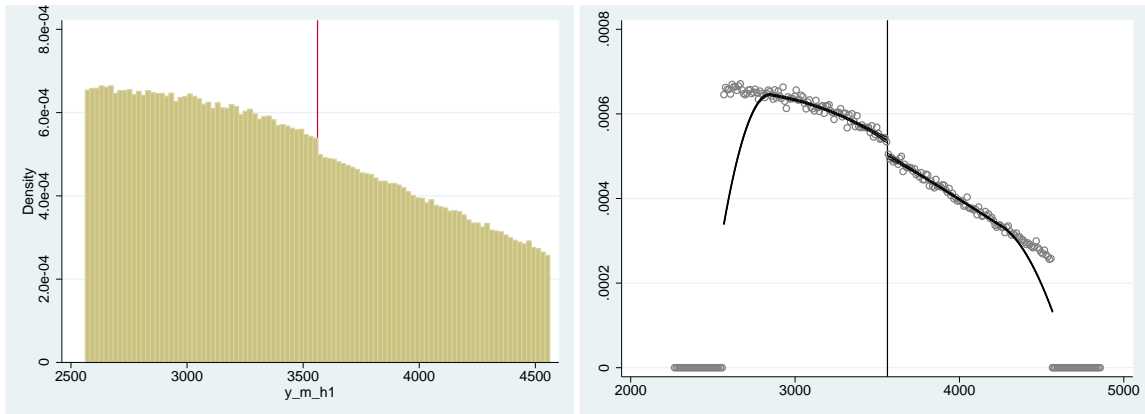
Most data sets suitable for this type of incidence analysis only include gross earnings. Labor costs and net earnings would have to be inferred by means of equations (3) and (4). By applying this deterministic relationship measurement error is simply perpetuated. Assume, for example, that non-systematic measurement error masks an actual discontinuity in the distribution of gross earnings. The distributions of ‘calculated’ net earnings and labor costs will then be found discontinuous. The substantial interpretation of no shifting that would be consistent with these findings is in fact not true. The deterministic relationship between the earnings concepts can still be exploited: We might obtain indirect evidence that noise in the data weakens identification at the earnings cap. The differences between the discontinuity estimates in the respective distributions might not reflect the drop in the marginal SSC rate. Alternatively statistically significant discontinuities are detected in less than two distributions.

3 Simulation exercises

The objective of this section twofold: First, we demonstrate the method on simulated data. Second, we assess the potential threat for identification posed by a violation of assumptions or data problems. The latent wage distribution used for our simulation exercises is assumed to be normally distributed and simulated on the basis of the parameters from the 2006 GSES wage data for West Germany (mean=2,913 €/month, s.d.=1,481 €/month, N=1.6 million). Based on these parameters, a gross earnings distribution is calculated that features a negative discontinuity. Its size is consistent with a full shifting of employers' SSC to employees. All simulations are repeated 1000 times with the point estimates equaling the mean among repetitions and the reported confidence intervals containing 95% of estimated values.

The left panel of Fig. 1 shows the resulting gross wage distribution around the German earnings cap of health insurance in 2006 where the marginal SSC rate drops by 7% (for employees and employers, respectively). The right panel of Fig. 1 applies the McCrary test as described in section 2.2 and correspondingly finds a negative discontinuity of 7.3% (see first row in table 1). The confidence interval translates to the interval of shifting parameters (-1.29,-0.68).

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



Source: Own simulations.

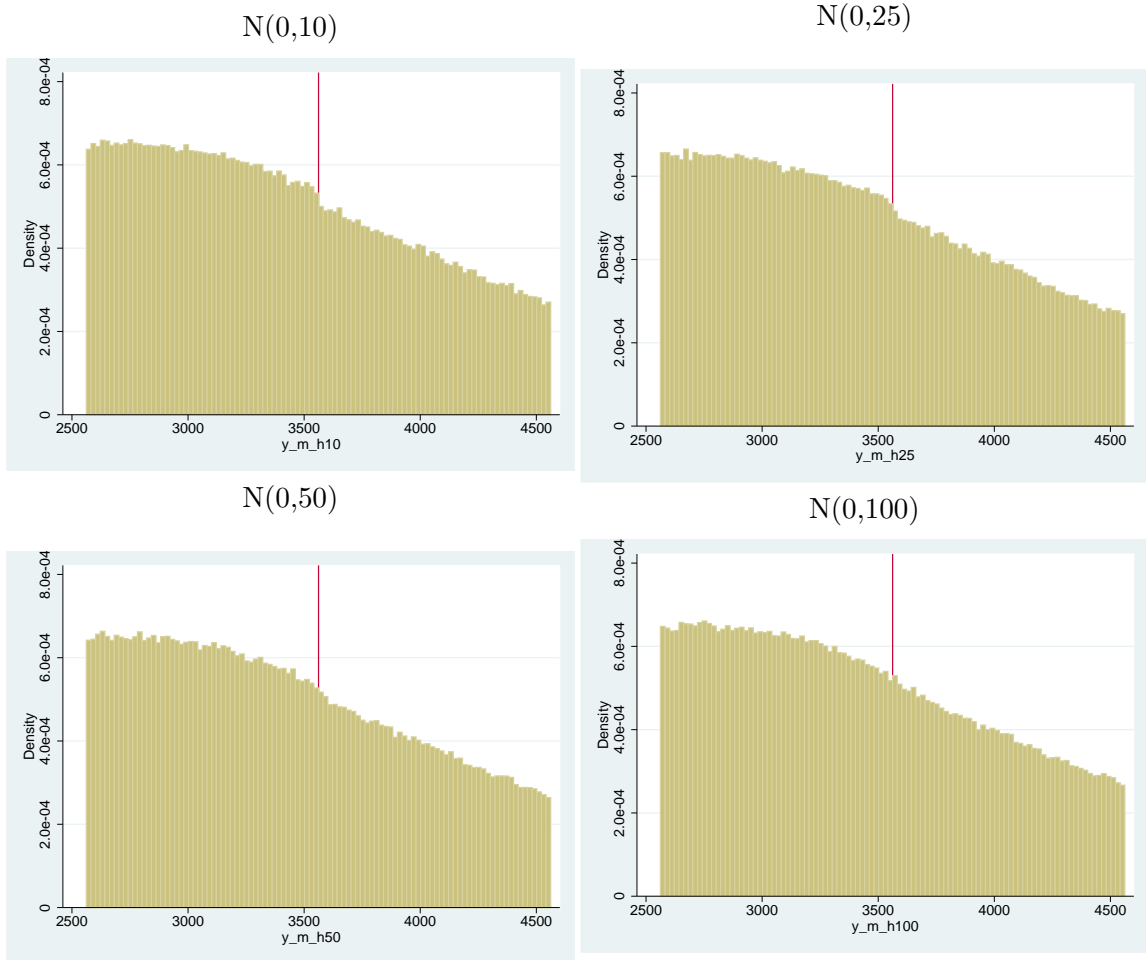
The parametric approach also identifies the correct discontinuity (see right panel of table 1). The confidence interval is slightly narrower than for the McCrary-test. However, although the sample size and variance is taken from the GSES wage data, the simulated earnings distribution appears to feature less noise than the observed earnings distribution. Therefore, confidence intervals which are estimated based on the simulated data represent a lower bound of uncertainty.⁴

On the basis of this gross earnings distribution we simulate potential distortions and vary their extent to assess their potential to distort identification. Table 1 re-

⁴We will mimic the actual noise more realistically in our simulations in the next version of this paper.

ports the corresponding discontinuity estimates and their confidence intervals. First, we exemplify the consequences of different data problems using a simulation of non-systematic measurement error. The random element p is added to the gross earnings. We assume p to be normally distributed with mean zero. The standard deviation (s.d.) varies between 10 €/month (upper left panel of Fig. 2) and 100 €/month (lower right panel of Fig. 2).

Fig. 2: Distributions of simulated gross earnings with normally distributed random measurement error



Source: Own simulations.

The estimated magnitude of the discontinuity is reported in rows 2 to 5 of table 1. For both the parametric and the non-parametric approach the estimate becomes weaker by increasing the s.d. of the added distortion. With the McCrary-test a s.d. of 25 €/month roughly halves the discontinuity estimate while statistical significance is lost with a s.d. of 50 €/month. The parametric approach is much less affected by the distortion. The estimate is still statistically significant for a distortion with a s.d. of 100 €/month, albeit the estimated size is just one third of the true value. A reason for the discrepancy between both approaches is that observations close to the earnings cap have less weight in the parametric estimation.

Tab. 1: Discontinuity estimates – simulated gross earnings

Simulated problem		McCrary		Polynomials	
		p.e.	CI	p.e.	CI
(1)	Basic	-0.73	[-.094 ; -.050]	-0.68	[-.08 ; -.055]
<i>Random component</i>					
(2)	$p \sim N(0, 10)$	-0.55	[-.078 ; -.033]	-0.62	[-.075 ; -.05]
(3)	$p \sim N(0, 25)$	-0.30	[-.057 ; -.004]	-0.55	[-.067 ; -.043]
(4)	$p \sim N(0, 50)$	-0.19	[-.042 ; .005]	-0.44	[-.057 ; -.03]
(5)	$p \sim N(0, 100)$	-0.05	[-.026 ; .017]	-0.26	[-.039 ; -.014]
<i>Timing</i>					
(6)	$r \sim U(0, 50)$	-0.63	[-.084 ; -.041]	.066	[-.08 ; -.054]
(7)	$r \sim U(0, 100)$	-0.41	[-.062 ; -.018]	.063	[-.076 ; -.05]
(8)	$r \sim U(0, 150)$	-0.19	[-.041 ; .004]	.057	[-.071 ; -.045]
<i>Rand. comp. and Timing</i>					
(9)	$p \sim N(0, 25); r \sim U(0, 100)$	-0.16	[-.043 ; .010]	-0.06	[-.072 ; -.048]
<i>Behavioral responses</i>					
(10)	$P(e \sim U(0, 0.7)) = 0.65;$ $P(e = 0) = 0.35; s = 1$	-0.59	[-.093 ; -.023]	.054	[.033 ; .075]
(11)	$P(e \sim U(0, 0.7)) = 0.65;$ $P(e = 0) = 0.35; s = 1 - e$	-0.57	[-.097 ; -.018]	.021	[-.001 ; .041]
(12)	$P(e \sim U(0, 0.7)) = 0.825;$ $P(e = 0) = 0.175; s = 1 - e$	-0.40	[-.106 ; .022]	.081	[.054 ; .107]
<i>Heterogeneous incidence</i>					
(13)	$s \sim U(0, 1)$	-0.37	[-.056 ; -.016]	-0.35	[-.047 ; -.022]
(14)	$s \sim U(0, 1)$	-0.36	[-.063 ; -.011]	-0.34	[-.046 ; -.022]

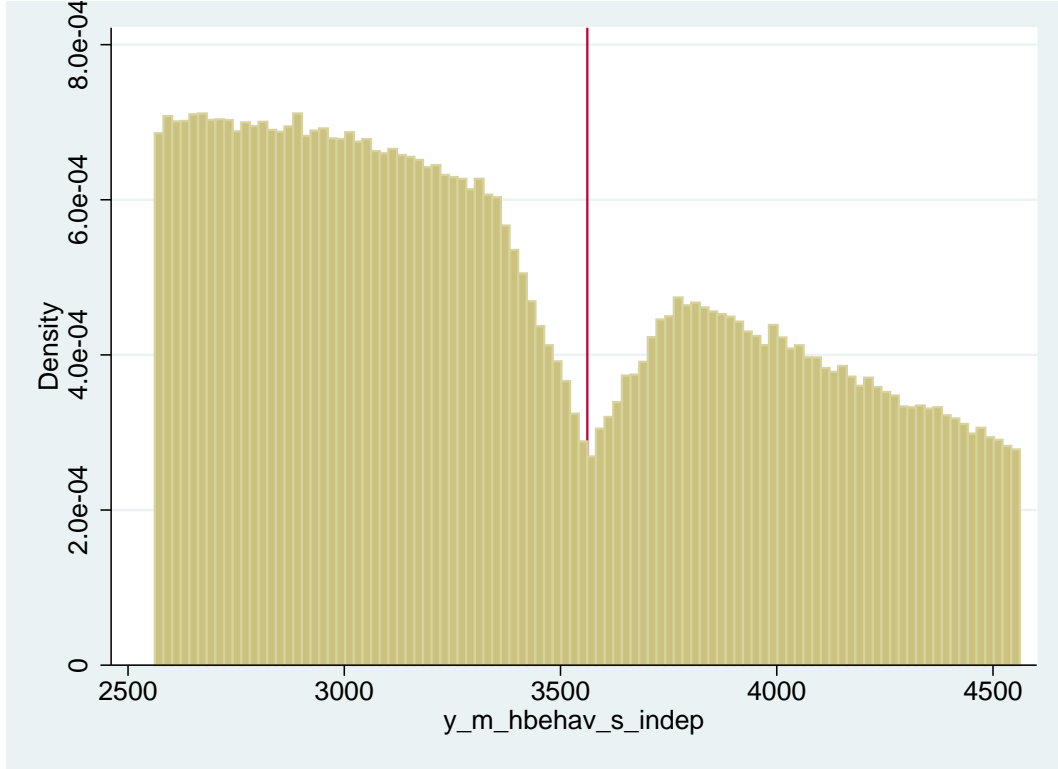
Notes: p.e. – point estimate, CI – 95% confidence interval.

Source: GSES 2001, 2006; own calculations.

We also simulate optimization frictions by *delaying* the wage rate responses to crossing the earnings cap. Within our static framework, this is modeled such that the drop in the marginal tax rate is first accounted for when earnings are above a random threshold $r > \bar{y}$. The location of the cap is, however, perceived correctly. r is simulated to be uniformly distributed within a varying interval of 0 to up to 150€/month. Rows 6 to 8 of Tab. 1 pertain to these simulations. Using the McCrary-test, delaying the renegotiation up to 100€/month decreases the point estimate to -0.4 while statistical significance is retained. With an upper bound of the delay of 150€/month no significant discontinuity can be detected. Again, the parametric estimates are much less impacted by the distortion such that it is only decreased by 0.01 due to a delay with an upper bound of 150€/month. In general, confidence intervals seem to increase slightly due to the introduced distortions for the McCrary-test but not for the parametric approach.

The two potential problems of measurement error and optimization frictions are likely to occur simultaneously. Row 9 of Tab. 1, therefore, reports estimates

Fig. 3: Distribution of simulated gross earnings with behavioral responses



based on a distribution distorted by non-systematic measurement error as well as optimization frictions. Using the McCrary-test no significant discontinuity is found if both distortions are combined although they are not large enough to entirely destroy identification separately (see rows 3 and 7). Strangely, the estimate based on the parametric approach is even slightly closer to the true value if both distortion are combined relative to adding only random measurement error.

Fig. 3 depicts the effect of behavioral responses which are simulated according to the model in Saez (2010). For roughly two thirds of the employees the labor supply elasticity, e , is simulated to be uniformly distributed in the interval $(0, 0.7)$ but independent from economic incidence which is still assumed to be fully with employees. For the remaining employees the assigned elasticity is set to zero such that the distribution features a dip instead of a gap around the earnings cap. Nevertheless, row (10) of Tab. 1 shows that the estimate of the discontinuity based on the McCrary-test is hardly distorted. However, the estimate is less precise than without behavioral responses. By contrast, the estimate based on the parametric approach is completely out of place and even changes signs. It seems that the parametric approach is not flexible enough to fit the dip.

Rows (11) and (12) of Tab. 1 let economic incidence be directly determined by the labor supply elasticity ($s = 1 - e$). This relationship is based on the neo-classical labor market model where highly elastic employees bear less of the SSC burden. In row (11) the elasticity is distributed as in row (10). Therefore, an average of

$s = 0.77$ results in a true discontinuity of -0.054 . The McCrary test almost exactly delivers that value. In row (12) the fraction with $e = 0$ is reduced to 0.175. The true discontinuity decreases to -0.05 . Although the point estimate is only slightly distorted (-0.040) it is not significantly different from 0 at the 5% confidence level. Again, the estimates based on the parametric approach are completely out of place.

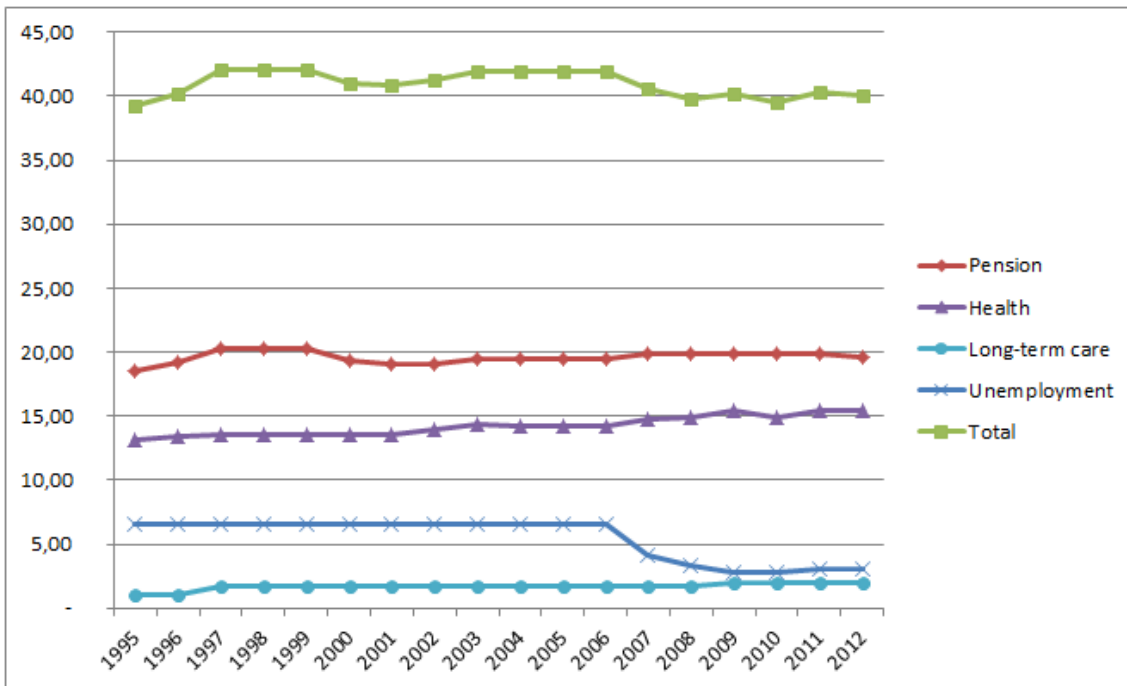
The last row of Tab. 1 pertains to simulating the violation of the assumption of a homogeneous shifting parameter. Shifting s is simulated to follow a uniform distribution in the interval $(0,1)$ and, therefore, equals 0.5 on average. The estimates vary between -0.034 and -0.037 suggesting that in the case of heterogeneous economic incidence, the outlined approach identifies average economic incidence.

4 Application to Germany

4.1 Institutions

The German social security system consists of pension, health, unemployment, and long-term care insurance. The contribution rates are flat with daily gross earnings as tax base. Unit 2004 SSC have been equally shared between employees and employers. Since 2005 0.9 pp.⁵ are exclusively paid by employees. Most important for this study, marginal SSC rates only apply up to earnings caps which differ between pension (and unemployment) insurance and health (and long-term care) insurance⁶ as well as East and West Germany. The evolutions of the SSC rates and earnings caps since 1995 are shown in Fig. 4 and 5. 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. 4: Development of SSC rates over time



Source: <http://www.statistik.baden-wuerttemberg.de>

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).

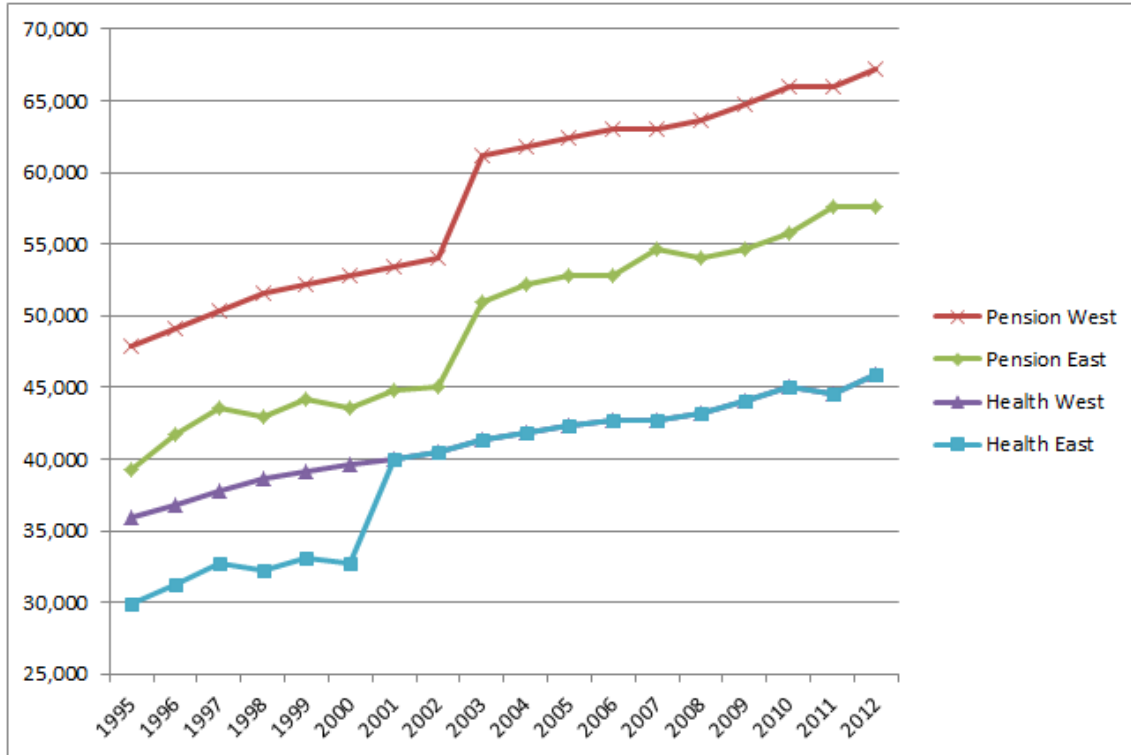
Both SSC rates and earnings caps are subject to yearly gradual changes which are difficult to exploit for estimating economic incidence. Considerable discontinuous changes are rare with the strong increases of the earnings cap of health insurance in East Germany in 2001 and pension insurance in 2003 as notable exceptions. While the former is evaluated by Neumann (2014), the latter is difficult to analyze because

⁵Plus 0.25 pp. for childless employees

⁶For the sake of readability we do not refer to unemployment and long-term care insurance in the following.

the largest German panel data set is right-censored at this earnings cap. This demonstrates the advantage of the cross-sectional approach outlined above which is not dependent on policy reforms.

Fig. 5: Development of earnings caps over time



Source: <http://www.statistik.baden-wuerttemberg.de>

One peculiarity of the German social security system which might have an impact on the analysis is the possibility to substitute public for private health insurance. The possibility of opting for private health insurance depends on exceeding the income threshold for compulsory insurance (*Versicherungspflichtgrenze*; except for the self-employed and civil servants which we exclude here). It had been equal to the earnings cap of health insurance until both caps were decoupled in 2003 and has been between the earnings caps of health and pension insurance since then. The employees' share of SSC of a privately insured employee does not depend on earnings but on personal characteristics. Employers' pay half of private contributions but only up to the maximum value of employers' SSC of a publicly insured employee.

Private health insurance only impacts our analysis if it affects the change in the marginal SSC rate at the analyzed earnings cap which is only the case for the earnings cap of health insurance before 2003⁷. Being slightly below or above the earnings cap might change the whole insurance system. Although we do not observe whether an employee is privately health insured, it is not obvious how severely

⁷As 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 section 4.2 for a detailed discussion

this distorts the analysis due to two reasons. First, as the cap on employers' SSC corresponds in both regimes, their treatment by crossing the threshold is comparable as long as private is at least as expensive as public health insurance which usually is the case (Neumann, 2014). Second, the switch to private health insurance usually occurs earliest in the year after earnings crossed the threshold for the first time. That suggests that a large fraction of employees with earnings very closely above the earnings cap are still publicly health insured (Neumann, 2014).

4.2 Data

4.2.1 The GSES

We use the German Structure of Earnings Survey (GSES, *Verdienststrukturerhebung*) which is a repeated cross-section. 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. Reliability of the individual earnings information is thus better than in surveys which is why non-random measurement error arguably is not a big issue (although we cannot determine its magnitude). We primarily use monthly earnings information that refers to the month of October. As 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), non-systematic measurement error is neither an issue.

This is the only data set for Germany with separate measures of gross earnings and employees' SSC. The independent variable on actually paid SSC includes an employee's contributions (also voluntary) to social insurance (i.e. the employee's share of contributions to the pension, unemployment, health and care insurance). This also includes contributions to private health insurance as well as occupational insurance schemes. Explicit information on participation in different schemes is not available, though. This impacts identification in the analysis of the distribution of net earnings. For example, privately health insured employees with $y > \bar{y}^h$, $y \leq \bar{y}^p$ have *observed* net earnings $d = y - ssc_{obs} = y(1 - t_e^p) - ssc_{private}^h$. The location of the earnings cap in the net earnings distribution is $\bar{d}_i^h = \bar{y}(1 - t_e^p) - ssc_{i,private}^h$. However, as we do not observe, whether an employee is privately health insured, we approximate \bar{d}_i^h by $\hat{d}^h = \bar{y}(1 - t_e^p) - t_e^h \bar{y}^h$. This might blur a potential discontinuity.

There is no independent information on the employer's SSC. We focus this note therefore on tests of gross earnings y and net earnings d and do not separately consider labor costs z . The GSES some further limitations. Small firms (below 10 employees) are exempted because of the administrative burden. Single sectors are also missing; it is thus not fully representative for the German employers and employees. Compared to other administrative data sets, the GSES's sample size is

not as comprehensive in a given wave and region. To circumvent this limitation, we pool the two waves of data currently available to us (2001 and 2006).

4.2.2 Sample

We use two waves of data for the years 2001 and 2006 in the current version of the paper.⁸ 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, furthermore, exclude part time and marginally employed persons to limit behavioral adjustments. (The earnings caps are hardly relevant for these types of employees, anyway). In general only those employees working at least 35 hours per week are considered in the sample. The sample comprises 641,635 (34.39%) employees for the wave 2001 and 1,224,300 (65.61%) employees for the wave 2006. We analyze the pooled sample as well as single waves. East and West Germany are pooled generally, although the pension cap still differs between those regions.

4.2.3 Indexation and pooling

We analyze discontinuities at the caps for *health* and *pension* insurance \bar{y}^h, \bar{y}^p . There is variation in caps \bar{y}_{rt} over time t and – for the pension insurance cap – by region r (i.e. between West and East Germany). Therefore we index all y_{rt} in the following simple way: $y_{rt}^{ind} = y_{rt} - \bar{y}_{rt}$. The resulting y_{rt}^{ind} is measured in deviations from the threshold. This allows us to pool data for different r and t – at the moment waves 2001 and 2006 as well as East and West Germany – to get a larger sample size and more power for the estimations performed. Note that legal earnings caps are defined in terms of gross earnings y . The indexation of net earnings d at the cap is a little more intricate since the SSC rates applied change due to the health insurance earnings cap \bar{y}^h .

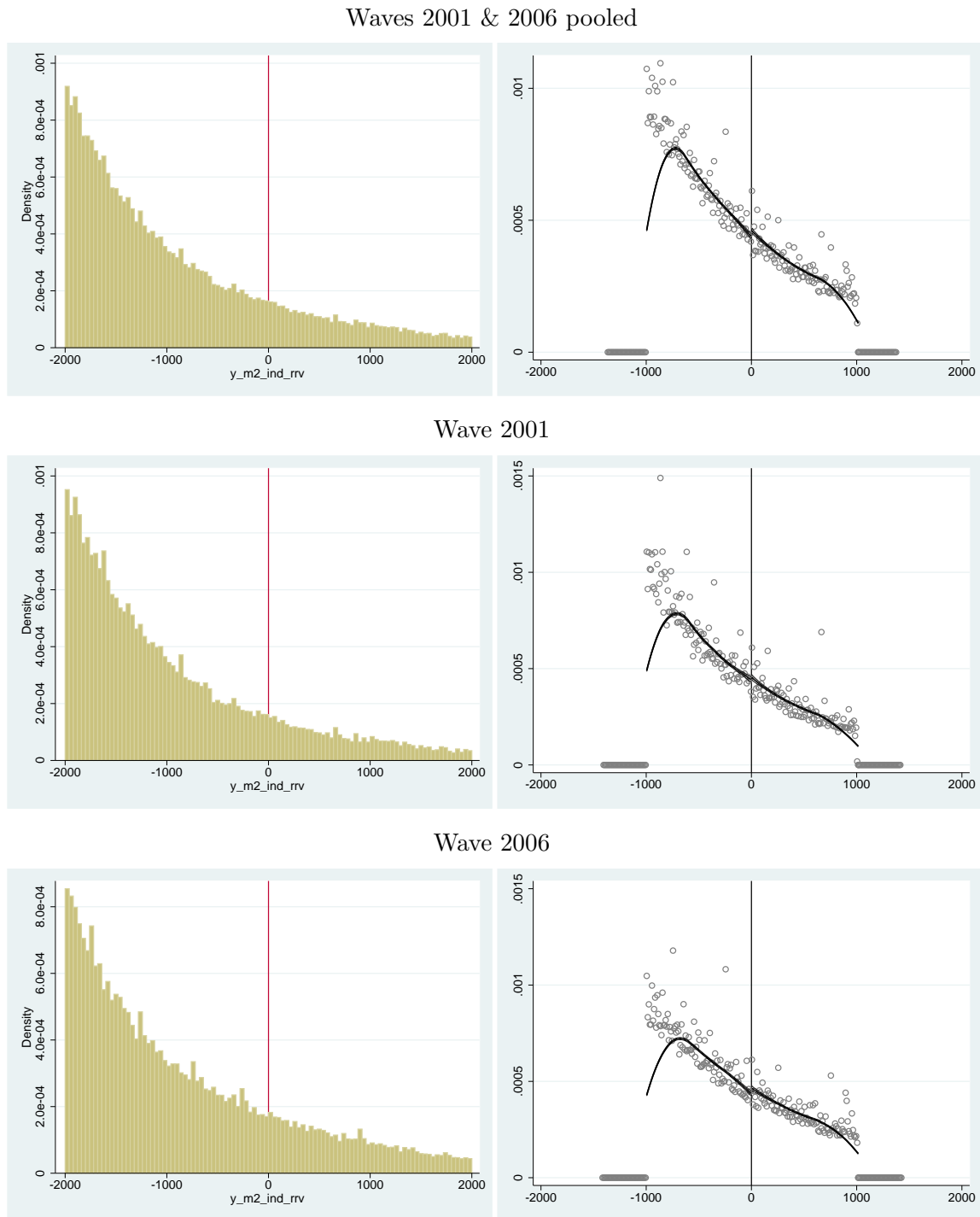
4.3 Empirical results

4.3.1 Gross earnings

It is hard to identify a clear-cut discontinuity in histograms of the gross earnings distribution for the pension (Fig. 6) or health (Fig. 7) cap. This holds for the pooled sample and for single waves. Applying the McCrary test, however, yields a statistically significant positive discontinuity estimate for the pension cap in the pooled sample (see Tab. 2). This is also true for 2006 but not for 2001 where the point estimate is positive but not statistically significant. Applying polynomials, point estimates are similar but the estimation seems to be much less precise resulting in insignificant estimates. This is surprising because in the simulations the parametric approach usually delivered more precise estimates.

⁸We will exploit all four available waves in the next version of the paper.

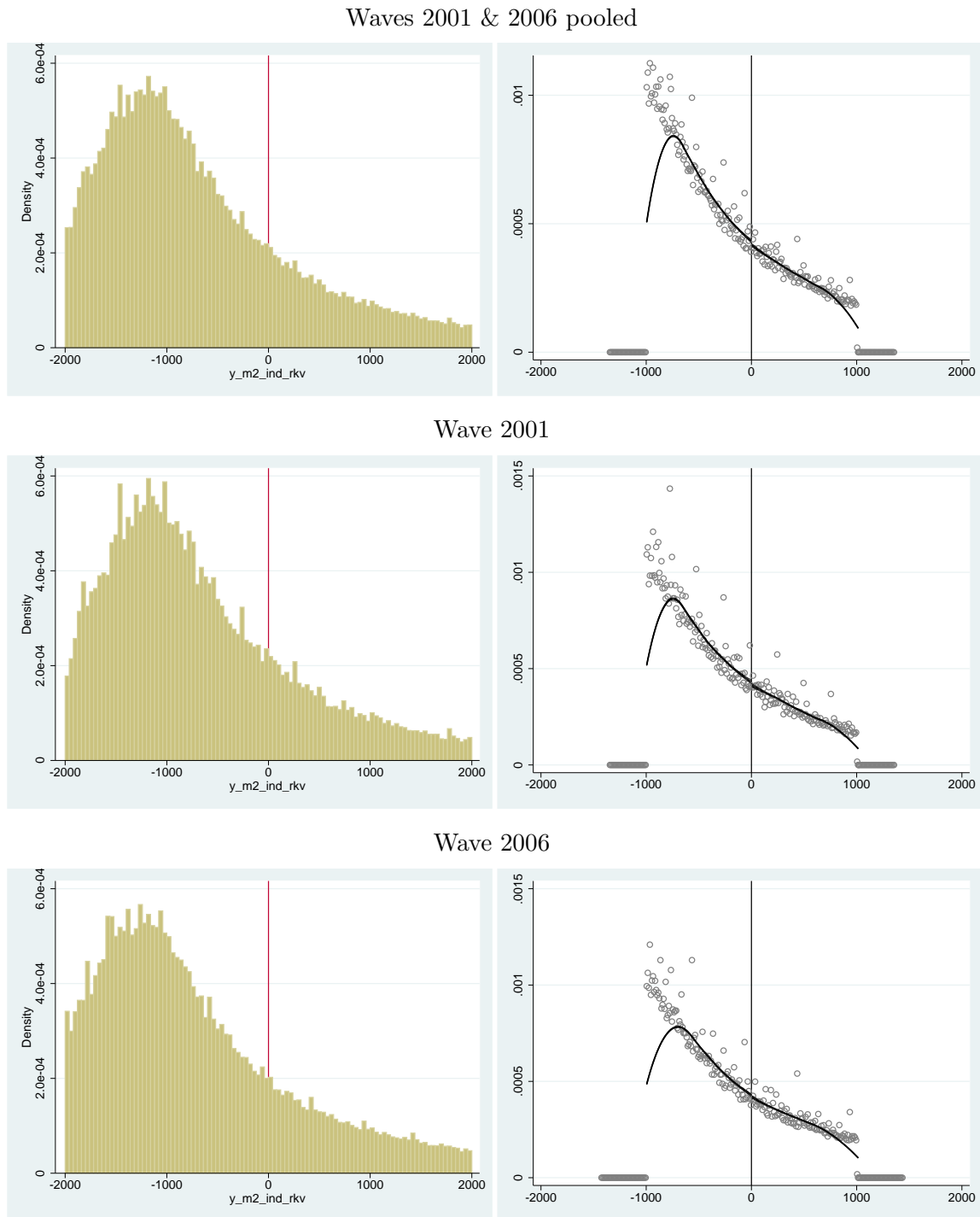
Fig. 6: Distribution of gross earnings, histograms & discontinuity estimates, pension insurance cap



Source: GSES 2001, 2006; own calculations.

At the health cap we find a statistically significant negative discontinuity with the McCrary test for the pooled sample and in 2001. Again, the point estimates produced by the parametric approach are similar but not statistically significant. Histograms for single waves are more spiky than for the pooled sample. Accordingly, the estimates based on the pooled sample are more precise.

Fig. 7: Distribution of gross earnings, histograms & discontinuity estimates, health insurance cap



Source: GSES 2001, 2006; own calculations.

Taken at face value this implies that the economic incidence of SSC to pension (health) insurance is a bit more with employers (employees) compared to the legal incidence. Although this pattern is found for both single waves as well as for pooled data, shifting of SSC to pension (health) insurance is more pronounced in 2006 (2001). More concretely, the estimates of the McCrary test for the pooled sample

mean that approximately 42% of the employees' SSC to pension insurance is shifted to employers whereas approximately 38% of the employers' SSC to health insurance is shifted to employees. The main difference between the two schemes is the link between contributions and benefits. However, as it is stronger for pension insurance, theory predicts that economic incidence of SSC to pension insurance are rather shifted to wages (Summers, 1989) which is the opposite of the observed pattern. Thus, other than different income levels (the cap is on average about 3336, 3562 for health and 3732 (4448), 4400 (5250) for pension insurance in East (West) Germany for the years 2001, 2006 in € per month), there are no explanations for such a discrepancy.

Tab. 2: Discontinuity estimates – gross earnings

Wave	McCrary		Polynomials	
	p.e.	t	p.e.	t
<i>Pension cap</i>				
Pooled	.055	3.52	.049	.75
2001	.032	1.44	.052	.53
2006	.068	3.49	.047	.54
<i>Health cap</i>				
Pooled	-.031	-3.21	-.038	-.72
2001	-.043	-2.71	-.051	-.61
2006	-.017	-1.58	-.03	-.44

Notes: p.e. – point estimate, t – t-value.

Source: GSES 2001, 2006; own calculations.

In order to informally test how likely a discontinuity of a certain magnitude is generated by random sampling error [XXX: dafür sind doch eigentlich die CIs da, oder?], we test robustness by estimating *placebo discontinuities* at arbitrary thresholds (Tab. 3). For this purpose we use the distribution above the pension cap and below the health cap but not in between. We find several statistically significant estimates at arbitrary thresholds which almost reach (exceed) the magnitude of the estimate at the true pension (health) cap. Thus, we cannot rule out that the estimates at the earnings caps in gross earnings are generated by random noise in the data. Therefore the patterns in the results should be interpreted very cautiously.

Further, we test the robustness of the estimates by varying the bin size used to depict the earnings distribution. Table 4 shows that the estimates are quite robust in that regard.

Tab. 3: Discontinuity estimates – gross earnings, sensitivity w.r.t. breakpoint, pooled

Breakpoint	McCrory		Polynomials	
	p.e.	t	p.e.	t
$\bar{y}^h - 1000$	-.05	-8.25	-.034	-.87
$\bar{y}^h - 800$.041	5.99	.034	.76
$\bar{y}^h - 600$.033	4.64	.039	.77
$\bar{y}^h - 400$	-.004	-.47	-.018	-.34
$\bar{y}^h - 200$.007	.82	.025	.45
$\bar{y}^p + 200$	-.011	-.7	.006	.1
$\bar{y}^p + 400$.047	2.82	.042	.56
$\bar{y}^p + 600$.017	.91	.014	.17
$\bar{y}^p + 800$	-.031	-1.54	.034	-.44
$\bar{y}^p + 1000$.048	2.39	.035	.41

Notes: p.e. – point estimate, t – t-value.

Source: GSES 2001, 2006; own calculations.

Tab. 4: Discontinuity estimates – gross earnings, sensitivity w.r.t. bin size

Bin Size	Pension				Health			
	McCrory		Polynomials		McCrory		Polynomials	
	p.e.	t	p.e.	t	p.e.	t	p.e.	t
1	.059	4.03	.055	.83	-.028	-3.34	-.039	-.67
100	.044	2.67	.064	2.64	-.027	-3.09	-.042	-.95

Notes: p.e. – point estimate, t – t-value.

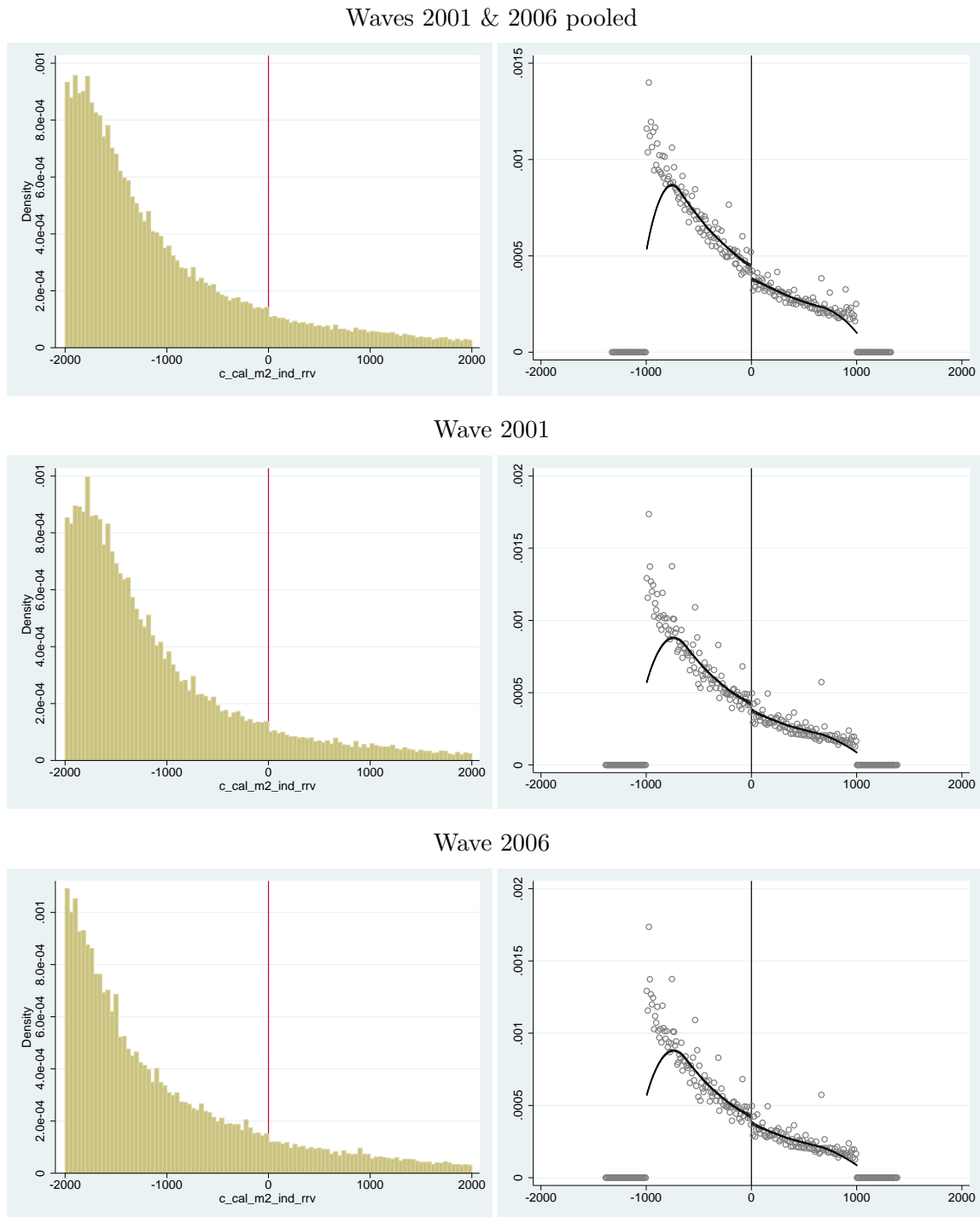
Source: GSES 2001, 2006; own calculations.

4.3.2 Net earnings

We compare the distributions of ‘calculated’ and ‘observed’ net earnings (as defined in sub-section 2.2 above). There are visible discontinuities in the histograms of *calculated* net earnings, more distinctly for the pension (Fig. 8) than for the health insurance cap (Fig. 9). This manifests itself also in the significant and large discontinuity estimates for both the McCrory and the parametric tests (left panel of Tab. 5).

As derived in section 2, the difference between the discontinuities in the gross and *calculated* net earnings distributions by definition corresponds to the drop in the marginal SSC rate (for one side of the market), which amounts to approximately 13 pp. at the pension cap and 8 pp. at the health cap. For health insurance this is approximately the case for all estimates. For pension insurance, however, only the McCrory estimates in 2001 are approximately consistent. The differences between the estimates of gross and net earnings is much too large for pooled data (21.6 pp.)

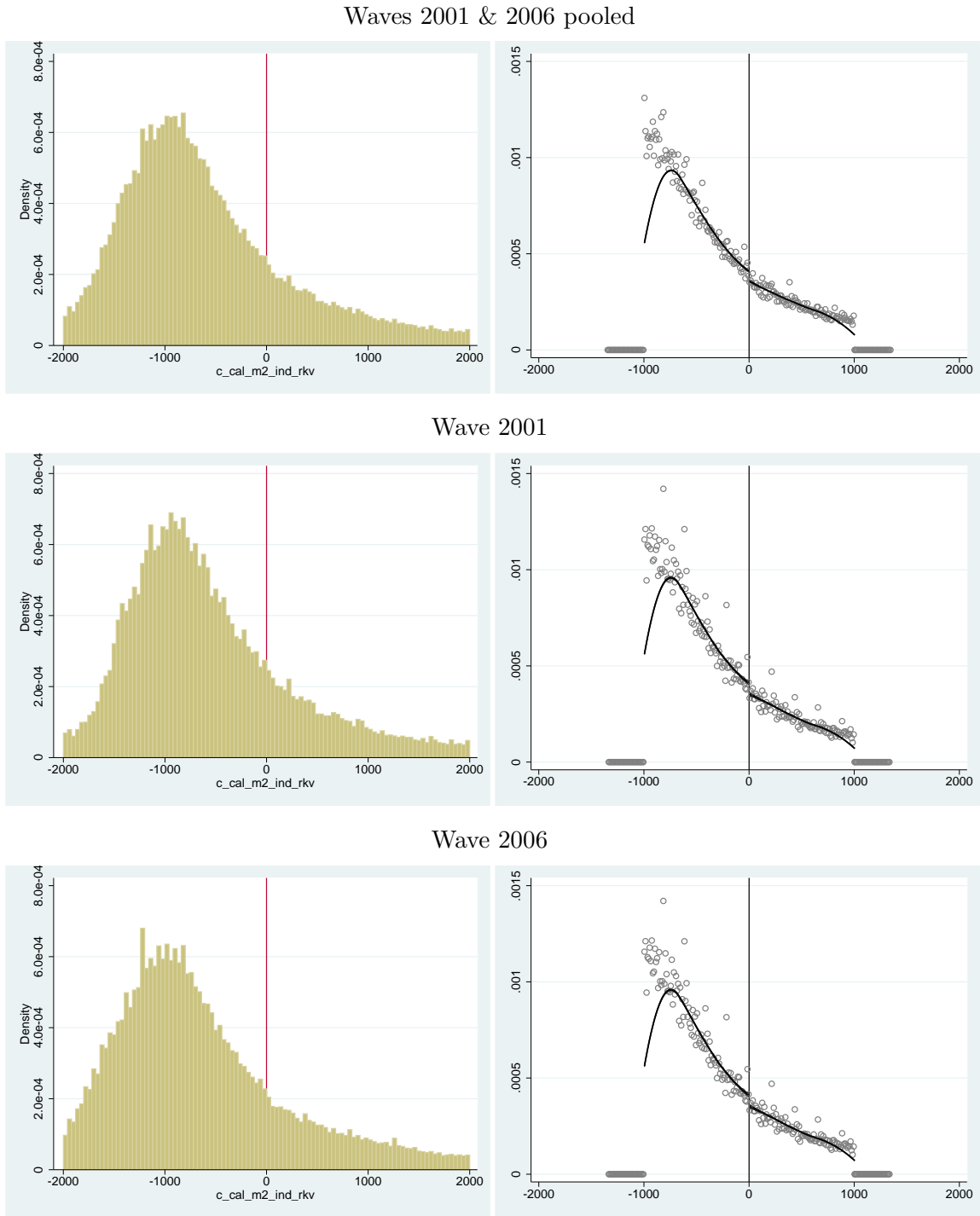
Fig. 8: Distribution of ‘calculated’ net earnings, histograms & discontinuity estimates, pension insurance cap



Source: GSES 2001, 2006; own calculations.

as well as 2006 (25.3 pp.). This suggests that the analysis of the earnings cap of pension insurance is very imprecise while this is not the case for health insurance. One reason for that difference might be that the amount of observations is much larger around the health cap than around the pension cap. Taken at face value, the estimates of *calculated* net earnings imply that economic incidence of both systems

Fig. 9: Distribution of ‘calculated’ net earnings, histograms & discontinuity estimates, health insurance cap



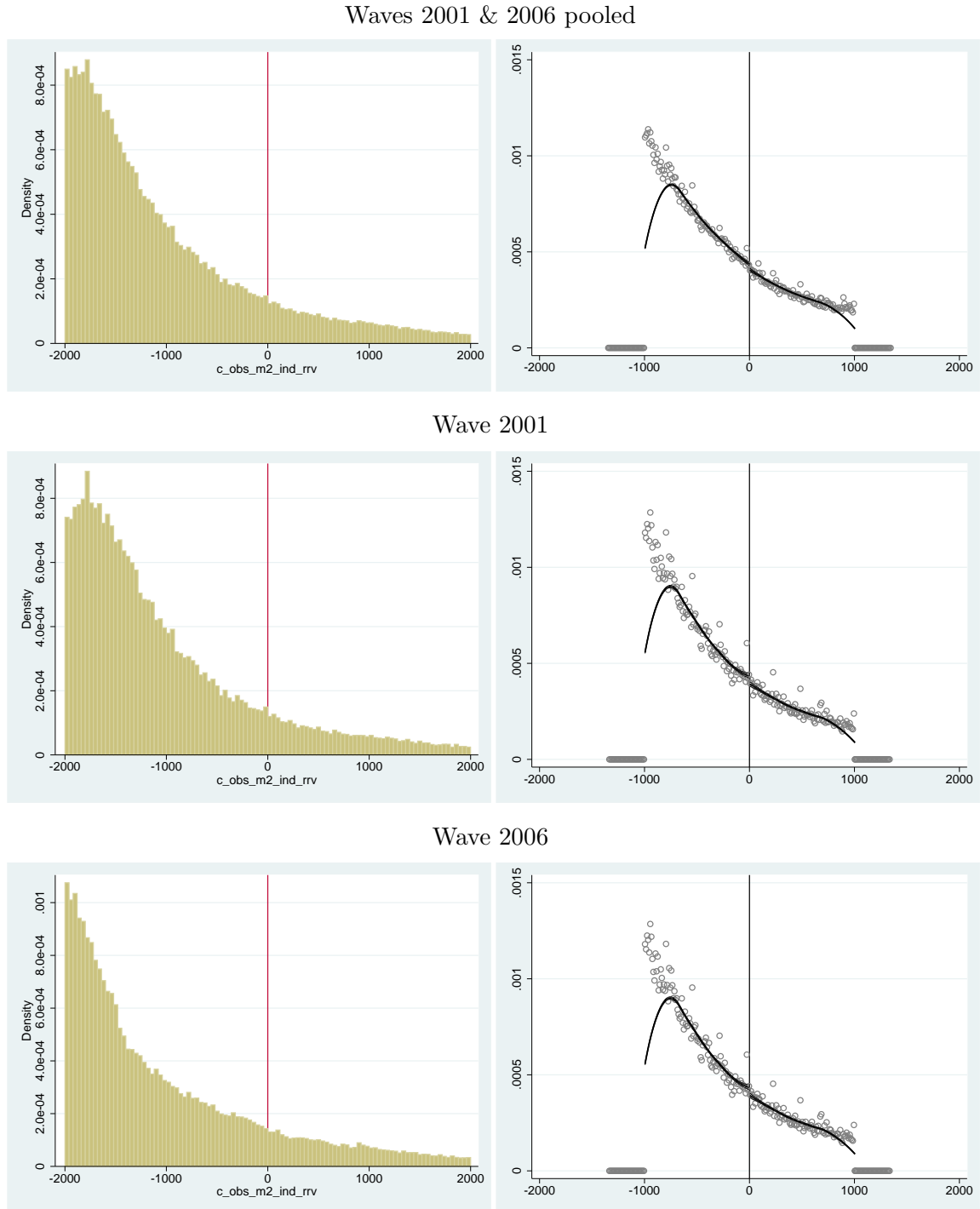
Source: GSES 2001, 2006; own calculations.

is slightly more with employees.

Turning to *observed* net earnings we see that discontinuities are much less pronounced at the caps compared to *calculated* earnings. This holds for the pension (Fig. 10) and the health cap (Fig. 11). This is also reflected by on average smaller point estimates (see right panel of Tab. 5). However, in most cases the estimates

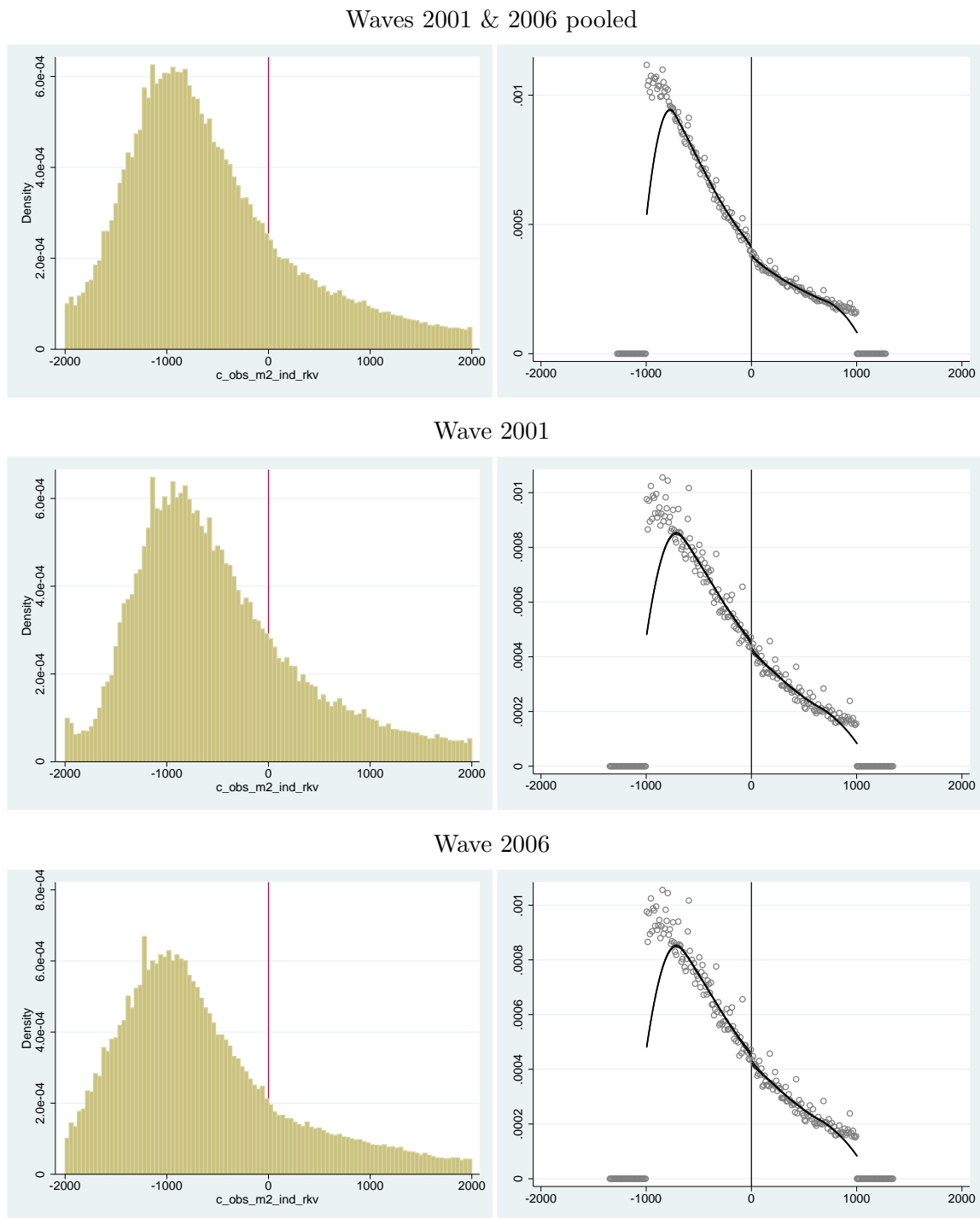
are still statistically significant. We get nearly identical discontinuity estimates for the pension cap for ‘calculated’ and ‘observed’ earnings (-.1) in 2001. In 2006 there is a wide discrepancy. For the health insurance cap we find the reverse pattern: there is a marked discrepancy between ‘calculated’ and ‘observed’ earnings in 2001, whereas for 2006 both discontinuity estimates almost coincide (at -.11).

Fig. 10: Distribution of ‘observed’ net earnings, histograms & discontinuity estimates, pension insurance cap



Source: GSES 2001, 2006; own calculations.

Fig. 11: Distribution of ‘observed’ net earnings, histograms & discontinuity estimates, health insurance cap



Source: GSES 2001, 2006; own calculations.

As discussed above, the main systematic difference between ‘calculated’ and ‘observed’ net earnings are contributions to private health insurance (and other voluntary contributions). As this induces the locations of the caps in the observed net earnings distribution to vary in an unknown way, estimates are biased towards zero which is consistent with the smaller point estimates found for ‘observed’ net

Tab. 5: Discontinuity estimates – net earnings

Wave	Calculated				Observed			
	McCrary		Polynomials		McCrary		Polynomials	
	p.e.	t	p.e.	t	p.e.	t	p.e.	t
<i>Pension cap</i>								
All	-.161	-10.04	-.181	-3.3	-.065	-4.41	-.073	-2.35
2001	-.105	-4.68	-.196	-2.25	-.097	-4.59	-.065	-1.16
2006	-.185	-9.84	-.166	-2.07	-.031	-1.49	-.073	-2.35
<i>Health cap</i>								
All	-.13	-14.65	-.146	-2.6	-.082	-8.47	-.085	-3.49
2001	-.148	-9.99	-.139	-1.75	-.06	-4.56	-.047	-1.1
2006	-.114	-11.37	-.147	-2.04	-.106	-8.07	-.11	-4.3

Notes: p.e. – point estimate, t – t-value.

Source: GSES 2001, 2006; own calculations.

earnings. In 2006 the threshold for compulsory health insurance had been raised and lay between the health and the pension cap. Therefore, the only clean estimate for observed net earnings (at least with respect to private health insurance) is the cap for health insurance in 2006 where indeed both estimates coincide. However, the difference between the estimates of pension insurance in 2001 and 2006 cannot be explained by private health insurance.

Tab. 6: Discontinuity estimates – net earnings, sensitivity w.r.t. breakpoint

Breakpoint	Calculated				Observed			
	McCrary		Polynomials		McCrary		Polynomials	
	p.e.	t	p.e.	t	p.e.	t	p.e.	t
$\bar{y}^h - 1000$.058	10.56	.022	.58	.012	1.84	-.001	-.03
$\bar{y}^h - 800$	-.05	-9	.021	.56	-.032	-5.19	-.015	-.85
$\bar{y}^h - 600$.006	1.11	.001	.03	.024	3.93	.02	1.05
$\bar{y}^h - 400$	-.043	-5.82	-.025	-.49	-.019	-2.26	-.018	-.82
$\bar{y}^h - 200$	-.008	-.89	.018	.36	.012	1.27	.03	1.29
$\bar{y}^p + 200$.03	1.77	.052	.7	.053	2.79	.039	1.11
$\bar{y}^p + 400$.039	2.21	.045	.58	.02	1.08	.029	.79
$\bar{y}^p + 600$	-.026	-1.42	-.044	-.55	-.046	-2.39	-.068	-1.88
$\bar{y}^p + 800$	-.046	-2.1	-.012	-.16	-.056	-2.51	.002	.04
$\bar{y}^p + 1000$	-.043	-2.18	-.038	-.57	-.036	-1.68	-.023	-.59

Notes: p.e. – point estimate, t – t-value.

Source: GSES 2001, 2006; own calculations.

Table 6 estimates placebo discontinuities at arbitrary thresholds. Almost none of the point estimates comes close to the estimates at the actual caps. This is

especially true for arbitrary thresholds below the health cap. This, again, suggests that the estimate for health insurance is much more robust due to the larger amount of observations in that region of the earnings distribution.

Taking the analyses of gross and net earnings together, the results suggest that the distributions of gross earnings at most feature small discontinuities and the distributions of net earnings feature a robust and substantial negative discontinuity. This implies that neither employees nor employers can shift their entire burden to the respective other side of the market.

5 Conclusions

This paper refines an idea by Alvaredo and Saez (2007). They develop a cross-sectional estimator of the economic incidence of social security contributions (SSC). The main advantage of the approach is that it does not rely on reform-induced, exogenous changes in SSC parameters over time. This circumvents problems like non-observed hours of work, limited validity of potential control groups and the restriction to short-term responses. We argue, however, that these advantages come at the cost of other rather restrictive assumptions like for example the absence of optimization frictions such that the drop in the marginal SSC rate is accounted for immediately after crossing the threshold. The approach is also reliant on a very high data quality. First, the sample size has to be sufficiently large to reduce random noise in earnings distributions. Second, the amount of earnings subject to SSC has to be measured with very little measurement error.

Conducting simulation exercises we gauge the effect of random measurement error as well as optimization frictions and show that both are able to destroy identification. By contrast, we show that allowing for behavioural responses as well as heterogeneous economic incidence does not significantly distort identification. We then apply the approach to the earnings caps of SSC in Germany. We exploit the specific advantages of the German Structure of Earnings Survey (GSES) for the empirical analysis. The data provides a sufficient sample size for distributional analyses, precise and detailed measures of gross earnings as well as a direct measure of employees' SSC. The latter allows us to independently analyze discontinuities at earnings caps for the gross as well as net earnings distributions. We, therefore, do not rely on a single cross-sectional distribution which strongly increases the robustness of the analysis.

We measure small discontinuities at gross earnings thresholds for the pension and health insurance. Those discontinuities are mostly statistically significant. Since discontinuities go into opposite directions for the pension and health insurance and robustness checks have shown they might just be random noise in the data, we tend not to over-interpret these findings in terms of incidence. We find substantial (negative) and statistically significant discontinuities in 'observed' net earnings. Subject to the discussed assumptions, these results jointly suggest that economic and legal incidence are similar.

References

- Alvaredo, Facundo and Emmanuel Saez**, “The Effects of Payroll Taxes on Earnings: Evidence from Spanish Administrative Data,” Working Paper 2007.
- Benmarker, Helge, Erik Mellander, and Björn Öckert**, “Do regional payroll tax reductions boost employment?,” *Labour Economics*, October 2009, 16 (5), 480 – 489.
- 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.
- 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.
- 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, German Institute for Economic Research (DIW Berlin), mimeo 2014.
- Saez, Emmanuel**, “Do Taxpayers Bunch at Kink Points?,” *American Economic Journal: Economic Policy*, 2010, 2 (3), 180–212.
- Summers, Lawrence H**, “Some Simple Economics of Mandated Benefits,” *American Economic Review*, May 1989, 79 (2), 177–83.