Visibility of taxes and the size of government

Libor Dušek*†

August 7, 2002

Abstract

The paper studies the popular hypothesis that "hiding" taxes from taxpayers, such as by nominally dividing the payroll tax between employees and employers, leads to higher taxes. To test the hypothesis, I analyze payroll taxes and pension spending in a sample of 89 countries and find that countries where employers nominally pay a higher share of the payroll tax tend to have lower taxes and smaller pension programs, contradicting the visibility hypothesis. In an effort to rationalize these counterintuitive results, I derive an interest group model of the size of government where taxpayers have biased information about taxes and their incidence. The model shows that the government may get bigger when taxpayers become better informed if making the tax more visible to one taxpayer group makes it less visible to another group, which I argue is plausible in some real-world applications.

JEL classification: H10, H22, H55

Keywords: public choice, size of government, fiscal illusion, tax incidence, social security, payroll tax

1 Introduction

Positive theories of the size of government (for example, Peltzman (1980), Meltzer and Richard (1981), Becker (1983)) generally assume that taxpayers have perfect information about the taxes they pay, their incidence and their deadweight costs. Yet the opposite idea, that taxpayers are not correctly informed about taxes, has also enjoyed considerable support. Governments are sometimes accused of reducing the "visibility" of taxes so that they could pluck more money from taxpayers who are less prone to oppose taxes that are "hidden" from them. This proposition was well formulated in Buchanan and Wagner (1977) and recently has often been popularized by Milton Friedman: "Which taxes should be abolished? The least visible... You need taxation, but taxation should be visible so that people know what they are paying. Otherwise you give everyone the impression he is getting something for nothing."[1] For example, the European value-added tax is considered to be a less visible tax than the sales tax used by most states in the USA, because the VAT is already included in price while the sales tax is added onto the posted price. Similarly, income tax withholding presumably hides the income

*Graduate student, Department of Economics, University of Chicago. E-mail: ldusek@midway.uchicago.edu.

†I greatly appreciate the comments and suggestions from Petr Barton, Gary Becker, Steven Levitt, Casey Mulligan, and seminar participants at the University of Chicago. I would like to give special thanks to professor Mulligan for providing most of the data used in the paper. The usual disclaimer applies.

[1] Milton Friedman in Prague (1997).

tax from workers, compared to the old way of collecting income tax in a single payment at the end of the year.

The suspicion that many taxes are intentionally "hidden" or "invisible" (these two terms will be used interchangeably) has been widely shared among libertarian thinkers. Their desire to make taxes more visible in order to make them lower has already produced a few specific proposals. For example, the Mackinac Center for Public Policy is promoting "The Right to Know Payroll Form", a modification of the current payroll stub. Currently, half of the payroll tax in the USA (the revenue from which finances Social Security and Medicare) is deducted from the employees' paychecks while the other half is paid by the employers. This presumably hides half of the tax from workers. The proposed new payroll stub would also list also the employer's half of the tax.[2] Dean Stansel (1998) of the CATO institute proposes first to repeal the income tax withholding, and ultimates to replace the income and payroll taxes with a national sales tax. That, according to Stansel, "would give Americans a much clearer picture of the cost of government every time they make a purchase".

Some of these claims have already received empirical scrutiny in the "fiscal illusion" literature[3], which generally relates the size of local or state governments to some measure of voters' misperception about the true costs of government. In the pioneering paper, Wagner (1976) tested a "tax complexity hypothesis", asserting that tax systems that rely on a few large sources of revenue will be opposed more by taxpayers than tax system that rely on numerous smaller sources. He regressed expenditures of 50 large city governments on a number of control variables and a Herfindahl index of tax sources. The Herfindahl index has an obvious intuitive appeal for measuring the degree in which the tax system relies on few large tax sources. Consistent with the hypothesis, Wagner found a large negative effect of increasing tax concentration on city expenditures. Another support for the hypothesis was found by Pommerehne and Schneider (1978). On the contrary, Clotfelter (1976) studied U.S. state-level spending on higher education and did not find any significant relationship between the level of spending and Herfindahl index of tax sources. He also studied the relationship between spending and the states' reliance on more visible taxes (i.e., direct taxes). His founding was contrary to the illusion hypothesis: heavier reliance on more visible taxes was associated with more, not less, spending.

This paper contributes to our understanding of visibility of taxes in two ways. First, in section 2 it provides a new (and, I will argue, more convincing) test of the visibility hypothesis. The existing fiscal illusion literature suffers from one shortcoming: its measures of visibility are in some sense also measures of the efficiency of the tax system. For example, a tax system that spreads the tax burden more equally among many tax sources should generally be more efficient than a tax system relying on one tax source. However, we should expect more efficient tax systems to be associated with bigger governments (Becker and Mulligan (1998)). Therefore, Wagner's and others' findings may reflect the effect of efficiency, rather than visibility, of taxes on the size of government. [4]

To overcome this problem, I use the split of the payroll tax between employers and employees as a measure of visibility, and exploit the international variation in this measure to test the visibility hypothesis. Using the split of payroll tax has several advantages: how the payroll tax is divided between employer and employee has no effect on the efficiency of the payroll tax, and, as the standard tax incidence theory tells us, it has no effect on the division of real tax burden

[2]The new payroll stub has been already adopted voluntarily by a number of employers, including the State of Michigan.

[3]For a recent survey, see Oates (1991)

[4]For more arguments about the problems with empirical fiscal illusion literature, see Oates (1991).

between employer and employee.[5] It is a mere accounting device, which might, however, affect workers' perception about their tax burden. Surprisingly, I find that the effect of the split goes *against* the visibility hypothesis - the payroll tax is lower in countries where employers pay a larger fraction of it - or is insignificant.

Second, the existing literature on visibility has been only empirical, without specifically modelling the process through which taxpayers' misinformation translates into the size of government.[6] In section 3 I attempt to fill this gap and present a model of political competition between subsidy recipients and taxpayer groups who have imperfect information about the tax incidence. The model provides an important link between visibility of taxes and tax incidence, and helps explain the counterintuitive empirical results found in this paper or in Clotfelter (1976). The model predicts that the government may get bigger or smaller as taxes become more visible, depending on the initial bias in taxpayers' information. Section 4 discusses some alternative explanations and section 5 concludes.

2 Empirical investigation

The international differences in the design of payroll taxes offer an opportunity to test the visibility argument empirically. Typically, the payroll tax is paid as a fixed percentage of the contracted wage (gross wage), denoted w_g, and the revenue is used to finance social security and other welfare programs. A tax rate t_A is imposed on the employer while tax rate t_B is imposed on the employee. The total wage costs to the employer is $w^* = w_g(1 + t_A)$ and the take-home wage to the employee is $w_h = w_g(1 - t_B)$. The government revenue is $w_g(t_A + t_B)$. The share of the tax nominally paid by the employer, a^N, is thus

$$a^N = \frac{w_g t_A}{w_g(t_A + t_B)} = \frac{t_A}{t_A + t_B}. \tag{1}$$

I measure the effective tax rate as a share of the tax revenue to the total wage costs:[7]

$$T = \frac{t_A + t_B}{1 + t_A} \tag{2}$$

The employer's share varies substantially among countries, ranging from zero in the Netherlands to one in Pakistan. Most countries are somewhere between the two extremes, but typically the tax is split fifty-fifty or is tilted more heavily on the employer. The question I am trying to answer is: Do countries with higher employer's share have higher tax rates or bigger social security programs? To do so, I regress the fraction of pension spending to GDP in 89 countries against a number of variables, including the employer's share of payroll tax. In addition these regressions, I examine whether the tax rates themselves depend on the employer's share, use

[5] There are caveats to this statement. First, the split of payroll tax matters for real incidence if minimum wage is binding. Second, it has an impact on the income tax bill of the employee.

[6] The literature on the so called "flypaper effect", which is another source of fiscal illusion, does provide explicit modelling. See for example Logan (1986). However, in the flypaper effect illusion, the grants from central to the local governments confuse taxpayers about the marginal tax price of local public goods. This is a substantially different subject from the visibility hypothesis.

[7] The sum of the two tax rates, $\mathsf{t}_A + \mathsf{t}_B$, is not a good measure since the equilibrium gross wage adjusts to the split of the tax.

instrumental variables to correct for possible measurement error in the employer's share, and study whether the employer's share has any effect on tax rates in the future.

2.1 The Data

The main source of data was the publication Social Security Programs Throughout the World, published irregularly by the U.S. Social Security Administration. The publication reports basic features, including statutory tax rates, of government pension programs, sickness and maternity programs, work injury insurance, unemployment insurance, and family allowances. I added observations for years 1985, 1989 and 1999 to the dataset for 1995 assembled by Mulligan and Sala-i-Martin (1999). That dataset already contained the ratio of social security expenditure to GDP (till 1989), real GDP per capita measured in 1985 dollars (till 1989), and a share of the people aged 60+ in the population. Unfortunately, the years for which I have observations on payroll taxes do not match perfectly with the years for which I have observations for other control variables.[8] The final sample contains 89 countries from all continents. 27 countries in the sample were OECD members in 1999.

The employer's share of the tax, a^N was computed from the statutory tax rates. Certain issues in computing the split warrant discussion. First, the payroll tax, while much simpler than income tax, is hardly ever a flat tax rate applied over the entire range of income. Some countries (for example Canada, Congo, Luxembourg or Mexico) exempt the low range of income from the tax and exactly half of the countries in the sample impose a cap beyond which additional wage income is not taxed. Twenty-five countries in the sample (mostly developing countries) exempt some parts of the labor force from the payroll tax. These factors may affect the size of the social security program. I deal with them by including two dummy variables indicating whether the tax is capped and whether there are exemptions. Second, while most countries use a single tax rate, 24 countries (among them Japan, Finland or Seychelles) use progressive rates or rates that vary across industries. When the structure of progressive rates is relatively simple, the publication reports all income brackets and corresponding tax rates. In that case, I used the top tax rate to compute employer's share. When the rate structure is more complex, the publication usually reports the average tax rate, which I used to compute employer's share. When different tax rates apply to blue collar versus white collar workers, or private versus public sector workers, I used the tax rates for blue collar and private sector workers. Third, 45 countries have two separate payroll taxes earmarked to finance the pension program (henceforth referred to as the social security tax) and the sickness and maternity program (henceforth referred to as the medical tax). The remaining countries have the medical program financed either from a joint payroll tax or from other sources. This leads to an upward bias in the measure of T for countries that have a joint payroll tax. Since I am ultimately interested only in the social security tax, I deal with this problem by introducing a dummy variable equal to one if the social security tax is also used to finance the medical program.

2.2 The Results

Table 1 presents the summary statistics. The mean employer's share is 59 percent. Equal split of the tax is most popular among governments, as it us used by 23 countries (in 1999),

[8] One relevant control that is missing in the sample is a measure of income inequality, which has been found an important source of the growth of government in Peltzman (1980).

including the United States. Within a country, employer's share tends to be very stable over time. The correlation between the share in 1985 and 1999 is 0.74, and 36 countries did not change the share at all over those 14 years. On average, employers share has declined slightly (by 1.6 percentage points) although some countries made radical decreases (Peru from 66% to 0, Poland from 100% to 50%) or increases (Hungary from 40% to 73%). The OECD countries do not differ from other countries in terms of the split, although they have higher pension spending and higher fraction of the elderly. Figures 1 and 2 plot the tax rates on employer and employee. The general pattern is that countries that tax the employer more heavily also tax the employees more heavily, although this pattern seems to be weaker for OECD countries. Figures 3 and 4 plot the effective social security tax rate (defined in equation (2)) against employer's share. They do not show any clear relationship - in particular, note the spread of tax rates for countries with equal split.

Tables 2 and 3 present the results of cross sectional regression. The dependent variable is the pension spending as a share of GDP, averaged over 1980-1989. Regressions in Table 2 have the employer's share in 1985 on the right-hand side while regressions in Table 3 have the employer's share averaged over 1985-1999 on the right hand side. Countries that do not finance social security from payroll taxes were excluded. Dummy variables for caps and exemptions were included, but the coefficients were insignificant and are not reported here. Like in Mulligan and Sala-i-Martin (1999), the results cast serious doubt on the conventional visibility argument. In the basic specifications (column (1) of Tables 2 and 3), the coefficients on the employer's share are -0.0246 and -0.0199, respectively, and they are significant at 5 and 10 percent level, respectively. The size of the coefficients is not dramatic. It implies, for example, that if the United States increased the employer's share from current 50% to 60%, the pension spending as share of GDP would eventually drop by 0.2 percentage points, from current 6.25[9] percent to 6.05 percent.

Suspecting that the "determinants" of social security programs may operate differently in developed and developing countries, I reran the regression separately for countries that are currently OECD members and for others. The results are presented in columns (2) and (3) of Tables 2 and 3. The negative effect of employer's share turns out to be much stronger in non-OECD countries (the coefficient falls to -0.0506 or -0.0463 and is strongly significant), while for OECD countries it remains negative but loses statistical significance. The coefficients on control variables are also worth some discussion. The fraction of the elderly in the population remains the single most important factor affecting the size of social security programs. The coefficient varies between 0.5 and 0.7 and indicates that increases in the elderly population translate into less than one-for-one increases in the social security programs. However, the size of the program is much more sensitive to the fraction of elderly population in the OECD than in non-OECD countries. Economic development appears to have a negligible effect on the social security program, once the correlation between development and the fraction of the elderly is accounted for. Still it is worth noting that among non-OECD countries, the richer ones tend to have a bigger social security program, while the opposite seems to be the case among the OECD countries.

For reasons explained later in section 3, the relationship between employer's share and the size of the program may by U-shaped rather than linear. I explore this possibility by adding a square of the employer's share into the regression. The results are reported in columns (4) through (6) of Tables 2 and 3. Indeed, the coefficients on both the linear and the quadratic term are now strongly significant in all but one specification. The coefficient on the linear term is always negative and the coefficient on the quadratic term is always positive. The adjusted R^2 improves somewhat and the coefficients on other variables are virtually unchanged. The

[9]The last observation for the United States, for year 1997, in the IMF government finance statistics.

magnitude of coefficients implies that the size of the program is decreasing at low values of employer's share but increasing at high values. For example, column (4) of Table 3 implies that the change from 50% split to 60% split would reduce the social security program in the U.S. by 0.26 percentage points, which is similar in magnitude to the prediction of the linear specification. The size of the program is minimized when the share is equal to 0.62, which is surprisingly close to the actual average.

As mentioned earlier, the structure of payroll taxes often includes caps, exemptions, progressive rates, or rates differing between sectors. These complications could have induced a measurement error in the employer's share. I try to correct for the error by an instrumental variable estimator, using the split of the medical tax as an instrument for the split of the social security tax. The split of the medical tax is correlated with the split of the social security tax (see Table 1); the correlation across all observations is 0.2917, being lowest in 1985 (0.173) and highest in 1995 (0.432). For the United States, the two splits are identical. I would argue that while the split of the medical tax might influence the medical tax rate or the size of publicly financed health care, it should not influence the social security tax rate or the size of publicly financed pensions. Therefore the split of the medical tax could be a valid instrument. The disadvantage of the instrumental variable estimator is that many countries do not have a separate medical tax, and therefore half of the sample could not be used.

The results are presented in Table 4. When the employer's share in 1985 is used as the explaining variable, the coefficient on employer's share becomes positive but insignificant. When the quadratic term is included, both coefficients are slightly positive but also insignificant. When the average employer's share over 1985-1999 is used as the explaining variable, the coefficients have the same sign as in the linear regressions, although they are smaller in absolute values and insignificant.[10]

Pension spending is one way of measuring the size of the social security program. However, part of the spending may be financed by sources other than the social security tax, or part of the tax revenues may be used to financed other spending. Ultimately we are interested in whether countries where employers pay a higher share of the payroll tax have higher *taxes*. Therefore I use social security tax rates, computed as in equation (2) and averaged over the years 1985, 1989, 1995, and 1999, as the dependent variable. The results are reported in Table 5 for both OLS and IV regressions. In the linear OLS specification (columns (1)-(3)), the coefficients on employer's share are negative but they are not significant. In the quadratic OLS specification (columns (4)-(6)) the coefficients on share are significant only for the OECD subsample and again they imply that the tax rate is falling for employer's share below 0.65 and rising thereafter. IV estimation produced coefficients with the same signs as in Table 4, but they have large standard errors. The fit of the models is fairly poor, however, particularly for the non-OECD subsample in OLS. The adjusted R^2 does not exceed 0.30, except for the OLS regressions on the OECD subsample.

Two additional interesting results emerge from Table 5. One is the very large coefficient on the fraction of the elderly in the OECD subsample (above 1.5). Since one would expect that the fraction of the elderly translates at most one-for-one into pension spending, is there something wrong with the results? Note, however, that the ratio of pension spending to the GDP (let's for now assume it is equal to the ratio of payroll tax revenue to the GDP) and the payroll tax rate are not the same thing. Let R be the average payroll tax revenue, Y the gross domestic product and α_{LT} the share of taxable wage income in GDP. By accounting identity, the average

[10] I also produced separate estimates for OECD and non-OECD countries, but due to very small sample, in some cases the IV estimation produced implausible estimates with huge standard errors. Therefore they are not reported here.

tax rate is

$$T = \frac{R}{\alpha_{LT} Y} = \frac{1}{\alpha_{LT}} (\frac{R}{Y})$$

Therefore the coefficients in regressions in Table 5 need to be scaled by α_{LT} if they are to be compared with coefficients in regressions in Tables 2-4. I do not have information on α_{LT}, but its upper bound must be the ratio of labor income to GDP (if all labor income is taxed), which is around 0.67. Caps on payroll taxes and exemptions make the true α_{LT} smaller. Multiplying the coefficient on the fraction of the elderly by the share of labor income gives $1.5180 \times .67 = 1.02$, and so virtually any cap or exemption would reduce α_{LT} sufficiently below 0.67 so that the scaled coefficient would fall below 1, into the "reasonable" range. The value of α_{LT} that would equalize the scaled coefficient in Table 5 with the coefficient in Table 3 is $\alpha_{LT} = 0.48$. This in turn would imply that about 72 percent ($= 0.48/0.67$) of labor income is, on average, subject to payroll taxes. I do not know how the implied α_{LT} compares with the actual. If it is too low, that would in turn imply that as the faction of the elderly increases and there is more "demand" for pension spending, the role of payroll taxes as a source of pension financing increases.

The other interesting finding is the coefficient on the dummy variable indicating whether the medical program is financed from a joint payroll tax as opposed to separate medical and social security taxes. If there are no unobservable differences between countries that use a joint tax and those that use a separate medical tax, the coefficient should be equal to the average medical tax rate in the countries that have a separate medical tax. In other words, the sum of social security and medical tax rates should equal the joint payroll tax rates in the countries that do not use separate medical and social security taxes. However, the results indicate that this is not the case. The coefficient on the "joint" dummy is 0.0565 for the whole sample and 0.0284 for the OECD countries, while the average medical tax rate in the countries that have it is 6.22 percent (whole sample), and 8.16 percent (OECD). It appears that the countries that choose to finance the health program by a separate medical tax have higher medical tax rates than countries that "hide" the medical tax into a joint tax with social security.

Finally I look at the behavior of tax rates over time. Table 6 shows the regression of social security tax rate in 1995 on the employer's share and the social security tax rate in 1985. That is, given the tax rates that countries had in 1985, did countries with higher employer's share had higher tax rates ten years later? The answer is negative. Indeed, countries with higher employer's share seem to have reduced the tax rates relative to countries with low share, although the relationship is weak in all specifications except column (1). Not surprisingly, the tax rate in 1985 best "explains" the tax rate ten years later, notably for the non-OECD countries. The coefficient on the tax rate in 1985 is 0.94 for non-OECD countries, which suggest a very high degree of persistence. Indeed, 24 non-OECD countries in the sample have not changed their social security tax rates at all during these 10 years, which can be said only of 5 OECD countries (among which Australia and New Zealand have had zero rates). Tax rates in OECD countries appear more sensitive to the underlying factors, notably the fraction of the elderly in the population.

The results from the panel data (fixed effects) regression are shown in Table 7. Unfortunately, my data for pension spending end in 1989, so I did only the regression for tax rates as the dependent variable. I also did not have data for GDP after 1989, but I added the ratio of total government revenue to GDP as an alternative control variable. In the linear specification, all coefficients on employer's share are negative and they are significant in the pooled regression and in OECD subsample. When significant, the coefficients are higher in absolute values than in the cross section regressions (Table 5). They would imply that an increase in employer's

share by 10 percentage points would reduce the tax rate by 0.797 percentage points in OECD countries. Given that the average tax rate in the OECD countries is 14%, this seems quite high. In the quadratic specification, the results are mixed. Coefficients are significant only in the non-OECD subsample, with the linear term being negative and the quadratic term positive, as in the cross section regressions. The panel analysis yields a very low coefficient on the fraction of the elderly in the OECD countries (0.17 and 0.11, respectively), but the influence of the elderly appears to be captured in the coefficient on the overall size of government. It should be noted that since the employer's share remains relatively stable over time, the results are driven mainly by variation between countries rather than within countries over time. The fraction of variance due to between variation ranges from 0.62 to 0.92. Therefore, the panel analysis does not add much new information than was already conveyed in the cross section. And the R^2 is very low in the non-OECD subsample.

To summarize, the results disqualify the claims that hiding the payroll tax from employees leads to higher taxes or higher social security spending. *Not a single regression* shows a positive and significant relationship between employer's share of the social security tax and the tax rate or pension spending. Out of 18 linear specifications, the coefficient on employer's share is negative and significant 7 times, negative and insignificant 10 times, and positive but insignificant only once. Out of 18 quadratic specifications, coefficients on employer's share are significant 6 times. In these cases one can detect a positive relationship between employer's share and the taxes, but only in the upper range of the share (above 60 or 66 percent). The results suggest that either the taxpayers are not being fooled, or even if they are being fooled, it has the opposite effect.

3 Theoretical model

The theoretical section of the paper develops a model of the size of government in which taxpayers are wrongly informed about the taxes, and rationalizes why taxes may indeed be lower if they are "invisible". But before that I should briefly mention whether taxpayers are indeed wrongly informed and why we might expect them to be.

For one thing, economists themselves do not know a lot about taxes and their incidence, and often have different views (for a controversy on the incidence of corporation income tax, compare Krzyzaniak and Musgrave (1963), Cragg, Harberger and Mieszkowski (1967) and Gordon (1967)). But probably the best justification for why taxpayers may wrongly perceive taxes is offered in Buchanan and Wagner (1977, p.127-134) and is based on the rational ignorance argument: It is an intellectually difficult task to learn about all complicated taxes and their consequences, yet each individual taxpayer has very little influence to change them. Therefore it is rational for him not to invest in knowledge about taxes.

There have been surveys on people's awareness of taxes. Goetz (1977) provides a summary of some these studies. One study found that "only slightly more than one-half of the respondents were able to estimate their true income tax payments within a plus or minus 10 percent range". Another study concluded that "about 30 percent of taxpayers in the above-average $10,000+ category were found to be under misapprehension concerning their marginal tax rates."[11] Bises (1990) surveyed Italian workers about the income tax and found that "only 8% of the respondents appear to know their average tax rate with sufficient approximation (± 2 percentage

[11] The quotes are from Goetz (1977).

points), while 57% overestimate it. ... The stated average rate is about one and a half times higher than the true one."

3.1 Visibility of taxes and tax incidence

Given that taxpayers are not perfectly informed about taxes and some taxes are hidden, what is the relationship between visibility and the size of government? The conventional visibility argument seems indisputable. If a tax is hidden from a taxpayer, how could this not make it easier for the government to tax him more? But the argument tacitly assumes that there is only one taxpayer (or taxpayer group) and hence ignores one of the central issues in public finance, tax incidence. If tax incidence is brought into the analysis, the visibility argument becomes less straightforward. If there are two taxpayers who bear the tax burden, the tax has to be made less visible to both taxpayers, or, at minimum, less visible to one taxpayer without being made more visible to the other taxpayer if the government is to be made bigger. This may be very hard in practice. More likely, making the tax less visible to one taxpayer makes it more visible to the other taxpayer. Below I present two arguments to support this claim, based on nominal tax incidence and information about the real incidence.

Presumably, nominal tax incidence is tied to visibility. Taxpayers are likely to be more aware of the tax that they nominally pay than of the tax that someone else nominally pays. Therefore, hiding the tax from one taxpayer by having some other taxpayer write a check to the government could make the tax more visible to the other taxpayer. The split of the payroll tax (recall that the average employer's share is 57-60%) is a good example. While almost 60% of the tax may be hidden from workers, it may be more visible to employers. They need not realize that wages would rise if the tax were repealed, and therefore they may think that they actually bear their share of the tax (or at least a part of it). They will have a reason to oppose the tax, thus replacing some, or perhaps all, of the reduction in opposition by the workers.

Firms indeed do oppose their share of the payroll tax, while workers would like to shift the nominal burden on employers. An important recent example is Germany's debated pension reform, driven in part by the desire to cut German firms' labor cost and thus increase their competitiveness.[12]. However, recent empirical research on the incidence of the payroll tax[13] seems to be coming to a consensus that the tax is fully shifted into wages. Hence the tax reform should have no effect on the labor costs of German firms. Yet if the tax were nominally paid only by workers, the employers probably would not fight for tax cuts.

Alternatively, consider that taxpayers do know that nominal incidence does not matter for who ultimately bears the tax burden, but they are not perfectly informed about the real incidence. If their knowledge improves over time, some taxpayers learn that they in fact pay lower taxes than they originally though, but other taxpayers inevitably learn that they pay higher taxes. The idea that making tax less visible to one taxpayer inevitably makes the tax more visible to another taxpayer is the crucial assumption of the model.

[12] Financial Times, September 27, 2000.
[13] Gruber (1995), Anderson and Meyer (1998).

3.2 The Model

In order to formalize the misperception, I extend the framework provided by Becker and Mulligan (1998). The model is general and applies to any tax, not just payroll tax. The size of government is determined by competition between pressure groups, taxpayers and subsidy recipients. To build the tax incidence into the framework, I assume that there are two taxpayer groups, A (for example, consumers) and B (producers), and one group of subsidy recipients, C. The government collects a tax revenue T from a single tax, the real incidence of which may fall on both taxpayer groups. The taxpayer group A bears a tax burden of T_A and taxpayer group B bears a tax burden of T_B. Of course, the real tax incidence does not depend on the on the nominal tax incidence, i.e., which group actually writes the check to the government. A simple accounting identity requires $T_A + T_B = T$.

An important assumption is that the political process cannot shift the real tax burden between taxpayers. That is, the real incidence is determined only by the (exogenous) taxing technology available to the government and the underlying economy. The government may increase the tax and thus increase the tax burden falling on both taxpayers, but cannot alter the taxing technology such that it would shift the real tax burden from taxpayer A to taxpayer B and vice versa. For analytical simplicity, I assume that, regardless of T, the groups bear a constant proportion a^R and $b^R = 1 - a^R$, respectively, of the tax: $T_A = a^R T$, $T_B = (1 - a^R)T$.[14] I take the natural assumption that $a^R \in [0, 1]$, although some studies of tax incidence in imperfectly competitive markets[15] find that there is overshifting of the tax into the price, so the consumers are bearing more than the tax and $a^R > 1$.

As in the Becker-Mulligan model, what ultimately limits the size of government is the deadweight costs of taxes $\Delta(T)$ and the deadweight costs of subsidies $\Sigma(G)$. Marginal deadweight costs are assumed to be positive and increasing, although for some values of T and G the total deadweight costs may be negative if the tax or subsidy induces an efficiency-improving change

[14] The ratios a^R and b^R are derived as follows for a partial equilibrium model of an excise tax t imposed on each unit of the good sold: The quantity supplied and demanded as a function of price are $S(p)$ and $D(p)$, respectively. Let p^*be the market equilibrium price without the tax, p_t be the after-tax price received by producers, and so $p_t + t$ is the after-tax price paid by consumers. p_t, and the after-tax quantity q_t, are given by

$$S(p_t) = D(p_t + t) = q_t$$

The tax revenue is

$$T = tS(p_t) = tD(p_t + t)$$

of which the consumers bear

$$T_A = a^R T = \frac{(p_t + t - p^*)}{t} T$$

and the producers bear

$$T_A = b^R T = (1 - a^R)T = \frac{(p^* - p_t)}{t} T.$$

In this general formulation, a^R need not be constant as T changes, although it is constant, for example, when the supply and demand curves are linear. The assumption that a^R remains constant is taken for analytical simplicity. In that case, we can conveniently define the after-tax prices as explicit functions of p^*, a^R, and t:

$$\begin{aligned} p_t &= p^* - (1 - a^R)t \\ p_t + t &= p^* + a^R t \end{aligned}$$

[15] Besley and Rosen (1998).

in behavior. Like the tax, the deadweight cost may fall on both taxpayers: $\Delta_A + \Delta_B = \Delta(T)$.[16]

Tax revenue T is determined by the political pressures produced by each group. Let A, B, and C denote the amount of resources (time or money) devoted by each group to political pressure. The groups are assumed to allocate their resources efficiently between various pressure-producing activities, such as campaign contributions, bribes, advertising, lobbying etc. These specifics of political pressure, as well as the mechanics of the political system through which the pressures are translated into outcomes, are subsumed into the pressure function, $F(A, B, C)$, the value of which denotes the amount of tax taken from taxpayers in the political equilibrium. Since the government can only give what it takes, we have the balanced budget condition:

$$T = F(A, B, C) = G$$

More pressure by any of the taxpayers reduces the size of government, and more pressure by the subsidy recipients increases the size of government. Both are subject to diminishing returns:

$$F_A < 0, F_B < 0, F_C > 0, F_{AA} > 0, F_{BB} > 0, F_{CC} < 0$$

In order to isolate the effects of tax incidence from differences in political influence, I assume that the pressure function is symmetric in A and B, so that both taxpayers are equally politically powerful:

$$F(A, B, C) = F(B, A, C)$$

To formalize the taxpayers' misperception about the tax or its incidence, I introduce parameters $a^\phi \in [0, 1]$ and $b^\phi \in [0, 1]$ to represent the *perceived* tax incidence, as opposed to the real tax incidence a^R, b^R. They say what share of the total tax each taxpayer thinks he is paying: Taxpayer A thinks he is paying $a^\phi T$ and taxpayer B thinks he is paying $b^\phi T$. Such a way of modelling misperception can be rationalized in several ways. First, the taxpayers are perfectly aware that the government is collecting T in taxes, but they have a wrong information about its incidence. Alternatively, the taxpayers are aware only of a part of the tax that the government is collecting, $a^\phi T$ or $b^\phi T$ (the tax is "invisible"), but they believe that they fully bear the tax that they are aware of. Intermediate cases are also possible: For example, while the government is collecting T in taxes, consumers may be aware only of 80% of it, and further think that 75%

[16] For the sales tax, Δ_A and Δ_B are derived as follows. Total deadweight costs of the tax are

$$\Delta = \int_{p_t}^{p^*} \mathsf{S}(\mathsf{p})\mathsf{dp} + \int_{p^*}^{p_t+t} \mathsf{D}(\mathsf{p})\mathsf{dp} - \mathsf{T}$$

of which the consumers and producers bear

$$\begin{aligned}
\Delta_A &= \int_{p^*}^{p_t+t} \mathsf{D}(\mathsf{p})\mathsf{dp} - \mathsf{a}^R\mathsf{T} = \int_{p^*}^{p_t+t} \mathsf{D}(\mathsf{p})\mathsf{dp} - (\mathsf{p}_t + \mathsf{t} - \mathsf{p}^*)\mathsf{D}(\mathsf{p}_t + \mathsf{t}) \\
\Delta_B &= \int_{p_t}^{p^*} \mathsf{S}(\mathsf{p})\mathsf{dp} - (1 - \mathsf{a}^R)\mathsf{T} = \int_{p_t}^{p^*} \mathsf{S}(\mathsf{p})\mathsf{dp} - (\mathsf{p}^* - \mathsf{p}_t)\mathsf{D}(\mathsf{p}_t + \mathsf{t})
\end{aligned}$$

The taxpayers' shares of total deadweight costs need not be the same as their shares of the tax revenue. They would be the same if the following condition holds:

$$\frac{\Delta_A}{\Delta} = \frac{\mathsf{T}_A}{\mathsf{T}} \Leftrightarrow \frac{\int_{p^*}^{p_t+t} \mathsf{D}(\mathsf{p})\mathsf{dp}}{\int_{p_t}^{p^*} \mathsf{S}(\mathsf{p})\mathsf{dp}} = \frac{\mathsf{a}^R}{1 - \mathsf{a}^R}$$

of the tax is borne by them. Therefore $a^\phi = 0.6$. What ultimately matters for the behavior is how much taxpayers think they are paying, $a^\phi T$ or $b^\phi T$.[17]

The parameters a^ϕ and b^ϕ summarize all factors that influence taxpayers' awareness of taxes and their incidence: the nominal tax incidence, the knowledge that economists have about the real incidence (which is very imperfect so far), the degree in which such knowledge has been communicated to the taxpayers, or the method of tax collection. The perceived tax incidence is assumed to be exogenous to the model, although in reality it may be influenced by the political competition itself. For example, if taxpayer group A underperceives its tax incidence, group B may direct some of its political pressure toward educating members of group A about the real tax incidence in order to mobilize them to participate more actively in the political process. Unfortunately, we do not have a theory of how taxpayers acquire knowledge about taxes and their incidence and how they form perceptions. While such an extension would be undoubtedly interesting to study, I will restrict the scope of this paper to exploring how *given* misperceptions affect the size of government. [18]

One need not impose any relationship between a^ϕ and b^ϕ, but that would come at a cost of having too much flexibility: Any small government could be explained by asserting that both taxpayers think they bear the entire share (a^ϕ and b^ϕ are close to 1), and any large government could be explained by asserting that the tax is "hidden" from both taxpayers (a^ϕ and b^ϕ are close to 0). The interesting question is what happens if the perception (or "visibility") of the tax shifts from one taxpayer to another. Therefore I impose the requirement that $db^\phi/da^\phi < 0$, without fixing the absolute values.[19] As with the real tax incidence I assume that a^ϕ and b^ϕ remain constant as T changes.

The perception of the tax incidence has implications for the taxpayers' perception of the incidence of deadweight cost. The main result, analyzed below, is that if taxpayer's perceived incidence is higher than the real incidence, not only is his perceived deadweight cost greater than his real deadweight cost, but it may be actually greater than the *total* deadweight cost borne by both taxpayers.

Figure 5, a textbook depiction of a sales tax, demonstrates the link between the perception of the tax and the perception of deadweight cost. If $a^\phi = a^R$ and $b^\phi = b^R$, both consumers and producers correctly estimate the tax incidence, and thus consumers correctly think that they are paying a tax T_A and suffering a deadweight cost Δ_A, and producers think that they are paying a tax T_B and suffering a deadweight cost Δ_B. They know that if the tax were repealed, producers would receive price p^* instead of p_t, consumers would pay price p^* instead of $p_t + t$, and a quantity q^* instead of q_t would be produced and consumed.

How do consumers perceive deadweight costs if they misperceive the tax incidence? At one extreme, $a^\phi = 0$ implies that consumers think the tax is not reflected in price (either because

[17] Note that even if the taxpayer misperception is rationalized by "invisibility" of the tax, the perceived tax incidence may still exceed the real tax incidence. In the example above, if $a^R < 0.6$, the taxpayer thinks he pays more than he actually does, because the invisibility is outweighed by a pesimistic belief about incidence.

[18] The analysis presented here is not limited to the direct redistribution through taxation, but can also be applied to redistribution by regulation. For example, cross-subsidies mandated by regulators are common in the network industries such as telecommunications. Rural cosumers pay line charges that are below costs, and phone operators recover the loss from above-costs charges to urban consumers. However, the "revenue" from the taxed consumers can be raised in ways which are presumably more or less visible to them. The tax may be hidden in a higher price, which is presumably the least visible form. Many consumers need not even know that they are paying some tax. Alternatively, the tax can be shown as a special item on the telephone bill, in which case the consumer has an immediate information about the tax he is paying.

[19] A useful special case is $a^\phi + b^\phi = 1$. It would describe a situation when both the firms and the workers naively think that they bear their nominal share of the payroll tax. In that case $db^\phi/da^\phi = -1$.

they think they do not bear the tax or they do not know about it). Therefore they expect no reduction in the after-tax price if the tax were repealed, and no change in their behavior. They expect they would continue buying quantity q_t at a price $p_t + t$. Hence their perceived deadweight cost, Δ_A^ϕ, is also zero.

At the other extreme, $a^\phi = 1$, and consumers are fully aware of the tax plus they think the tax is shifted dollar for dollar into the price. If the tax were repealed, they think the price that they pay would fall to p_t, and they would be buying quantity $D(p_t)$. Consumer thus perceive a deadweight cost Δ_A^ϕ equal to the area $(\Delta_A + \Delta_B + \Delta_C + \Delta_D)$ in Figure 2, which is not only larger than the real deadweight cost borne by them, Δ_A, but also larger than the total real deadweight cost borne by both taxpayers, $\Delta_A + \Delta_B$.[20] This is because the overperception of tax incidence implies an overestimate of the change in consumers' behavior in response to the tax.

More generally, the deadweight costs perceived by taxpayers A and B are expressed as functions $\Delta_A^\phi(a^\phi T)$ and $\Delta_B^\phi(b^\phi T)$, respectively.[21] The assumption that the perceived deadweight costs are functions only of the perceived tax, regardless of its decomposition into total tax and the perception, is taken for analytical convenience.[22] It follows from the previous discussion that

$$\begin{aligned} \Delta_A^\phi(0) &= 0 \\ \Delta_A^\phi(a^R T) &= \Delta_A(T) \\ \Delta_A^\phi(T) &\geq \Delta(T), \text{ with strict equality iff } a^R = 1 \end{aligned}$$

and analogously for B. The perceived deadweight costs functions need not be symmetric for A and B, as Figure 6 demonstrates. Because the supply curve is more inelastic beyond the equilibrium point, the producers' perceived deadweight cost when $b^\phi = 1$ (the pentagon ABCEF) is smaller than the consumers' perceived deadweight cost when $a^\phi = 1$ (triangle BCD). The perceived deadweight costs are convex and I further assume they are continuous.

[20] For example, when supply and demand curves are linear with the same slope, $\Delta_A^\phi = 4\Delta_A = 2(\Delta_A + \Delta_B)$, i.e., the consumers' perceived deadweight cost is four times as high as consumers' real deadweight cost and two times as high as the total deadweight cost.

[21] For the sales tax, the perceived deadweight cost function is derived as follows for consumers: Let $\mathsf{p}_A^\phi(\mathsf{a}^\phi)$ denote the perceived tax-free price, i.e., the price that consumers' believe would be the market equilibrium price if the tax were repealed. It is defined by

$$\mathsf{p}_A^\phi(\mathsf{a}^\phi) = \mathsf{p}_t + (1 - \mathsf{a}^\phi)\mathsf{t}$$

In other words, perceived tax incidence a^ϕ implies implies that consumers think the price they pay would drop by $\mathsf{a}^\phi\mathsf{t}$, from $\mathsf{p}_t + \mathsf{t}$ to $\mathsf{p}_t + (1 - \mathsf{a}^\phi)\mathsf{t}$. They think they would buy $\mathsf{D}(\mathsf{p}_A^\phi(\mathsf{a}^\phi))$ units of the good instead of $\mathsf{D}(\mathsf{p}_t + \mathsf{t})$, and therefore they think the tax reduces consumer surplus by

$$\int_{p_A^\phi(a^\phi)}^{p_t + t} \mathsf{D}(\mathsf{p})\mathsf{dp}$$

Since $\mathsf{a}^\phi\mathsf{T}$ of the lost consumer surples is the perceived tax, the perceived deadweight costs is

$$\Delta_A^\phi = \int_{p_A^\phi(a^\phi)}^{p_t + t} \mathsf{D}(\mathsf{p})\mathsf{dp} - \mathsf{a}^\phi\mathsf{T}$$

The perceived deadweight costs for producers is derived analogously. It is straightforward to show that for the sales tax, the marginal perceived deadweight costs is always positive and increasing with respect to both t and a^ϕ.

[22] It is satisfied, for example, if the demand and supply functions are linear. A general formulation would be $\Delta_A^\phi(\mathsf{a}^\phi, \mathsf{T})$, $\Delta_B^\phi(\mathsf{b}^\phi, \mathsf{T})$.

It is interesting to note that if $a^\phi + b^\phi = 1$, the sum of perceived deadweight costs $\Delta_A^\phi + \Delta_B^\phi$ is *minimized* with respect to a^ϕ when $a^\phi = a^R$, i.e., when taxpayers perceive the tax incidence correctly.

In the political "game", each taxpayer group is minimizing its perceived costs of redistribution, which include the perceived tax, perceived deadweight costs, and the expenditure on political pressure. The subsidy recipient group is maximizing the subsidy net of the deadweight cost and the expenditure on political pressure. Each group is solving the problem:

$$\begin{aligned} A &: \min_A \ \Pi^A = a^\phi T + \Delta_A^\phi(a^\phi T) + A \\ B &: \min_B \ \Pi^B = b^\phi T + \Delta_B^\phi(b^\phi T) + B \\ C &: \max_C \ \Pi^C = G - \Sigma(G) - C \end{aligned} \tag{3}$$

In the political equilibrium (the existence of which is assumed), each group is spending an optimal amount on pressure, given the pressures by other groups. The following first-order conditions are satisfied:

$$\begin{aligned} a^\phi F_A(1 + \Delta_A^{\phi\prime}) &= -1 \\ b^\phi F_B(1 + \Delta_B^{\phi\prime}) &= -1 \\ F_C(1 - \Sigma') &= 1 \end{aligned} \tag{4}$$

The taxpayers are exerting political pressure up to the point where a dollar of expenditure yields a dollar reduction in the perceived tax and the perceived deadweight costs.

3.3 Changes in perceptions: why ignorance may be bliss

The main question is what happens to the size of government when the perceptions change, notably, when a^ϕ rises and b^ϕ declines. The direct effect induces group A to produce more pressure (since the tax and its deadweight costs "feel" more) and group B to produce less pressure (since the tax and its deadweight costs "feel" less). In response to that, C will counteract by producing more or less pressure, and A and B will also adjust their pressure in response to change in C's pressure and also to the change in each other's pressure. In the new equilibrium, each group is exerting optimal pressure given the pressures by other groups. The interactions between groups are quite complex and, unfortunately, the comparative statics on the first-order conditions, while technically solvable, do not yield much insight into the equilibrium responses.[23]

[23] The equilibrium responses can be computed by totally differentiating the first-order conditions:

$$\begin{aligned} \Pi^A_{AA}\mathsf{dA} + \Pi^A_{AB}\mathsf{dB} + \Pi^A_{AC}\mathsf{dC} + \Pi^A_{Aa^\phi}\mathsf{da}^\phi &= 0 \\ \Pi^B_{BA}\mathsf{dA} + \Pi^B_{BB}\mathsf{dB} + \Pi^B_{BC}\mathsf{dC} + \Pi^B_{Bb^\phi}(\mathsf{db}^\phi/\mathsf{da}^\phi)\mathsf{da}^\phi &= 0 \\ \Pi^C_{CA}\mathsf{dA} + \Pi^C_{CB}\mathsf{dB} + \Pi^C_{CC}\mathsf{dC} &= 0 \end{aligned}$$

Let Π denote the matrix of the cross derivatives with respect to $\mathsf{A}, \mathsf{B}, \mathsf{C}$. Then we get the following vector of reaction functions:

$$\begin{pmatrix} \mathsf{dA}/\mathsf{da}^\phi \\ \mathsf{dB}/\mathsf{da}^\phi \\ \mathsf{dC}/\mathsf{da}^\phi \end{pmatrix} = -\Pi^{-1} \begin{pmatrix} \Pi^A_{Aa} \\ \Pi^B_{Bb^\phi}(\mathsf{db}^\phi/\mathsf{da}^\phi) \\ 0 \end{pmatrix}$$

Although it comes at a cost of losing generality, more structure needs to be imposed on the pressure function. I therefore assume that $F(A,B,C) = F(A+B,C)$; i.e., the pressures by groups A and B are perfect substitutes, and only the total pressure of taxpayers matters for the size of government. It can be shown that only one taxpayer will be active. Taxpayer A will be the active taxpayer if $a^{\phi}(1+\Delta_A^{\phi\prime}) > b^{\phi}(1+\Delta_B^{\phi\prime})$, i.e., if the tax burden is "felt" more by taxpayer A on the margin. Large changes in perceptions may result in discontinuous switches in pressure between groups. Through the rest of the paper I will assume that the model parameters are such that taxpayer A is the active taxpayer. Analytically, the model now becomes a slight modification of Becker and Mulligan's (1998) model of two taxpayers.

An exogenous change in one active group's pressure induces a change in the opposing group's pressure according to

$$A_C = \frac{a^{\phi 2}(-F_A)F_C\Delta_A^{\phi\prime\prime} - F_{AC}/(-F_A)}{a^{\phi 2}F_A^2\Delta_A^{\phi\prime\prime} + F_{AA}/(-F_A)} \text{ and } C_A = \frac{-F_AF_C\Sigma'' + F_{CA}/F_C}{F_C^2\Sigma'' - F_{CC}/F_C}$$

A_C and C_A denote the slopes of the reaction functions of groups A and C. The changes in equilibrium pressures by both groups are obtained by comparative statics:

$$\begin{aligned} \frac{dA}{da^{\phi}} &= \frac{-\Pi^A_{Aa^{\phi}}}{\Pi^A_{AA}(1-A_CC_A)} \\ \frac{dC}{da^{\phi}} &= C_A\frac{dA}{da^{\phi}} \end{aligned} \tag{5}$$

where $-\Pi^A_{Aa^{\phi}} = (-F_A)(1+\Delta_A^{\phi\prime\prime}) - a^{\phi}FF_A\Delta_A^{\phi\prime\prime} > 0$ is the change in A's perceived marginal income due to a change in the perceived tax incidence, and Π^A_{AA} is the second derivative of A's objective function, which is negative by the second order condition. The change in the size of government is determined by the change in the pressure by both groups:

$$\frac{dT}{da^{\phi}} = F_A\frac{dA}{da^{\phi}} + F_C\frac{dC}{da^{\phi}} = (F_A + F_CC')\frac{dA}{da^{\phi}} \tag{6}$$

As in Becker and Mulligan (1998), I want to rule out implausible reaction functions where an increase in C's pressure in reaction to an increase in A's pressure would be so large that the government would actually get bigger, even though A was exogenously given advantage in the political process due to an increase in his perceived tax incidence. The government gets smaller as a^{ϕ} increases if $(F_A + F_CC') < 0$, which (as can be verified) holds if

$$\frac{F_AF_{CC}}{F_C} > F_{CA}$$

This is the strategic separability condition introduced by Becker and Mulligan (1998) and it imposes an upper bound on the cross derivative of the pressure function so that C is not induced to increase its pressure "too much". Intuitively, it should not be optimal for the subsidy

This general result is not very transparent because it involves an inverse of a 3x3 matrix. The usual way out, i.e., assuming that some elements of the Π matrix are zero, is not applicable here. For example, $\Pi^A_{AB} = -\mathsf{a}^{\phi}\mathsf{F}_{AB}(1+\Delta_A^{\phi\prime}) - \mathsf{a}^{\phi}\mathsf{F}_A\mathsf{F}_B\Delta_A^{\phi\prime\prime}$. Hence even if $\mathsf{F}_{AB} = 0$ (the marginal "political product" of group A is not affected by B's pressure), it is not true that $\Pi^A_{AB} = 0$ because of the term $\Delta_A^{\phi\prime\prime}$, which is always positive. A's marginal income is still affected by B's pressure through the effect on perceived marginal deadweight costs.

recipient to "override" an exogenous increase in the taxpayer's additional pressure. We can see that the condition does not depend on taxpayers' perceptions.

The result that the government gets smaller if the perceived tax incidence of the active taxpayer increases is not surprising as such. A more interesting question is what happens to the size of government if the taxes become more visible in the sense that the perceived tax incidence gets closer to the true incidence. As the analysis below demonstrates, the government need not shrink when taxpayers become better informed about taxes and their incidence.

Let me study the case when $a^R > b^R$ and also $a^R(1+\Delta'_A) > b^R(1+\Delta'_B)$, i.e., when the tax falls mainly on taxpayer A and the deadweight cost functions Δ_A and Δ_B are such that A would be the active taxpayer under perfect information.[24] The results depends on the relation between perceived and real incidence, and four different scenarios may occur:

1. $a^R > a^\phi > b^\phi > b^R$. The perception was biased toward the average, but it remains in the right order, so that the taxpayer A is active. More visibility implies that a^ϕ goes up, the active taxpayer will produce more pressure and the government will get smaller.

2. $a^\phi > a^R > b^R > b^\phi$. The perception is biased toward the extreme, i.e., taxpayer A thinks he bears a larger tax burden than he actually does. When provided with better information, he realizes he was spending too much on political pressure, therefore he will reduce pressure and the government will get bigger. The tax is perceived more by taxpayer B, but he remains inactive.

3. $a^R > b^\phi > a^\phi > b^R$. The perception is biased toward the average, but it is biased so much that the taxpayers switch roles: Taxpayer B is active although he should not be. Better information induces B to cut his pressure while A remains inactive, and so the government gets bigger.

4. $b^\phi > a^R > b^R > a^\phi$. The perception is completely reversed from the truth so that taxpayer B thinks he is paying most of the tax while in fact he is paying very little. When provided with better information, B will reduce his pressure and the government will get bigger.

As we can see, the government may get smaller as well as bigger, depending whether the perceived tax incidence of the *active* taxpayer goes up or down. Scenarios 2, 3, and 4 demonstrate the "ignorance may be bliss" result: the government is actually smaller when taxpayers are misinformed about taxes. The intuition is simple and hinges on the key assumption that while reducing "visibility" decreases the perception of one taxpayer, it increases the perception to the other taxpayer. As long as the taxpayer whose perception has been increased is the active one, the government gets smaller.

3.4 Predictions for the size of government

In this subsection I return back to the split of the payroll tax and describe the model's predictions for the size of government as a function if the split. I adopt the naive assumption that on the nature of taxpayers' misperceptions, that the perceived tax incidence is equal to the nominal incidence, hence $a^\phi = a^N$ and $b^\phi = 1 - a^N$. Based on the empirical studies quoted above, I assume that the tax is fully born by workers due to inelastic labor supply. The after

[24] Another, less obvious case, is when $\mathsf{a}^R > \mathsf{b}^R$ but B should be the active taxpayer under perfect information because $\mathsf{a}^R(1+\Delta'_A) < \mathsf{b}^R(1+\Delta'_B)$; i.e, the deadweight cost falls disproportionately on him. The remaining two cases are when $\mathsf{a}^R < \mathsf{b}^R$ and B should be the active taxpayer and finaly when $\mathsf{a}^R < \mathsf{b}^R$ but A should be active. The analysis of these cases is analogous to the case presented here.

tax wage to the employer is fixed at w^* regardless of the tax and is equal to the marginal product of labor. The gross (contracted) wage w_g and the workers' take-home wage w_h are in equilibrium determined by

$$w_g = \frac{w^*}{1+t_A} \tag{7}$$

$$w_h = w_g(1-t_B) = w^*\frac{1-t_B}{1+t_A} \tag{8}$$

The tax rate T is defined as in equation (2). It follows that the real deadweight cost of the tax is zero. This is probably an overstatement, although payroll tax is regarded as one of the most efficient ways of raising revenue. Given the split, workers think they pay $(1-a^N)Tw^*$ in taxes, but their perceived deadweight costs is always zero since they would not supply more labor if the tax were repealed and their take-home wage rose. The employers think they pay $a^N Tw^*$ and their perceived deadweight costs is positive, increasing and convex, since they would demand more labor if the tax were repealed and their wage costs would drop (as they believe) from w^* to w_g.

When $a^N = 0$, only the workers are politically active. As the split of the tax tilts more toward the employer, the workers reduce pressure while the employers remain inactive and the government grows. At some critical point a_c^N there is a switch and the employers become the active taxpayer group while the workers become inactive. The critical point is defined implicitly by $a_c^N(1+\Delta_A^{\phi\prime}(a_c^N Tw^*)) = 1-a_c^N$. Because the perceived deadweight cost is positive for employers but zero to workers, the employers have a stronger incentive to be politically active and therefore $a_c^N < 1/2$. As the split increases beyond the switch point, the employers produce more pressure and the government gets smaller until $a^N = 1$, where the government is minimized. In fact, the government is globally minimized at $a^N = 1$, since the pressure by employers when $a^N = 1$ is greater than the pressure by workers when $a^N = 0$, because the employers perceive higher deadweight costs.

How do these predictions compare with the data? In a vast majority of countries (76 out of 89) the employer's share is greater or equal 1/2, hence the intensity of employers' political pressure should drive most of the data, and the observed taxes should be declining with employer's share. The results from the linear specifications are consistent with this prediction. To verify the global prediction that the tax should be lower if the employers pay the entire tax than if the workers pay the entire tax, I repeated the linear regressions in Tables 2, 3, 5, 6 and 7 on countries where the share is either below 0.25 or above 0.75. (The sample is small, however, from 8 to 11 countries) The results support the prediction. The coefficients on share are always negative, very similar to those found on the full sample, and in two cases are strongly significant. Finally, the model's prediction that the tax should be increasing for low values of the split and declining for higher values was the reason why I included a square term in the regressions (columns (4)-(6) in all Tables). Unfortunately, when significant, the linear coefficient is negative and the quadratic coefficient is positive, indicating a U-shape rather than an inverted U-shape. The tax appears to be minimized when the split is around 2/3 rather than being maximized for at a split somewhere below 1/2. Here the model fails. The reason may be that the assumption $F(A,B,C) = F(A+B,C)$ is too restrictive as it completely rules out the workers from the political process, or the relationship between nominal and perceived tax incidence is more complicated.

4 Alternative explanations

An effort to rationalize the negative relationship between the split of the payroll tax and the level of taxes by taxpayers' misperceptions runs into the problem that we do not have a theory of how taxpayers form their perceptions about taxes and how inputs into perceptions eventually translate into the values of a^{ϕ}, b^{ϕ}. Here I briefly discuss two reasons other than perceived tax incidence why the split may negatively influence the size of government.

First reason is the *subsidy recipients'* ignorance about tax incidence. They may think that increasing the tax on employer by 1 percentage point while reducing the tax on worker by 1 percentage point will not change the tax revenue, since they may not realize that the gross wage will adjust to the change in tax rates. But while such a change in tax rates may seem innocuous, it actually benefits the taxpayers through the equilibrium adjustment in wages. From equation 8 it follows that if $dt_B = -dt_A$, then

$$\frac{dw_h}{dt_A} = w^* \frac{t_A + t_B}{(1 + t_A)^2} > 0$$

The worker's take-home wage rises while the firm's labor cost is unchanged, therefore one taxpayer is better off and the other taxpayer is not worse off. Whether subsidy recipients are indeed being fooled this way is another question to which I do not have an answer.

The second reason is slow adjustment in nominal wages. As Hamermesh (1980) shows, if wages are not adjusting instantaneously, the allocation of nominal incidence matters for the real incidence when tax rates change. When the tax is reduced, the side which nominally paid the tax enjoys an extra benefit before the wage adjusts to the new equilibrium. Poterba, Rotemberg and Summers (1986) actually use the changes in tax rates to estimate the speed of adjustment in prices and wages in a general equilibrium setting. They find very slow adjustment in wages (in the US) - more than 5 years are needed to bring the wages close to the new equilibrium.

For the split of the payroll tax, this would imply that employers have a rational reason to oppose taxes even if they realize that they will not benefit in the long run. Before the gross wage rises to the equilibrium, they will pocket some of the tax reduction. And their gains from a tax reduction are larger the higher their share of the tax, and therefore they have a stronger incentive to be politically active. If the adjustments are indeed that slow as Poterba, Rotemberg and Summers indicate, the short-term gain could be substantial. The adjustment could be used to put the "misperception" on a firm ground. The "misperception" could be defined as the difference between what each taxpayer would save if the tax were repealed and the wages adjusted instantaneously, and what he would save if the tax were repealed but the wages adjusted as slowly as they do in the real world. The perceived tax incidence would be replaced by real tax incidence in a dynamic context with slow nominal wage adjustments. This concept is, however, a substantial departure from what we mean by visibility of taxes, although it is by itself an interesting subject for future research.

5 Conclusions

The goals of the paper were: first, to perform a new, and, as I argued in section 2, cleaner test of the popular belief that making taxes more visible (such as by eliminating the employer's portion of the payroll tax) would reduce the size of government. The conclusion was that

countries where employers pay a higher fraction of the payroll tax in fact appear to have lower taxes and definitely do not have higher taxes. The second goal was to provide the first rigorous model of taxpayer's misperceptions, and one of the conclusions of the model was that initiatives calling for more visible taxes may be counterproductive to their intended goal: They may fail to account for their effect on the perceived tax incidence of other taxpayers. Despite this, the traditional visibility argument could still be valid in contexts other than the payroll tax. For example, it is hard to imagine how income tax withholding, which presumably reduced the visibility of the income tax to workers, could have increased the visibility of the tax to some other taxpayers. Therefore the validity of the visibility argument has to be considered on a tax-by-tax basis, taking into account factors such as real tax incidence, deadweight costs, and especially the sources of misperception for all taxpayers.

References

[1] Anderson, Patricia M., and Bruce D. Meyer: *Using a Natural Experiment to Estimate the Effects of the Unemployment Insurance Payroll Tax on Wages, Employment, Claims, and Denials*, NBER working paper 6808 (1998).

[2] Becker, Gary S.: *A Theory of Competition among Pressure Groups for Political Influence*, Quarterly Journal of Economics (1983), Vol. 98(3): 371-400.

[3] Becker, Gary S. and Casey B. Mulligan: *Deadweight Costs and the Size of Government*, NBER working paper 6789 (1999).

[4] Besley, Timothy J., and Harvey S. Rosen: *Sales Taxes and Prices: An Empirical Analysis*, NBER working paper 6667 (1998).

[5] Bises, Bruno: *Income Tax Perception and Labour Supply in a Sample of Industry Workers*, Public Finance (1990), Vol. 45(1): 3-17.

[6] Buchanan, James M., and Richard E. Wagner: *Democracy in Deficit*, Academic Press (1977).

[7] Clotfelter, Charles T.: *Public Spending for Higher Education: An Empirical Test of Two Hypotheses*, Public Finance (1976), Vol. 31(2): 177-95.

[8] Cragg, John G., Arnold C. Harberger and Peter Mieszkowski: *Empirical Evidence on the Incidence of the Corporation Income Tax*, Journal of Political Economy (1967), Vol. 75(6): 811-821.

[9] Goetz, Charles J.: *Fiscal Illusion in State and Local Finance*, in Borcherding, Thomas E. (ed.): Budgets and Bureaucrats: The Sources of Government Growth, Duke University Press (1977).

[10] Gordon, Robert J.: *The Incidence of the Corporation Income Tax in U.S. Manufacturing*, American Economic Review (1967), Vol.57(4): 731-758.

[11] Gruber, Jonathan: *The Incidence of Payroll Taxation: Evidence from Chile*, NBER working paper 5053 (1995).

[12] Hamermesh, Daniel S.: *Factor Market Dynamics and the Incidence of Taxes and Subsidies*, Quarterly Journal of Economics (1980), Vol. 95(4): 751-764.

[13] Hines, James R., and Richard H. Thaler: *Anomalies: The Flypaper Effect*, Journal of Economic Perspectives (1995), Vol. 9(4): 217-226.

[14] Krzyzaniak, Marian, and Richard A. Musgrave: *The Shifting of the Corporation Income Tax*, Johns Hopkins Press (1963).

[15] Logan, Robert R.: *Fiscal Illusion and the Grantor Government*, Journal of Political Economy (1986), Vol. 94(6):1304-1318.

[16] Mackinac Center for Public Policy: *The Right to Know Payroll Form*, The Mackinac Center website (www.mackinac.org).

[17] Meltzer, Allan H. and Scott F. Richard: *A Rational Theory of the Size of Government*, Journal of Political Economy (1981), Vol. 89(51): 914-927.

[18] Milton Friedman in Prague, Liberalni institut Praha (1997).

[19] Mulligan, Casey B., and Xavier Sala-i-Martin: *Gerontocracy, Retirement and Social Security*, NBER working paper 7117 (1999).

[20] Oates, Wallace E.: *On the Nature and Measurement of Fiscal Illusion: A Survey*, reprinted in Oates, Wallace E. (ed.): *Studies in Fiscal Federalism*, Edward Elgar Publishing (1991): 431-448.

[21] Peltzman, Sam: *Toward a More General Theory of Regulation*, Journal of Law and Economics (1976), Vol. 19(2): 211-240.

[22] Peltzman, Sam: *The Growth of Government*, Journal of Law and Economics (1980), Vol. 23(2): 209-287.

[23] Pommerehne, Werner W. and Friedrich Schneider: *Fiscal Illusion, Political Institution and Local Public Spending*, Kyklos (1978), Vol. 31(3): 381-408.

[24] Poterba, James M, Julio J. Rotemberg and Lawrence H. Summers: *A Tax-Based Test for Nominal Rigidities*, American Economic Review (1986), Vol. 76(4): 659-675.

[25] Stansel, Dean: *The Hidden Burden of Taxation: How the Government Reduces Take-Home Pay*, Cato Policy Analysis No. 302 (1998).

[26] Wagner, Richard E: *Revenue Structure, Fiscal Illusion and Budgetary Choice*, Public Choice (1976), Vol. 25(25): 45-61.

Figure 1
Social Security Tax Rates, OECD, 1995

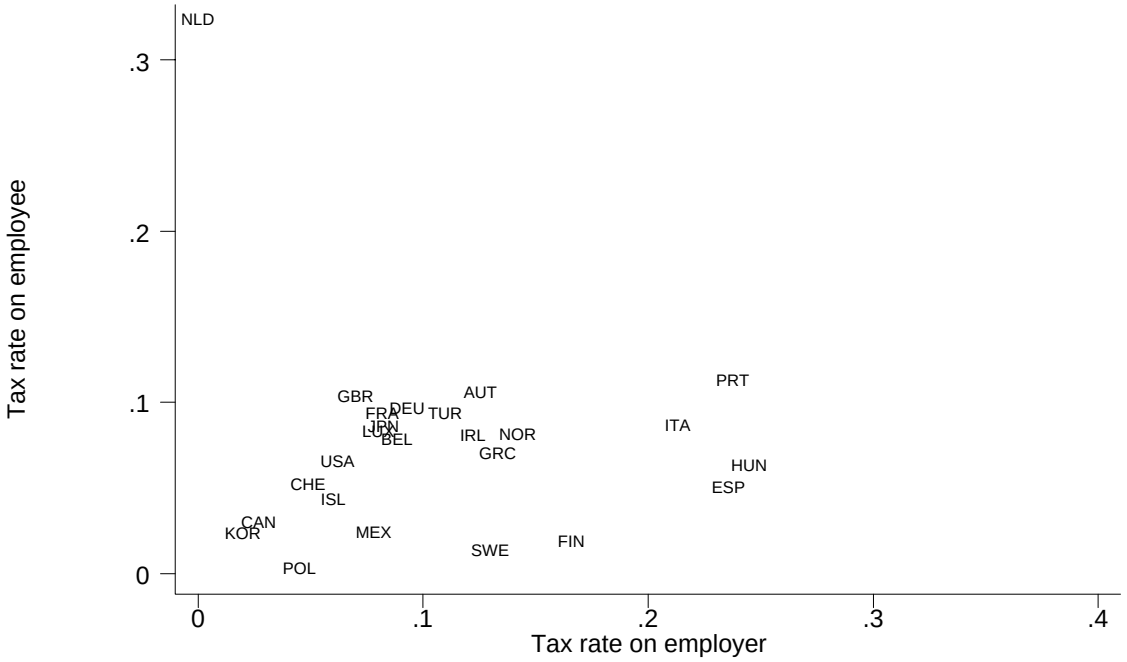


Figure 2
Social Security Tax Rates, non-OECD, 1995

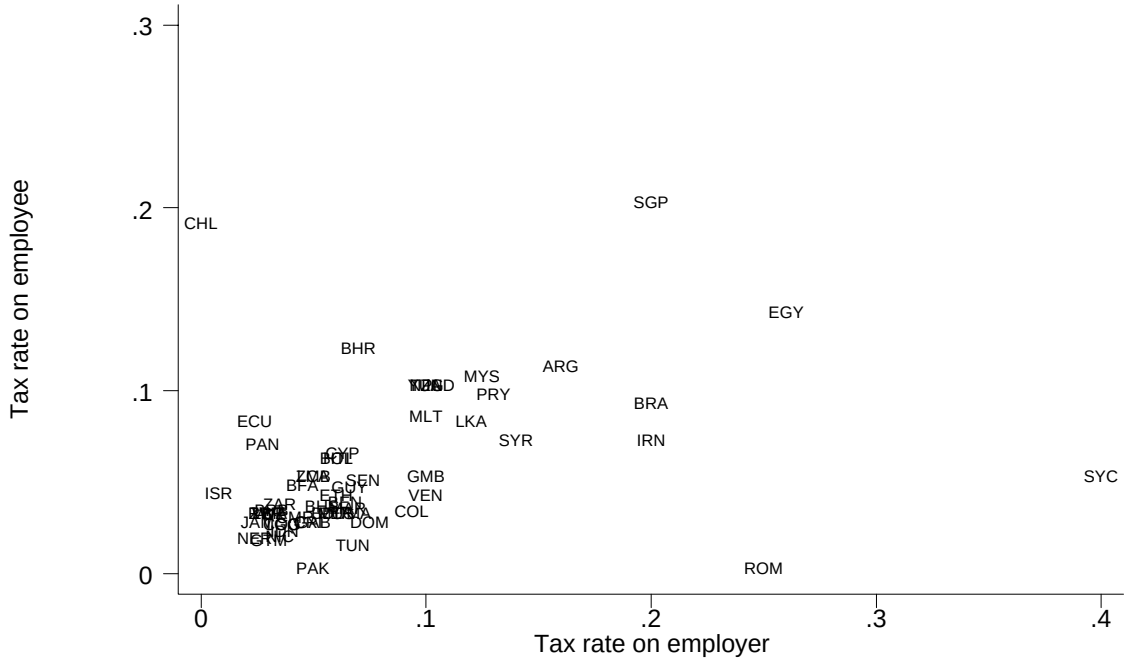


Figure 3
Levels of Tax Rate and Employer's Share, OECD, 1995

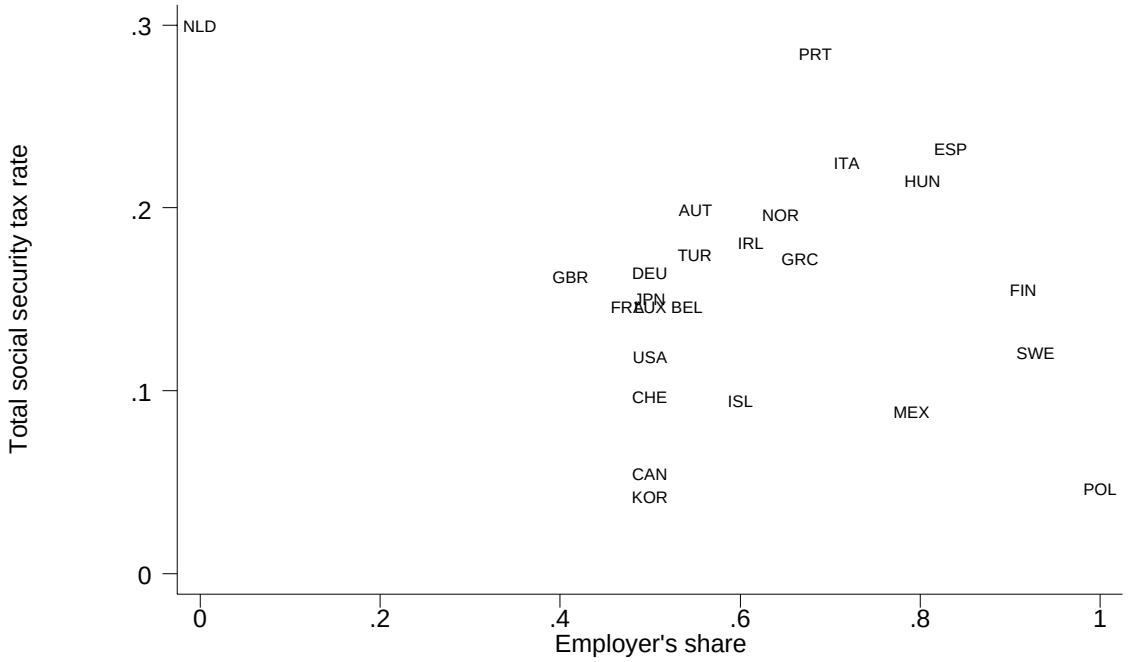


Figure 4
Levels of Tax Rate and Employer's Share, non-OECD, 1995

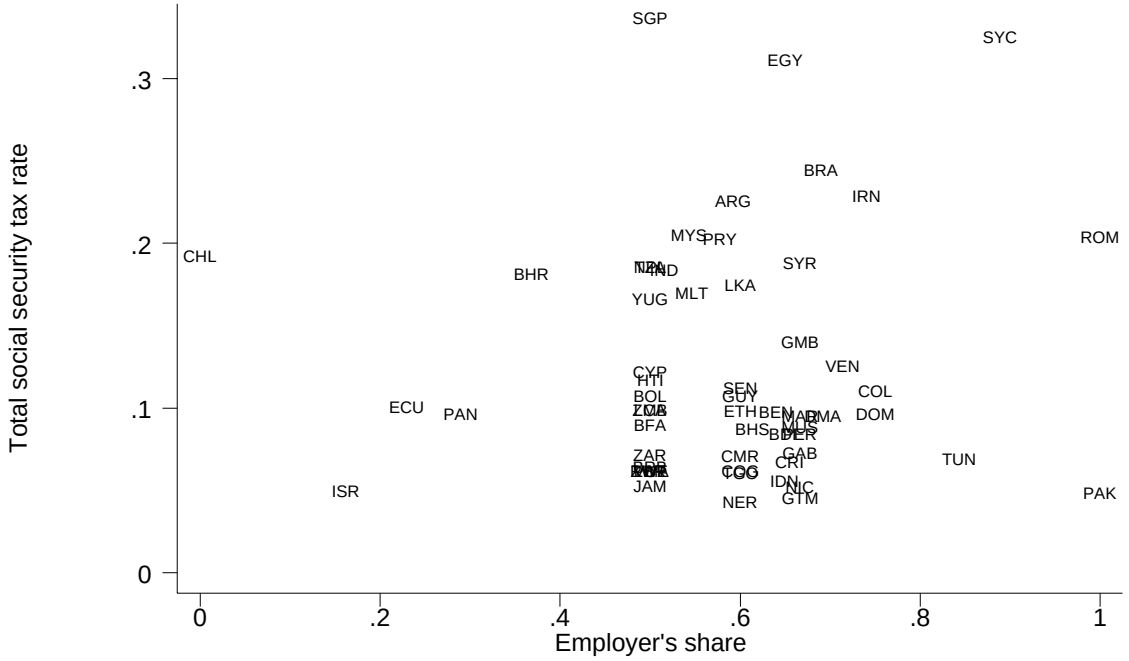


Figure 5

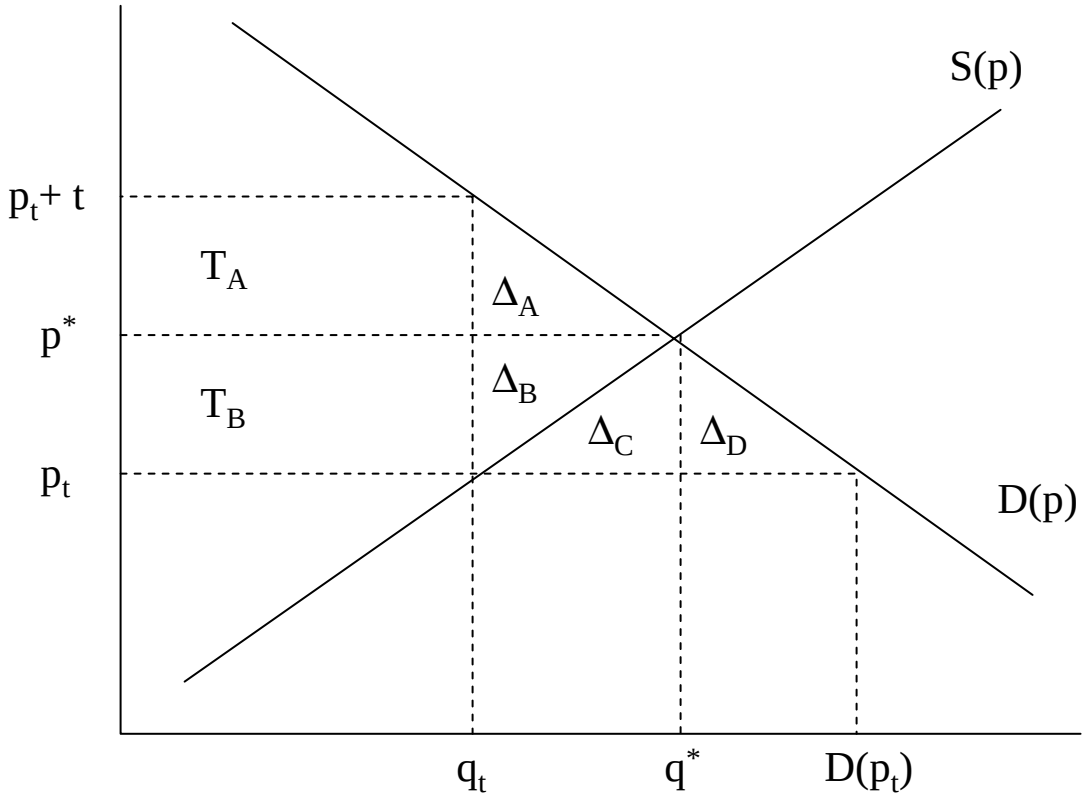
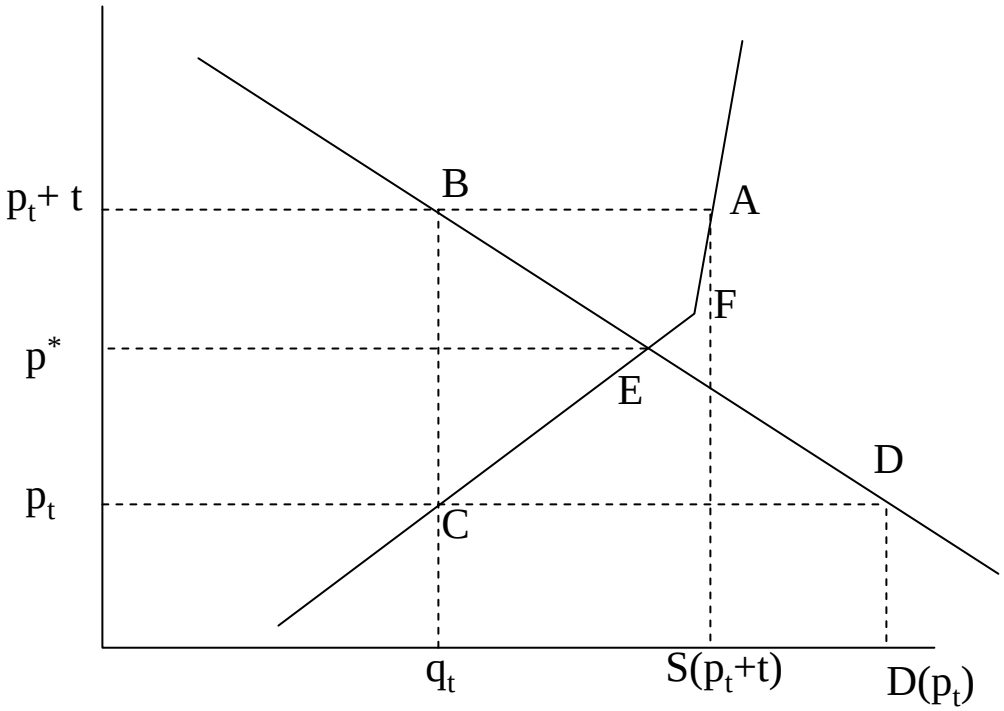


Figure 6



Social security system parameters

All countries

	Obs	Mean	Std. Dev	Min	Max
Employer's share of the social security tax (1995)	84	0.591	0.178	0.000	1.000
Social security tax rate (1995)	89	0.123	0.076	0.000	0.333
Tax rate on employer (1995)	89	0.084	0.071	0.000	0.400
Tax rate on employee (1995)	89	0.057	0.049	0.000	0.321
Pension spending as a share of GDP (1980-1989)	79	0.032	0.035	0.000	0.117
Ratio of the elderly (1980)	83	0.092	0.052	0.035	0.219

OECD countries

Employer's share of the social security tax (1995)	25	0.609	0.206	0.000	1.000
Social security tax rate (1995)	27	0.142	0.076	0.000	0.296
Tax rate on employer (1995)	27	0.100	0.072	0.000	0.245
Tax rate on employee (1995)	27	0.067	0.061	0.000	0.321
Pension spending as a share of GDP (1980-1989)	26	0.070	0.031	0.004	0.117
Ratio of the elderly (1980)	26	0.153	0.042	0.054	0.219

Non-OECD countries

Employer's share of the social security tax (1995)	59	0.583	0.167	0.000	1.000
Social security tax rate (1995)	62	0.114	0.075	0.000	0.333
Tax rate on employer (1995)	62	0.078	0.071	0.000	0.400
Tax rate on employee (1995)	62	0.052	0.042	0.000	0.200
Pension spending as a share of GDP (1980-1989)	53	0.013	0.018	0.000	0.072
Ratio of the elderly (1980)	57	0.065	0.027	0.035	0.141

Correlations of employer's share accross years

	1985	1989	1995	1999
1985	1			
1989	0.8938	1		
1995	0.87	0.9092	1	
1999	0.7376	0.8374	0.8519	1

Correlation between the split of the social security tax and the medical tax

Year	Obs	Correlation
1985	42	0.173
1989	44	0.261
1995	45	0.432
1999	45	0.297

Table 2
Cross sectional regressions

Independent variables	Dependent variable: Pension expenditure/GDP, 1980-89 average					
	Employer's share in 1985					
	(1) all countries	(2) OECD	(3) non-OECD	(4) all countries	(5) OECD	(6) non-OECD
Employer's share of social security tax	-0.0246	-0.0256	-0.0506	-0.0967	-0.2504	-0.0848
	(0.036)	(0.357)	(0.000)	(0.006)	(0.021)	(0.008)
Employer's share squared				0.0687	0.1844	0.0385
				(0.029)	(0.030)	(0.227)
1980 Log(GDP per capita)	0.0028	-0.0209	0.0059	0.0015	-0.0231	0.0054
	(0.417)	(0.127)	(0.036)	(0.662)	(0.063)	(0.056)
1980 fraction elderly	0.5375	0.6950	0.2414	0.5519	0.7972	0.2506
	(0.000)	(0.000)	(0.007)	(0.000)	(0.000)	(0.005)
Observations	67	23	44	67	23	44
Adjusted R^2	0.79	0.59	0.63	0.81	0.67	0.63

(p-values in parentheses)

Table 3
Cross sectional regressions (cont.)

Independent variables	Dependent variable: Pension expenditure/GDP, 1980-89 average					
	Employer's share, 1985-1999 average					
	(1) all countries	(2) OECD	(3) non-OECD	(4) all countries	(5) OECD	(6) non-OECD
Employer's share of social security tax	-0.0199	-0.0120	-0.0463	-0.1354	-0.2945	-0.0991
	(0.100)	(0.704)	(0.000)	(0.000)	(0.006)	(0.004)
Employer's share squared				0.1094	0.2411	0.0553
				(0.001)	(0.006)	(0.090)
1980 Log(GDP per capita)	0.0022	-0.0215	0.0055	0.0007	-0.0213	0.0047
	(0.521)	(0.131)	(0.053)	(0.835)	(0.070)	(0.091)
1980 fraction elderly	0.5514	0.7364	0.2717	0.5642	0.7922	0.2853
	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.001)
Observations	68	23	45	68	23	45
Adjusted R^2	0.79	0.57	0.61	0.82	0.63	0.72

(p-values in parentheses)

Table 4
Instrumental Variable Regression

Independent variables	Dependent variable: Pension expenditure/GDP, 1980-89 average			
	Employer's share in 1985		Employer's share, 1985-1999 average	
	(1) all countries	(2) all countries	(3) all countries	(4) all countries
Employer's share of social security tax	0.0245	0.0238	-0.0973	-0.1163
	(0.816)	(0.834)	(0.287)	(0.227)
Employer's share squared		0.0042		0.0362
		(0.957)		(0.687)
1980 Log(GDP per capita)	0.0006	0.0006	-0.0052	-0.0043
	(0.816)	(0.943)	(0.580)	(0.616)
1980 fraction elderly	0.6093	0.6118	0.7006	0.6869
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	40	40	46	46
Adjusted R^2	0.72	0.70	0.64	0.72

(p-values in parentheses)

Table 5
Regressions with tax rates as the dependent variable

Independent variables	Dependent variable: Social security statutory tax rate, 1985-99 average							
	Employer's share, 1985-1999 average							
	Cross-sectional regressions						IV regressions	
	(1) all countries	(2) OECD	(3) non-OECD	(4) all countries	(5) OECD	(6) non-OECD	(7) all countries	(8) all countries
Employer's share of social security tax	-0.0402	-0.0578	-0.0556	-0.1473	-0.6775	-0.0494	-0.0125	0.0029
	(0.350)	(0.358)	(0.327)	(0.311)	(0.002)	(0.782)	(0.954)	(0.992)
Employer's share squared				0.1000	0.5164	-0.0062		-0.0229
				(0.440)	(0.003)	(0.970)		(0.944)
1985 Log(GDP per capita)	0.0057	-0.1248	0.0201	0.0053	-0.1251	0.0201	-0.0132	-0.0137
	(0.613)	(0.001)	(0.121)	(0.639)	(0.000)	(0.125)	(0.589)	(0.598)
1990 fraction elderly	0.4211	1.5180	-0.0166	0.4115	1.5483	-0.0146	0.6516	0.6603
	(0.030)	(0.000)	(0.959)	(0.034)	(0.000)	(0.965)	(0.063)	(0.081)
Medical tax joint with social security	0.0565	0.0284	0.0480	0.0564	0.0253	0.0479	0.0948	0.0960
	(0.000)	(0.329)	(0.017)	(0.001)	(0.267)	(0.019)	(0.001)	(0.006)
Observations	77	24	53	77	24	53	50	50
Adjusted R^2	0.27	0.62	0.12	0.26	0.77	0.10	0.29	0.28

(p-values in parentheses)

Table 6
Regressions with lagged tax rates

Independent variables	Dependent variable: Social security statutory tax rate, 1995					
	Employer's share, 1985					
	(1) all countries	(2) OECD	(3) non-OECD	(4) all countries	(5) OECD	(6) non-OECD
Employer's share of social security tax	-0.0544	-0.1055	-0.0061	0.0445	-0.2815	0.0236
	(0.064)	(0.210)	(0.841)	(0.615)	(0.475)	(0.780)
Employer's share squared				-0.0929	0.1350	-0.0316
				(0.239)	(0.645)	(0.706)
Tax rate in 1985	0.7757	0.2275	0.9374	0.7835	0.1393	0.9379
	(0.000)	(0.423)	(0.000)	(0.000)	(0.688)	(0.000)
1985 Log(GDP per capita)	0.0050	-0.0648	-0.0013	0.0054	-0.0766	-0.0011
	(0.518)	(0.244)	(0.853)	(0.484)	(0.225)	(0.870)
1990 fraction elderly	-0.1008	0.7160	0.0992	-0.0970	0.8809	0.1097
	(0.446)	(0.224)	(0.565)	(0.462)	(0.215)	(0.534)
Medical tax joint with social security	-0.0111	-0.0020	-0.0149	-0.0117	0.0018	-0.0153
	(0.330)	(0.954)	(0.159)	(0.304)	(0.960)	(0.154)
Observations	74	23	51	74	23	51
Adjusted R^2	0.66	0.30	0.80	0.67	0.26	0.80

(p-values in parentheses)

Table 7
Panel data regressions

Independent variables	Dependent variable: Social security statutory tax rate					
	(1) all countries	(2) OECD	(3) non-OECD	(4) all countries	(5) OECD	(6) non-OECD
Employer's share of social security tax	-0.0509	-0.0797	-0.0069	-0.0272	0.1206	-0.1460
	(0.021)	(0.054)	(0.766)	(0.660)	(0.452)	(0.014)
Employer's share squared				-0.0214	-0.1514	0.1505
				(0.682)	(0.197)	(0.011)
Fraction elderly	0.4990	0.1669	0.6264	0.5021	0.1071	0.5683
	(0.000)	(0.557)	(0.010)	(0.000)	(0.708)	(0.019)
Total government revenue/GDP	0.0639	0.2990	-0.0086	0.0629	0.3320	0.0018
	(0.209)	(0.010)	(0.859)	(0.217)	(0.005)	(0.970)
Medical tax joint with social security	0.0323	0.0431	0.0227	0.0326	0.0447	0.0223
	(0.001)	(0.033)	(0.033)	(0.001)	(0.026)	(0.032)
Observations	218	79	139	218	79	139
Countries	71	24	47	71	24	47
Adjusted R^2	0.26	0.45	0.09	0.26	0.43	0.06

(p-values in parentheses)
(The panel covers year 1985, 1989, 1995, and 1999)