Optimal Taxation of Employment and Self-Employment: Evidence from Poland and Implications*

Tomasz Zawisza†
European University Institute, IFS

September 4, 2019

Abstract

I examine two central questions bearing on the optimal relative taxation of employment and self-employment income using the 2009 Polish reform. Firstly, I estimate the degree of substitution on the extensive margin between the employment and self-employment tax bases, as well as the intensive-margin elasticities for these tax bases separately. The baseline estimates of the intensive-margin elasticities are around 0.2 for the employed and around 0.7 for the self-employed. Secondly, I characterize optimal non-linear employment and self-employment tax schedules in a theoretical model, finding rationale for taxation of self-employment income at lower marginal and total tax rates.

JEL Codes: D31, H21, H24, H25, J24

Keywords: employment, self-employment, optimal taxation, occupational choice

*I gratefully acknowledge financial support from the ESRC, the Royal Economic Society and the Richard Kahn Fund at the University of Cambridge. This paper comprises two chapters from my doctoral thesis. I am especially grateful to Hamish Low, Sriya Iyer, Toke Aidt, Henrik Kleven, Wojciech Kopczuk, Emmanuel Saez and Florian Scheuer for comments and suggestions. I am grateful to the Polish Ministry of Finance, and in particular to Tomasz Szałwinski and Tomasz Mazur for help preparing and obtaining access to data. Declarations of interest: none.

†E-mail: tazawisza@gmail.com.

1 Introduction

The question of the relative tax treatment of different tax bases, and of the self-employment and employment tax bases in particular, is a pressing one for policymakers. It is also a significant topic of political debate. In early 2017, for instance, the Conservative government in Britain made a decision to increase national insurance contributions for the self-employed to bring them more in line with the rates faced by workers, justifying the proposal by arguing that it would introduce more equitable tax treatment between the employment and self-employment. However, it was subsequently retracted by the government in the face of widespread controversy.

A central result of public economics is that under certain conditions (Feldstein, 1995), the elasticity of taxable income (ETI) – the degree to which declared taxable income falls with increasing tax rates – is a sufficient statistic to measure the efficiency of taxation.[1] All other things being equal, a higher elasticity of taxable income is found to imply a greater efficiency loss and hence a lower optimal marginal tax rate. By implication, if self-employment income exhibits a higher taxable income elasticity than employment, it should *ceteris paribus* be taxed at lower marginal tax rates. Indeed, previous studies, such as Kopczuk (2015) or Saez (2010), suggest that self-employment income has a higher ETI than employment income, suggesting a *prima facie* case for taxing self-employment income at a lower marginal rate than income from employment.

However, it is also understood that a different treatment of income categories by the tax system may lead taxpayers to move between these categories. For instance, Gordon and Slemrod (2000) argue that the response to the Reagan 1986 tax cuts in the personal income tax base may have been driven mainly by taxpayers shifting from the corporate tax base. Such responses to tax reforms, which have been termed *fiscal externalities* (Saez et al., 2012), create further efficiency losses, and have dominated the policy discussion surrounding relative rate of taxation between employment and self-employment. Crawford and Freedman (2010) argue, for instance, that tax rates for the employed and the self-employed should be equalized, precisely on the grounds of reducing distortionary substitution between the two bases. As a result, there is a trade-off between adjusting marginal tax rates in response to higher elasticities among some categories on the one hand and creating incentives to switch tax categories on the other. The precise quantitative significance of these trade-offs has not been ascertained, however, neither in terms of estimating relevant parameters, nor their implications for optimal tax schedules.[2]

This paper seeks to address both the empirical and theoretical challenges in the con-

[1] Saez, Slemrod and Giertz (2012) provide a comprehensive survey. The literature highlights three factors as being important for finding optimal tax rates: the ETI, preferences for redistribution and the mass of individuals located at different points in the income distribution.

[2] Higher elasticities for workers may also incentivise the social planner to push individuals out of self-employment, and thereby be able to subject them to higher marginal tax rates.

text of employment and self-employment by taking advantage of a tax reform episode which occurred in Poland in 2009. In particular, I first estimate the parameters governing the efficiency cost of taxes in the presence of tax base shifting between these two tax bases. I then feeds these into a theoretical model of optimal taxation to work out quantitative implications in a stylized framework. Since the Polish tax reform altered the relative difference in the tax burdens between an employed individual and a self-employed individual with the same income, it allows the estimation of fiscal externalities between these two tax bases. Since the fraction of self-employed individuals who also report employment income is small (about a quarter of the total) as a first I focus on fiscal externalities on the extensive margin, whereby individuals switch entirely between self-employment and employment activity.[3] This is to my knowledge the first paper to estimate such extensive-margin externalities and asses their implications for optimal taxation. The reform also allows the estimation of traditional intensive-margin ETIs separately for the two main types of income under consideration, alongside the fiscal externalities. Looking at the intensive and extensive-margin responses to tax reform together within one tax system here is natural, since both matter for determining optimal relative rates rates of taxation.

Interestingly, I find stable own-elasticities of employment income across the income distribution in the 0.2-0.3 range. Own-elasticities of business income are higher, in the 0.5-0.7 range, although it is more difficult to estimate how stable these are across the income distribution because of small sample size. There are also non-trivial responses occurring at the extensive margin: a 10,000zł per annum ($2,672 as of May 2017[4]) change in the relative tax burden between the two tax bases increases the probability of a taxpayer filing linear business income, as opposed to progressive employment income, by 3.75 percentage points. The estimates allow us to find the contributions of the extensive margin to changes in the deadweight losses at the time of the 2009 reforms in Poland at 7% of the total.

I incorporate these empirical results into a model of general non-linear taxation of employees and the self-employed. In particular, I am interested in the implications of this behaviour for the design of socially optimal tax schedules. Lower marginal tax rates at a given point in the optimal tax schedule will lead to a higher tax differential higher up the schedule. As has already been mentioned, the phenomenon of individuals transitioning between tax bases to take advantage of favourable tax differentials introduces another distortion which acts in the opposite direction against a tax system which introduces large differences in the levels of taxation of self-employment and employment. Ultimately, it is a quantitative question whether an optimal tax schedule is likely to feature differential

[3] In an extension of the theoretical model in section 4.2 of the Online Appendix, I allow there to be fiscal externalities on both the intensive and extensive margins. In this scenario, individuals with both types of income adjust the relative amount of income declared in each tax base and for them an intensive-margin cross-elasticity becomes an addition to the optimal-tax formulae. Since the majority of taxpayers report either one or the other type of income, however, I here focus on the extensive-margin and leave a quantitative exploration of the richer model for future work.

[4] zł denotes the Polish currency, *złoty*.

tax treatment (in terms of marginal and/or total tax rates) for self-employment relative to employment. The approach is to augment the standard model of optimal taxation analyzed in detail by Mirrlees (1971), Diamond (1998) and Saez (2001), among others. Specifically, I introduce an additional tax base, the self-employment tax base, with possible switching on the extensive margin between the two, following Scheuer (2014). The results suggest that, for the calibrations employed here using the Polish dataset and estimates, self-employment should be taxed at a marginal rate lower than employment by around 10 to 20 percentage points for most of the income distribution. The presence of extensive margin transitions only moderately offsets the gains from lower marginal tax rates on self-employment income.[5]

The paper links to several strands in the public finance literature. Most closely related is Kopczuk (2015), who examines business income elasticities around the introduction of the linear tax option for business income in 2004 in Poland using a methodology also based on joint filing with a spouse. This study is complementary to his - a similar identification strategy is applied to a sample of business owners for intensive-margin elasticities. However, his paper does not examine the extensive-margin implications of differential taxation of tax bases. Kleven and Schultz (2014), Alstadster and Jacob (2016), Harju and Matikka (2016a) and Harju and Matikka (2016b) all examine cross-elasticities between business and employment income using reforms in Scandinanvian countries. However, they focus purely on intensive-margin responses, which are unlikely to be the dominant response as only a relatively small fraction of the self-employed also report employment income. From a theoretical perspective, Kleven et al. (2009) is a seminal contribution focusing on the joint taxation of a household involving a primary taxpayer and the spouse. The decision whether to work or not of a taxpayer's spouse in this paper shares many similarities with the choice to switch tax bases here. Scheuer (2012) applies an analogous framework to Kleven et al. (2009) to examine optimal non-linear taxation in environment with an extensive-margin occupational choice between employment and business. However, the calibrations of the relevant parameters employed do not incorporate the different elasticities between the tax bases and, moreover, the parameters driving the extensive-margin responses to taxation are not founded in estimates. Therefore, I regard this paper as complementary to his theoretical analysis.

The paper also deals with a number of methodological challenges which surrounds the identification of intensive-margin ETI parameters, and as such also makes a methodolog-

[5] As Slemrod and Kopczuk (2002) and Kopczuk (2005) have pointed out, parameters such as the elasticity of taxable income are not necessarily fixed, but depend on other policy considerations such as tax exemptions, the strictness of tax-enforcement and other contingent institutional features. This implies that policies such as tightening the rules around switching between employment and self-employment or enhancing third-party reporting for self-employment may be efficient options for policymakers. Thus, it should be borne in mind that the above conclusions are made, *taking the tax enforcement and reporting rules as given.* Incorporating non-tax changes in an analysis optimal tax policy in the context of self-employment is an important complementary avenue of research.

ical contribution. The dominant empirical approach has been to use tax reforms which create exogenous variation in marginal tax rates for some sub-groups of taxpayers while leaving them unchanged for others. Classic papers in this spirit include Feldstein (1995), Gruber and Saez (2002) and Kopczuk (2005). It is unclear whether these existing studies adequately controlled for the changes in taxable income which would have occurred anyway, even in the absence of tax reforms. Since tax reforms are usually concentrated in certain sub-sections of the income distribution, researchers have tended to assume that the pattern of income growth across the income distribution is stable over time, absent reforms. This, however, may not hold in practice. For instance, business cycle factors affect some parts of the income distribution more than others, income inequality trends may also change over time and, finally, people at different points in the distribution and different types of income may have different elasticities (Saez et al., 2012).

The Polish reform episode permits an identification strategy for intensive-margin elasticities which does not require the assumption of stable income dynamics over time across the income distribution. Specifically, the Polish system gives the option to file jointly with spouses. Joint filing with a lower-income spouse may allow an individual to enter a lower income tax bracket than under single filing, while an individual with the same income but a higher-income spouse may be forced to remain in a higher tax bracket. It is the latter taxpayers who experienced a large cut in MTRs as a result of Poland's 2009 tax reforms, while the former did not. I thus obtain a treatment and a control group which is independent of the position in the taxpayer's own position in the income distribution. For estimating extensive-margin elasticities, I rely both on an analogous estimation strategy, where individuals with spouses provide a control group to individuals transitioning to and from the linear business tax base, as well as transitions around income brackets.[6]

The paper is related to a number of more recent studies which have attempted to address some of the drawbacks of the classic approach in the empirical ETI literature (e.g. Burns and Ziliak (2017) and Weber (2014)). An important paper in this regard is Weber (2014), who argues that the traditional instruments in the spirit of Gruber-Saez are correlated with the idiosyncratic term under reasonable assumptions about the dynamics of the earnings process. Her proposed solution is to use longer lags of income in predicting changes in marginal tax rates, which she argues are not correlated with idiosyncratic earnings components with limited memory. When applied to the Polish case, I show that the Weber (2014) methodology does produce estimates closer to those produced by the method proposed here, especially at short horizons. Nonetheless, it still

[6] The main assumption for identifying intensive-margin elasticities used here is that, absent reforms, changes in the taxable income declared by a taxpayer are independent of his or her spouse's income level, once I condition for the taxpayer's own base-year income. Due to the availability of several years where no reforms occurred, I can verify its validity by conducting placebo tests something which is unfortunately rarely done in the literature (Kopczuk (2015) is a notable exception). Likewise, extensive-margin estimates are validated by placebo estimates in non-reform years.

produces somewhat divergent results and, additionally, the discrepancies increase the longer is the time horizon taken. Fundamentally, this alternative method still relies on an extrapolation of the pattern of earnings dynamics from past to the reform year and in this regard may still have limited relevance in situations where there is time-variation in earnings dynamics across the income distribution.

The structure of the remainder of the paper is as follows. Section 2 gives background information on the Polish tax system and the 2009 tax reform. Section 3 develops a model of tax reporting in an environment with multiple tax bases and demonstrates how fiscal externalities affect the formulae for the deadweight loss (DWL) of tax reforms. Section 4 outlines the dataset. Sections 5 and 6 present the empirical results on the intensive and extensive margin. Section 7 applies these results to calculate the DWL of the Polish tax reform. In Section 8, I present the calibration of the model of optimal taxation based on the empirical results in Sections 5 and 6. The paper concludes in Section 9.

2 Background: The 2009 Polish Tax Reform

The basic Polish personal income tax system has a progressive structure. For most individuals, income from standard employment contracts, non-standard contracts, such as commissions, as well as self-employment income, are aggregated and subjected to a progressive personal income tax schedule. Before 2009, the schedule had three tax bands with increasing marginal tax rates: 19%, 30% and 40%, and a small tax-free allowance at the bottom. However, unlike many other tax systems as of 2003 the Polish tax code gave taxpayers with self-employment income a choice regarding how it is taxed. Individuals with such income are given the option of having self-employment income taxed separately according to a linear 19% tax schedule.[7] The linear schedule deprives them of several tax advantages, including the tax-free allowance, the ability to claim deductions and the ability to share tax liability with a spouse. Income from assets is taxed separately according to a linear tax of 19%. The additional optional linear tax was granted on top of a number of advantages of reporting self-employment income. For instance, as is the case in many other countries, self-employed individuals can deduct business costs from their revenue, and deduct past losses in years with positive profits. Formally, an individual cannot be self-employed if they satisfy *all three* of the following conditions: 1) those commissioning the work bear responsibility for the business activity towards third-parties 2) the place and time of work are determined by those commissioning the work and 3) business risk for the activity is borne by those commissioning the work.

[7]Self-employment income is here defined as income from unincorporated businesses which act as pass-through entities (I otherwise refer to it as “business income”). Income from incorporated businesses is taxed separately, once under the Corporation Income Tax (CIT), which is 19%, and subsequently again under a 19% personal tax which applies to capital gains.

In principle, therefore, there is considerable scope for switching to self-employment by previously employed individuals.

If a taxpayer is married, they have the option of reducing their tax liability through joint filing, a feature the Polish tax system shares with many other tax systems, such as the US system. Income is summed over the taxpayer and the spouse: if the sum falls below twice a bracket threshold, a lower marginal rate is applied to the entire income, even if one of the taxpayers individually would have crossed the bracket and had been subject to a higher marginal rate. The averaging of incomes always creates an incentive to file jointly if the taxpayer would have otherwise been in a higher bracket than the spouse, or vice versa. Single parents can similarly share tax liability with their child, who acts as if they were a non-earning spouse.

The 2009 tax reform cut the the marginal tax rate in the bottom bracket by 1% and extended this tax bracket to cover what was previously the middle tax bracket, resulting in a 12% fall in the marginal tax rate for what was previously the middle tax bracket. The top rate was also reduced from from 40% to 32%. The reform therefore 'flattened' the tax schedule, resulting in a schedule with just two tax bands of 18% and 32%. The changes in marginal tax rates and thresholds are summarized in Table 1. Crucially, since the optional linear tax schedule for self-employment income remained unchanged, the reform also had the effect of reducing the differential in tax rates between the progressive and linear schedules, especially for individuals who would have previously fallen in the second and third tax brackets in the progressive schedule.[8]

In the periods 2004-2006 and 2008-2012, the tax brackets remained frozen in nominal terms, thus implying many individuals would have experienced a transition into higher tax brackets due to inflation and secular income growth (so-called 'bracket-creep'). No significant changes to the tax base occurred in my sample period, and available deductions remained fairly constant. This implies that the definition of taxable income remained fairly stable, and thus merits focus on taxable income as the variable of interest (other studies have tended to focus on broad income, which has often been seen as the measure less sensitive to changes in definitions of the tax base). Significant real-GDP growth occurred in 2007-2008, with a slight slowdown in 2009. This is problematic to the extent that changes in growth differentially affect different regions of the income distribution. However, it will be shown that the identification strategy presented allows us to control

[8] Since the identification strategy used in this paper relies on joint filing by spouses, it is important to note that following the reform, the combinations of income for which there were positive gains from joint filing shrunk. For instance, if the original filer was in the middle tax bracket with a spouse in the lower tax bracket, following the reform there was no longer a gain from filing jointly (although there was no financial loss to doing so either). I do indeed see a fall in the proportion of individuals in the second bracket reporting with a spouse from 65.3% to 61.4%, while I do not see a similar fall for those in the third bracket. I largely abstract from this issue, and it will be seen that the population of those who continued to file jointly throughout the period continues to serve as a source of suitable treatment and control groups.

Table 1: Marginal tax rates and tax bands (in zł).

Marginal tax rate	2007	2008	2009-12
Tax-free amount (0.00)	0-3,015	0-3,091	0-3,091
0.18			3,091-85,528
0.19	3,015-43,405	3,091-44,490	
0.30	43,405-85,528	44,490-85,528	
0.32			>85,528
0.40	>85,528	>85,528	

for business cycle effects. A final confounding factor consists of the lagged effects of the 2003 reform, which introduced the option of the linear business schedule. The first year in the dataset, 2004, was the first year following the introduction of the linear tax. Until 2005 I see a continuing year-on-year increase in the proportion of individuals filing the linear tax, and thereafter the take-up rate for this tax stabilizes in 2006 and 2007. Thus, I will exclude 2005 from the analysis of extensive-margin transitions.

A possible confounding event to the 2009 reform was a cut in the level of social security contributions for the financing of the disability insurance programme which occurred in 2007-2008. This is a contribution paid proportionately on employment income, but is a fixed rate for self-employed individuals equivalent to the rate paid by those employed full-time on the minimum wage. If I treat these social security contributions as a tax, this change would have caused a net increase in the tax gains from self-employment relative to employment at the lower end of the income distribution in 2007-2008 (and the reverse upper in the income distribution). On the extensive margin, my approach is to use the self-employed subject to progressive taxation as a control group for the self-employed subject to the linear schedule, or alternatively to examine changes along a band around a tax kink. This ought to control for the social security changes, as both groups would have been similarly in 2007-2008. On the intensive margin, the main identification strategy for the effect of 2009 tax reforms relies on assignment into treatment/control based spousal income, and controls for base-year income. Again, this should control for the effects of the 2007-2008 social-security changes, which were independent of spousal income.

It is important to note that the transition from employment to self-employment and linear taxation was subjected to some restrictions. For instance, a year had to elapse between a taxpayer being employed by a company, and subsequently being hired by that same company as a business-owner. I therefore focus on two-year differences when examining transitions between employment and self-employment.[9]

[9]The question of how optimal tax policy could in principle be affected if individual-specific fixed costs could be influenced by e.g. changes in reporting requirements is left for future work.

3 Motivation

A central feature of the 2009 reform was that it altered the relative incentives to file income under the linear business schedule and the progressive (business and employment) schedules. In this section, I outline a theoretical model of reporting behaviour which describes how intensive elasticities interact with extensive-margin income-shifting responses in determining the marginal deadweight loss (DWL) of tax reforms. Based on the predictions of this model, I will arrive at the empirical statistics sufficient for the calculation of the DWL. A discussion of how these extensive margins affect optimal tax formulae is left until Section 8.

3.1 DWL under Linear Taxation

In the model, I postulate the existence of two tax bases subject to linear taxation, as well as a fixed costs of reporting self-employment income for taxpayers. I assume there exists a mass of individuals, each characterized by a parameter θ influencing the marginal cost of declaring an extra unit of taxable income of either type, and a parameter ϕ, which is the fixed cost of declaring business income instead of employment income. The cumulative distribution functions in the population are $F(\theta)$ and $G_\theta(\phi)$, with marginal densities $f(\theta)$ and $g_\theta(\phi)$, where the distribution of fixed costs is allowed to depend on θ. Individuals are assumed to have a quasi-linear utility of the form

$$u(c, l, b; \theta) = c - 1\{b = 0, l > 0\}.\psi^L(l/\theta) - 1\{b > 0, l = 0\}.(\psi^B(b/(\tilde{\omega}\theta)) - \phi) \qquad (1)$$

where c is consumption, b is the amount of business income declared, l is the amount of labour income declared, and $1\{b = 0, l > 0\}$ as well as $1\{b > 0, l = 0\}$ are dummy variables equal to 1 if, respectively, any positive employment or self-employment income are declared. The function $\psi^K(.)$, where $K \in \{L, B\}$ indicates the tax-base, is convex and implies increasing marginal costs of producing an extra unit of taxable income as taxable income increases.

It is worth mentioning possible interpretations of this specification of utility. This fixed cost ϕ can encompass administrative costs, attitudes towards risk, and preferences towards the flexibility of self-employment (or, alternatively, easier access to evasion opportunities). The parameters θ and $\tilde{\omega}\theta$ can be interpreted as the hourly wage rates for an hour worked as an employee or as a self-employed individual, respectively, in a standard model of labour supply. The cost functions $\psi^K(.)$ presented above can then be interpreted as representing the increasing cost of additional hours-of-work. However, it is easy to broaden the model above to include situations where an individual has other margins of response, however.[10] Since the function $\psi^K(.)$ is linked to the elasticity of

[10]For instance, Chetty (2009) shows that in a model where an individual chooses hours to work, with

taxable income, and since previous studies suggest that the elasticities of taxable income are different for employment and self-employment, these are allowed to be different in the above model. The quasi-linear utility implies an absence of income effects, and has been previously used by Diamond (1998).[11]

In this simplified model, individuals choose between employment and business income, but cannot report both. As has already been mentioned, a minority of around a quarter of the self-employed also report employment income. The budget constraint for each individual is just $c \leq (1-\tau_L)l$ if the individual earns employment income, and $c \leq (1-\tau_B)b$ if the individual earns business income. On the intensive margin, conditional on being in a tax base, individuals will choose to report the level of income $1 - \tau_K = \psi'^K(k/\theta)/\theta$, where $k \in \{l, b\}$ denotes the level of income declared in a tax base. The solution to this yields the reported income supply functions $l(\theta)$ and $b(\theta)$. I may additionally define the indirect utility for each tax base, excluding fixed costs, as

$$v^K(\theta) = (1 - \tau_K)k(\theta) - \psi^K(k(\theta)/\theta).$$

An individual of type θ chooses the business tax base if the gain in indirect utility relative to the employment tax base exceeds the difference in associated fixed costs. The tax base choice for the individual is therefore determined by whether or not their fixed costs exceed the following threshold:

$$\widetilde{\phi}(\theta) = v^B(\theta) - v^L(\theta)$$

Consequently, the proportion of individuals of type θ reporting in the business tax base is simply $G_\theta(\widetilde{\phi})$, and the proportion in the employment tax base is simply $1 - G_\theta(\widetilde{\phi})$.

Let us suppose that the government increases the marginal tax rate on labour income τ_L by a small amount $d\tau_L$, with no change in the marginal tax rate on business income τ_B. The reform has two effects on tax revenue. The first is a "mechanical" increase in tax revenue from employees as a result of taxpayers with wage income facing a higher tax rate. The total size of this effect is

$$dM = \int_\Theta \Big[(1 - G_\theta(\widetilde{\phi}))l(\theta)\Big]\, dF\theta \times d\tau_L \tag{2}$$

This is the projected increase in tax revenue from all taxpayers with employment income, absent any behavioural response.

increasing marginal costs of hours, but also makes a choice about not reporting a proportion of income, which is costly, the welfare analysis of tax reforms presented below carries through. The advantage of using the simple framework above is that, as will be seen, elasticity parameters map straightforwardly onto coefficients in the utility function.

[11]Preliminary analysis suggests that income effects are in fact quantitatively small in the sample of taxpayers under consideration.

The second effect of the reform is to cause a behavioural response which changes the reported income of taxpayers in the employment tax base. In the model proposed here, the behavioural response will have both an intensive-margin and an extensive-margin component. The intensive-margin response is the one traditionally emphasized in the ETI literature (see for instance Saez et al. (2012)): the change $d\tau_L$will induce a taxpayer of type θ to reduce the level of reported employment income by $\frac{\partial l(\theta)}{\partial \tau_L} d\tau_L$. This can be re-expressed as $-\frac{l(\theta)}{(1-\tau_L)}\varepsilon_L(\theta)d\tau_L$, where $\varepsilon_L(\theta) = \frac{1-\tau_L}{l(\theta)}\frac{\partial l(\theta)}{\partial(1-\tau_L)}$ is the standard elasticity of employment income for individual θ with respect to the marginal tax rate on employment income.[12] The total amount of income lost through this response is the integral of the change in tax revenue for all employed individuals: $\left(\frac{\tau_L}{1-\tau_L}\right)\int_\Theta \left[(1-G_\theta(\widetilde{\phi}))l(\theta)\varepsilon_L(\theta)\right] dF\theta \times d\tau_L$.

However, for a taxpayer of type θ, the tax reform will also increase the threshold value of fixed costs at which it becomes optimal to switch from employment to business. This will induce a proportion of individuals with fixed costs below this value to switch to the business tax base. Specifically, the threshold $\widetilde{\phi}_\theta$ will change by

$$\frac{\partial \widetilde{\phi}_\theta}{\partial \tau_L} d\tau_L = -\frac{\partial\left(v^B(\theta) - v^L(\theta)\right)}{\partial(1-\tau_L)} d\tau_L = l(\theta)d\tau_L.$$

We may note that the change in the threshold for switching is therefore proportional to the change in the total quantity of tax paid on the employment income declared by an individual. The density of individuals of type θ who are induced to switch as a consequence of the reform is given by $g_\theta\left(\widetilde{\phi}\right) \times l(\theta)d\tau_L$. For each individual in employment who switches to the business tax base, the net loss in tax revenue will be equal to $\Delta T^{L,B}(\theta) = \tau_L l(\theta) - \tau_B b(\theta)$. Hence, for type θ, the amount of revenue lost on the extensive margin is the density of individuals induced to switch, multiplied by the difference in total tax rates between the tax bases across which they are switching: $g_\theta\left(\widetilde{\phi}\right) l(\theta)\Delta T^{L,B}(\theta) \times d\tau_L$. This term is novel, and captures the fiscal externality arising out of the income-shifting occurring on the extensive margin. It is a weighted average over all types θ by the probability density of each type, $f(\theta)$, and is increasing in the density of individuals induced to switch tax bases, $g_\theta\left(\widetilde{\phi}\right) l(\theta)$, as well as the difference in the total tax burden between the two tax bases, $\Delta T^{L,B}(\theta)$.

Summing up both the intensive and extensive margin, the total change in tax revenue due to the behavioural response is equal to

$$dB = -\left(\frac{\tau_L}{1-\tau_L}\right)\bar{\varepsilon}_L \times d\tau_L - \int_\Theta \left[g_\theta\left(\widetilde{\phi}\right) l(\theta)\Delta T^{L,B}(\theta)\right] dF\theta \times d\tau_L \qquad (3)$$

where I define $\bar{\varepsilon}_L = \int_\Theta \left[(1-G_\theta(\widetilde{\phi}))l(\theta)\varepsilon_L(\theta)\right] dF\theta$ as the aggregate elasticity of employment income to the marginal tax rate on employment, weighted by the level of employ-

[12] Since there are no income effects present, this is both the compensated and uncompensated elasticity.

ment income. The term $\int_\Theta \left[g_\theta \left(\widetilde{\phi} \right) l(\theta) \Delta T^{L,B}(\theta) \right] dF\theta$ is the extensive-margin response.[13] The total change in tax revenue dR due to the tax reform is just the sum of the mechanical and behavioural responses, $dR = dM + dB$.

The two terms accounting for the behavioural response are exactly equal to the marginal deadweight burden of the increase in the tax rate. This is a well-known result of the envelope theorem: due to the optimizing behaviour of taxpayers, the behavioural response to a small tax change creates no additional welfare loss aside from the mechanical effect. In the context of the above model, this is true both of responses at the extensive margins as well as the intensive margin. Individuals induced to switch to the business tax base are at the point of indifference between the two tax bases, and so incur no additional welfare loss due to switching. This is what allows us to measure the utility loss of the tax change in monetary terms purely in terms of the mechanical effect dM. Since tax revenue collected is $dR = dM + dB$, which is smaller than the utility loss dM as a result of the reform, the difference between the two, $-dB$, represents the extra amount lost in utility over and above the revenue collected.[14]

At the optimum, absent distributional preferences, the policymaker would attempt to equalise the marginal DWL between the two tax bases. If the elasticity of taxable income is higher for self-employment than for employment (as is found by some studies, such as Kopczuk (2015) and Saez (2010)), and there were no extensive-margin responses, the DWL formulae would imply that the marginal tax rate should be lower on the business-income tax base than on the employment tax base. However, the above analysis also implies that this motive would have to be tempered by the degree of switching on the extensive-margin due to the differential in total tax rates between the two tax bases.

The above analysis demonstrates that the presence of an extensive-margin response increases the size of this additional deadweight burden relative to a model with just an intensive-margin response.[15] It also suggests which statistics are sufficient to evaluate the

[13] This term can be readily expressed in terms of an extensive-margin elasticity $\xi_\theta(\widetilde{\phi}) = \frac{\Delta T^{L,B}(\theta)}{G_\theta(\widetilde{\phi})} \frac{\partial G_\theta(\widetilde{\phi})}{\partial(\Delta T^{L,B}(\theta))}$, which is just the elasticity of the fraction in self-employment with respect to relative tax gains of self-employment. Specifically, the extensive-margin response term becomes $\int_\Theta \left[G_\theta \left(\widetilde{\phi} \right) \eta_\theta(\widetilde{\phi}) \right] dF\theta$.

[14] The same analysis can be performed for the effects of a small increase in the business marginal tax rate $d\tau_B$, with analogous terms reflecting the intensive and extensive-margin behavioural responses. The marginal deadweight burden of an increase in the business tax rate would now be increasing in the elasticity of business income $\bar{\varepsilon}_B$. Specifically, the formula showing the change in revenue due to a small increase in the business tax rate would be

$$dR = \left\{ \int_\Theta \left[G_\theta(\widetilde{\phi}) b(\theta) \right] dF\theta - \left(\tfrac{\tau_B}{1-\tau_B} \right) \bar{\varepsilon}_B + \int_\Theta \left[g_\theta \left(\widetilde{\phi} \right) b(\theta) \Delta T^{L,B}(\theta) \right] dF\theta \right\} \times d\tau_B \ .$$

Presuming the marginal tax rate on employment income is initially higher, there is a difference in that the extensive-margin behavioural response is positive, i.e. attenuates the DWL. This is because with an initially higher tax rate τ_B, switching away from self-employment towards employment generate a positive fiscal externality.

[15] Piketty and Saez (2012) argue that extensive-margin responses do not alter the standard formulae for DWL. This is true in their model if the extensive-margin responses are between reporting some taxable

formula for DWL, and which we can in principle estimate. These would be the size of the aggregate tax elasticity of the employment tax base, corresponding to the aggregate intensive-margin response to the reform, as well as an estimate of the total mass of taxable income lost as a result of switches on the extensive margin. In the online appendix, I use an equation very similar to (3) to approximate the DWL to a tax reform with multiple tax brackets such as the 2009 Polish tax reform.

As has been pointed out by Chetty (2009), the elasticities of taxable income are sufficient for the welfare evaluation of tax reforms only if a number of assumptions hold. For instance, there has to be a lack transfers as a result of evasion to other individuals by the taxpayer, or a lack of optimisation frictions. In the absence of such frictions, and with only a private cost of hiding a share of income, a simple extension of the baseline model in Appendix H discusses how in the presence of such evasion responses elasticities of taxable income remain a sufficient statistic.[16]

4 Data

The dataset used for the empirical work is a proprietary dataset obtained from the Polish Ministry of Finance. It comprises a balanced panel of about a million Polish taxpayers over the years 2004 to 2012. This is a random sample selected out of the population of all taxpayers who reported any taxable income at least once in this period, and as a result is subject to attrition and replacement. The dataset contains information on individuals and spouses filing according to the progressive schedule, as well as self-employment individuals who chose to file self-employment income according to the linear schedule. In the latter case, individuals are legally prevented from filing jointly with a spouse and, as a result, I am unable to predict whether or not they have a spouse in practice. As with most tax return data, demographic information is limited, and contains only the taxpayer's age and gender. The number of children can be inferred only to the extent that the taxpayer claims deductions only available to those with children. Any deductions claimed by the individuals can be observed, as can be their county of residence.

The dataset allows us to link individuals who report both linear self-employment income and some employment income, and also to observe whether they report either self-employment or employment income, or both, if they report under the progressive schedule. Since less than a third of business owners report both employment and business

income and reporting none. However, the main empirical studies used for calculating DWL of taxation (Gruber and Saez (2002), and Kopczuk (2005), being notable examples) tend to employ a panel method, such that only individuals who continue to report a positive level of taxable income in the personal income tax base. Individuals who switch to other tax bases, such as the corporate tax base, are excluded from the analysis. The model in this paper implies that this should exaggerate the size of the income loss, which would be attenuated since individuals who switch are continued to be taxed at the rate τ_B.

[16]This holds if evasion involves hiding income completely from the tax authorities or reporting it as a part of business operating costs.

income, for the empirical section I consider these individuals to be members of the self-employment tax base, on the grounds that they bear the fixed cost of participating in this tax base. An extension of the empirical study to estimate intensive-margin cross-elasticities, which would be relevant if I included individuals who report positive levels of both types of income, is left for future study.

For the purpose of the analysis in this paper, I focus only on the impact of tax reforms on tax filers. In other words, I do not make spouses a subject of study. However, as will be seen, information on spouses contained in tax returns will play a crucial role in the identification of the effects of the Polish tax reforms.

5 Intensive-Margin Responses

Section 3 explained how intensive-margin elasticities and extensive-margin parameters interact in formulae for DWL. In this section, I employ the variation which arose as part of the Polish 2009 reforms to estimate responses on the intensive-margin and evaluate the relevant intensive-margin elasticities.

5.1 Empirical Strategy

The primary outcome of interest is the log-change in taxable income as a function of changes in marginal tax rates. Following much of the existing studies, the panel regression specifications for individuals reporting employment income and business income are:

$$\Delta \log z_{it}^k = \varepsilon^k \Delta \log (1 - \tau_{it}^k) + f_t^k(z_{it}^k) + \Delta \xi_{it}^k \tag{4}$$

where $k \in \{L, B\}$ indicates the tax base under consideration. The variable z_{it}^k is reported taxable income in tax base k, $1 - \tau_{it}^k$ is the net-of-marginal tax rate in tax base k. Here, $\Delta \log x_{it}^k = \log (x_{i,t+s}^k / x_{it}^k)$ is the log-change in each of the respective variables between base-year t and year $t + s$. Consequently, ε^k evaluates the percentage change in taxable income as a result of a 1% change in the net-of-tax-rate (i.e. the elasticity of taxable income).[17] The difference s will usually be taken to be one year, although longer differences will also be looked at to ensure that results are not driven by timing responses.

The differencing operation means that all time-constant variables are eliminated. However, mean-reversion continues to be a key concern when using panel data to estimate ETIs, and could in principle result in the error term being correlated with the instrument. In line with much of the literature, I therefore include a control for base-

[17] Income effects are here excluded from the theoretical and empirical specifications for reasons of tractability. The literature tends to show that these are in fact of low magnitude, and preliminary results suggests these do not substantially alter the main empirical results of the paper. A detailed study of their role in the present context is left for future study.

year income in the specification, $f_t^K(z_{it}^K)$ (see, for instance, Gruber and Saez (2002) and Kopczuk (2005)). For employment income, this will be a flexible piecewise linear function of log of base-year taxable income. For business income, due to a smaller sample size, I will include a linear control in the log of base-year income.

Choice of instruments. As has been widely recognized in the literature, the tax variables are endogenous due to the progressivity of the tax schedule, individuals who experience positive income growth will face higher marginal tax rates, resulting in biased estimates of the elasticity parameter. The variable for the change in the net-of-marginal tax rate therefore needs to be instrumented. The instrument for $\Delta \log(1-\tau_{it}^B)$ proposed here is a dummy equal to 1 if the income of the spouse is sufficiently high at time t for the couple to enter at least the second tax bracket under joint reporting, and 0 otherwise. The identification of the elasticities in this case comes from comparing the changes in taxable income for individuals who experienced the higher falls in the marginal tax rate associated with the middle and top brackets (for whom the instrument is equal to 1) with the changes in taxable income of those in the bottom bracket who experienced a negligible fall (for whom the instrument is equal to 0). For any level of base-year taxable income z_{it}^k in the range under consideration, it is possible to find both an individual who is 'treated' and an individual who falls into the 'control' group.[18] For the intensive-margin estimates, all income variables are deflated using average taxable income growth in the population, thus accounting for changes in tax rates due to 'bracket creep'.

The identification strategy can be helpfully illustrated diagrammatically. This is done in Figure 1, which compares the tax schedules faced by individuals who have a spouse declaring the same level of income as they are (or, alternatively, are single – the tax schedules in this case are the same, if I presume tax liability is divided equally between the couple), and individuals with a zero-income spouse. As can be seen, all individuals in the middle tax bracket (43,405zł-85,528zł), for instance, who have a spouse at the same level of income experience a cut in the marginal tax rate (represented by the slope of the tax schedule), as well as a cut in the total tax liability, represented by the vertical distance between the new and the old schedules. On the other hand, individuals who have the same basic income but who have a spouse with zero income experience a negligible cut in marginal taxes in the same income range, and also a negligible cut in total tax liability. In this way, the year-on-year change in income for individuals with a low-income spouse in this income range could plausibly provide a counterfactual for individuals in the same income range, but who have higher-earning spouses. This is precisely the logic of the instrumental variable strategy presented above.

This identification strategy is importantly different from that used in Gruber and

[18] This is possible for the income range $z_{it}^k \in [0, 2 \times \bar{z}^{(2)}]$, where $\bar{z}^{(2)}$ is the threshold for entering the middle bracket. Consequently, the estimates of ETIs are limited to this income range. However, this range spans approximately 95% of Polish taxpayers.

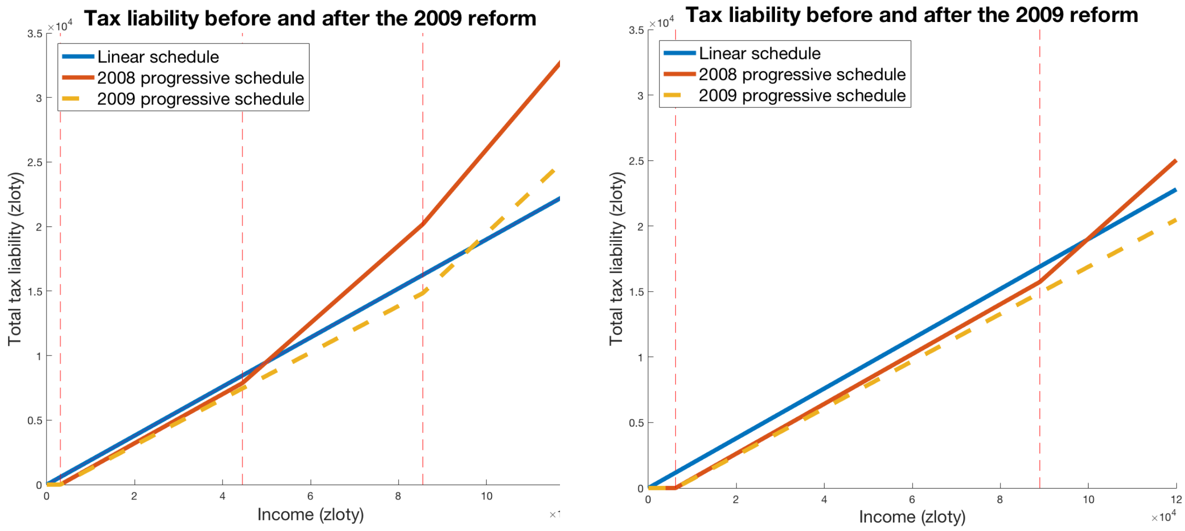


Figure 1: Effects of reform for individuals reporting who are: (i) single, or alternatively have a spouse declaring the same income level as the individual (left panel) and (ii) a spouse declaring zero income (right panel). The blue lines represent the linear self-employment schedule, which remains unchanged before and after the reform. The red line represents the progressive schedule (for employment or self-employment) before the reform, while the yellow line represents the progressive schedule (for employment or self-employment) after the reform. The x-axis represents nominal income, while the y-axis represents total nominal tax liability.

Saez (2002). Their approach would instead would use instruments based on membership of the highest two income brackets in the base year. In this alternative strategy, the counterfactual levels of income growth are provided from non-reform years – in particular, these are captured by the spline terms in the above regression framework. In terms of Figure 2, this would involve estimating a pattern of income growth for each income level on the x-axis using a flexible spline, and subsequently using this relationship and observed income growth in the group of individuals with income just below 43,405zł to provide the counterfactual for what the income change would have been above 43,405zł, absent reform. As will be shown, this strategy is not robust to changes in the pattern of income growth year-on-year, and may significantly bias estimates where stable income growth patterns are not observed.

Identifying assumptions. The crucial identifying assumption here is that the spouse's income being low enough for the taxpayer to fall into the lowest tax bracket is independent of unobserved variables driving the *change* in the taxpayer's taxable income, conditional on covariates.[19] This amounts to a common trends assumption which should hold in years outside of the reform, and it is possible to verify it empirically. To this end, Figure 2 plots the average level of income growth for the treatment and control groups of those

[19]Importantly, this does not require the assumption that there is not correlation in *levels* of income between spouses. Indeed, there is considerable correlation between the incomes of spouses in the dataset. Theoretically, this could be explained by a model in which there is separability between the taxable income of the primary taxpayer and their spouse.

reporting employment income for years preceding and following the reform, as well as the year of the reform itself. Both groups appear to show very similar patterns of income dynamics in the years outside of the reform for any given income level. In the year of the reform, however, the group pushed into the second and third brackets seems to have a discernibly higher rate of taxable income growth, and this appears to be relatively stable across the income distribution. To illustrate the 'first-stage' of this approach, Figure 3 plots changes in the net-of-tax rate among the two groups. As expected, net-of-tax rates change only slightly outside of the reform year for both groups (any actual changes arise due to moving between tax brackets among the two groups), while in the reform year the treatment group with high income spouses experiences a large increase in net-of-tax rates, while the control group with low income spouse groups experiences a much smaller change. This variation lies at the heart of using membership of either group as an instrument for the net-of-tax rate change.

Thus, it would be appear that the identifying assumption holds relatively well. Indeed, the graphs suggest that the methodology used in other ETI studies (Gruber and Saez 2002; Kopczuk 2005), which assumes a stable gradient of income growth from year to year outside of reform, would be questionable in the present case. For instance, the slope of the gradient of income growth in 2008 appears to be much steeper than that in 2010. Thus, the method would likely have produced a high elasticity estimate using the 2010 placebo, with 2008 as the baseline year. The method here allows us to sidestep this problem, as the patterns of income growth for the treatment and control groups are similar within each non-reform year, even if they differ from year to year.

Sample selection. Since mean-reversion is particularly acute at the very bottom of the distribution, I will exclude individuals with a base-year income below 20,000zł in 2004 terms from the sample of the employed, and a base-year income below 5,000zł in 2004 terms from the sample of business owners. As in Gruber and Saez (2002), I censor changes in log income at ± 7. This means dropping individuals who experienced a thousand-fold rise or fall in taxable income from year to year. A person will feature in the sample if they are observed in both the base year (the year in which the starting level of income is measured) and the target year (the end level of income, after a suitable number of years have elapsed). Thus, in principle analyzing differences over one year between 2008 and 2009 will involve different sample of taxpayers than analyzing differences over two years between 2007 and 2009, if some individuals only started reporting in 2008. However, restricting the sample to be the same does not significantly alter the size of the estimates.

For both tax bases, only individuals who report income in that tax base in all years only are included in the sample – thus, all individuals who report some amount of both types of income at any point are excluded. The same identification strategy is used both for the employment and business samples. This naturally relies on the sample of business owners who chose not to file according to the linear schedule.[20] Hence, I make the assumption that the selection of business owners into the linear tax base was independent of their elasticities of taxable income, and was driven by, for instance, preferences for deductions.

5.2 Results

The baseline results for the sample of individuals reporting employment income are shown in Table 2. The estimated elasticity of employment income ε^L is 0.218, and is highly statistically significant. All of the placebo estimates of the elasticity are statistically insignificant, supporting the validity of the exclusion restriction – belonging to the treatment group only appears to affect income growth only in the year where a significant change in marginal tax rates occurred (corresponding to Figures 2 and 3).

The elasticity estimates based on a 2-year difference between 2007-2009 is larger at 0.303. Over a longer horizon such as this, individuals who, together with their spouses, are only just assigned to the treatment or control groups based on their joint income in the base year, may be more likely to move between the treatment and control groups over the 2-year time-frame. For the 2-year differences, I therefore I exclude individuals who are within 10,000zł of being classified differently in the base year, based on their joint income with their spouse in the base year. These estimates together seem to point that the long-run response may have been even greater than that suggested by the immediate response and, crucially, I see no evidence of a timing response (with taxpayers shifting income between 2008 and 2009), at least in the employment sample. The placebo tests for 2-year differences perform slightly wore than for the 1-year difference estimates: for 2010-2012 they are statistically significant at the 5% level, although they are not statistically significant for the other two placebo estimates (2004-2006 and 2005-2007).

The identification strategy used here, together with the large size of the sample of employed individuals, allows us to estimate the ETI for small subsets of the distribution of employment income. The results are presented in Table 4. Above the 20,000 zł band, the ETI estimates show surprising stability, and all fall between 0.20 and 0.25.[21]

It has already been mentioned that the methodology used by Gruber and Saez (2002), which has often been replicated in other studies, relies on assuming a stable gradient of income growth from year to year. The essence of this methodology is to use non-reform (usually pre-reform) years to predict what the gradient income growth would have been in the reform year, absent tax reform. Thus, if there is an unusually flat downward gradient

[20] In a follow-up project using Polish adminsitrative data, I am pursuing the question of whether there was selection of higher-elasticity individuals into the linear tax base among the self-employed. As it stands, the estimates here and those in Kopczuk (2005) suggest a lower and upper bound on the elasticity of the self-employed population as a whole.

[21] It is worth noting that this spans most of the employment earnings distribution in the economy, with median earnings in Poland in 2008 being 31,680 zł.

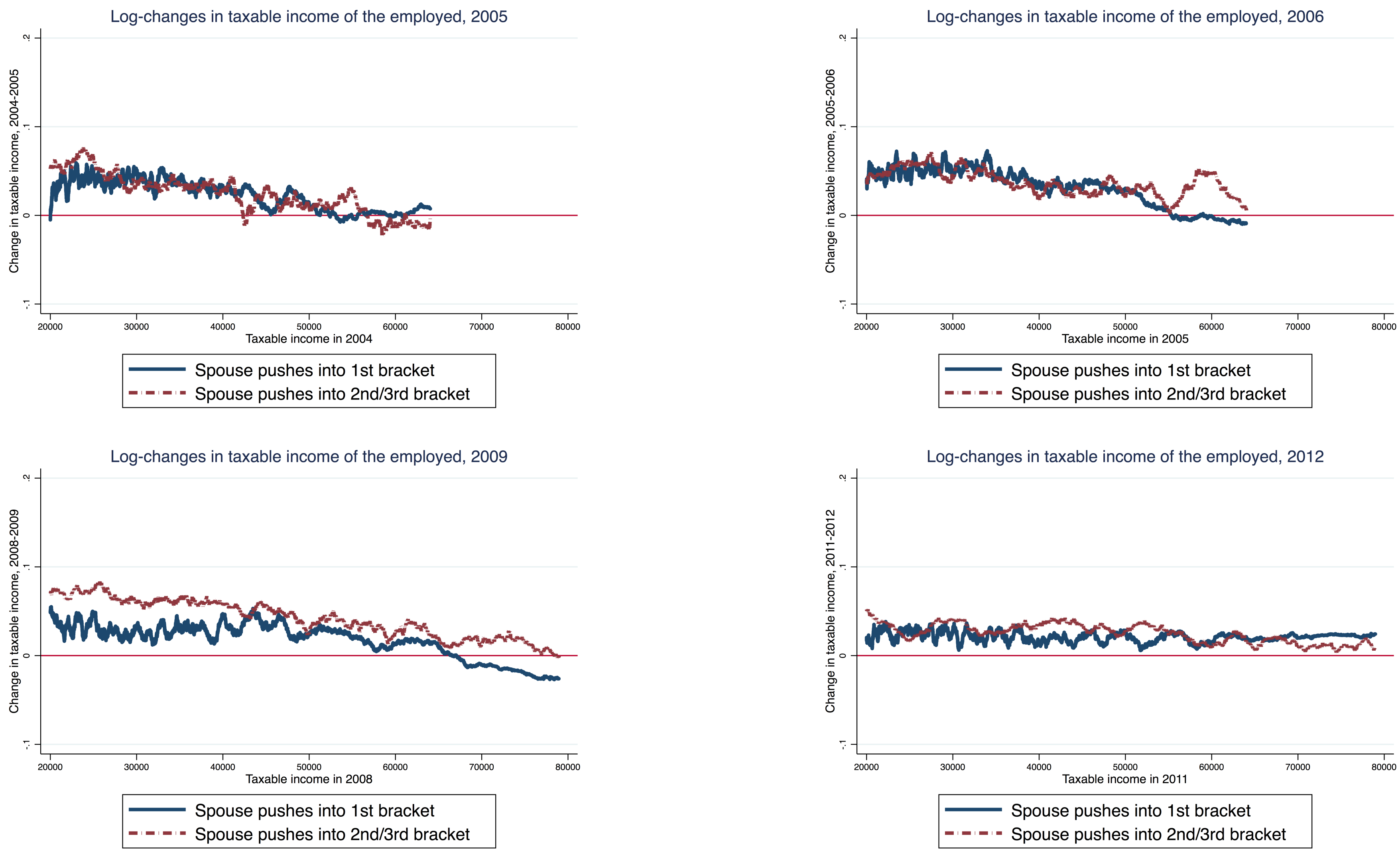


Figure 2: Taxable income growth rates for individuals whose joint filing with a spouse pushes them above the first tax kink (red line) and individuals who remain in the bottom tax bracket (blue line). The x-axis represents income in the base year, while the y-axis represents the log-change in taxable income.

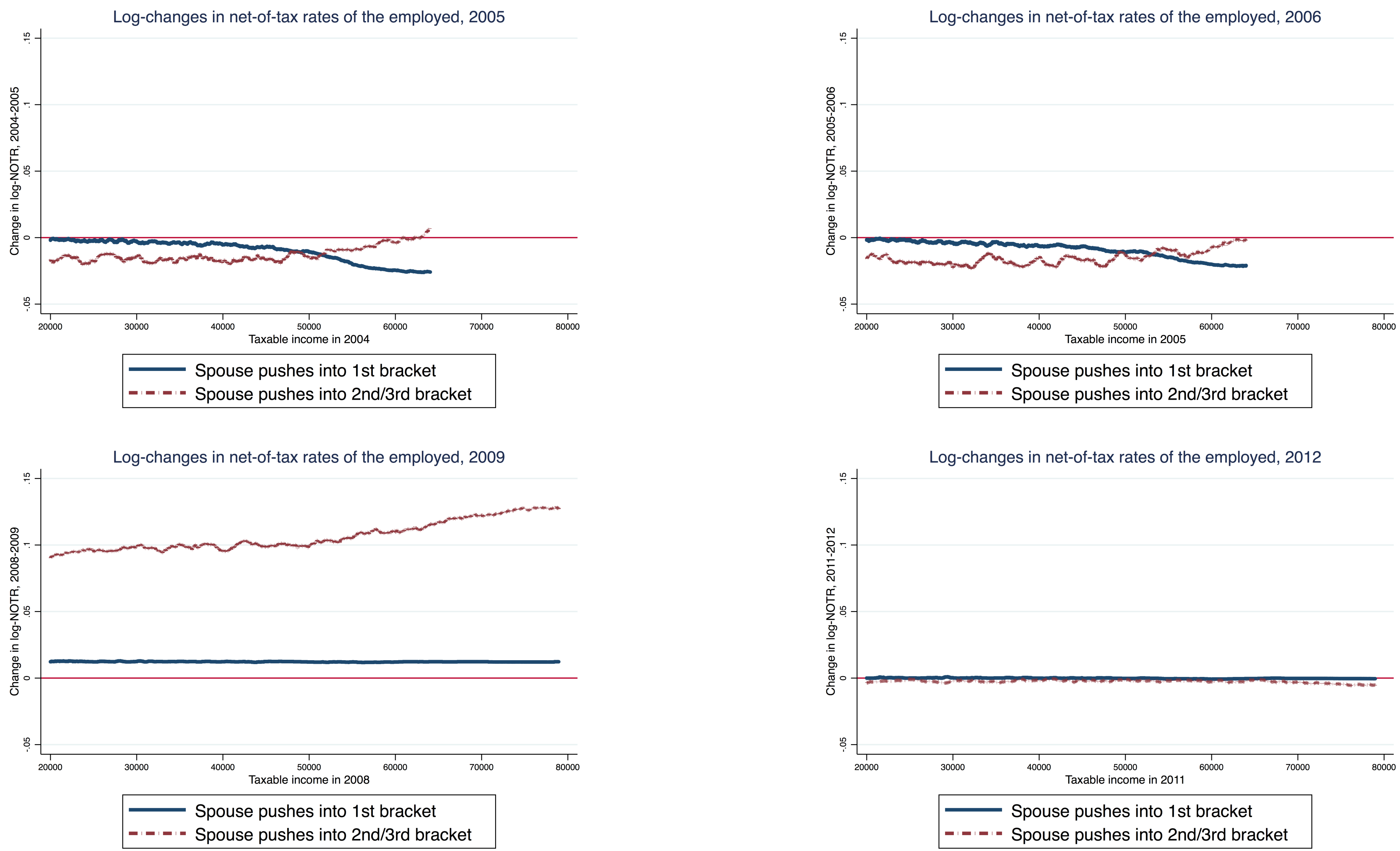


Figure 3: The first stage: (log) net-of-tax rate changes for individuals whose joint filing with a spouse pushes them above the first tax kink (red line) and individuals who remain in the bottom tax bracket (blue line). The x-axis represents income in the base year, while the y-axis represents the log-change in one minus the marginal tax rate (the net-of-tax rate).

of income growth in the year of the reform relative to previous years, which may reflect low levels of mean-reversion in any given year, the Gruber-Saez methodology would assign this reduction in mean-reversion to the effects of the reform. Thus, the approach would over-estimate the responsiveness of taxable income to a tax cut.

The results of applying the methodology used by Gruber and Saez (2002) to the present sample is illustrated in the final column of Table 2. It can be seen that the result is an elasticity of taxable income which is roughly 30% greater than the elasticity estimate based on the baseline methodology used here. Given the variation in the profile of income growth already noted in Figure 2, however, and the nature of the Gruber-Saez methodology, this is not surprising. Specifically, an estimate of the gradient of income growth for the years 2004-2005, 2005-2006 etc., is used here to predict the non-reform pattern of income growth in 2008-2009, and this is I have seen to be steeper in pre-reform years. As a consequence, this would have led to an under-prediction of the counterfactual and the higher estimates. In summary, one of the conditions for employment of the Gruber-Saez methodology, namely stability of income dynamics, appears to have been violated in the Polish case. The final column in Table 2 shows results of applying the more recent approach of Weber (2014). Here, the individual's income 2 periods before the base-year is used to construct a predicted change in net-of-tax rates in the reform year; at the same time, the spline in base-year income is combined with a spline in income 2 periods before the base-year. The purpose of both changes is to eliminate any remaining correlation between the instrument and the error term in the earnings growth process, which is assumed to have a limited memory form (e.g. MA(1) or MA(2)). As can be seen, it does align more closely with the baseline results in the first column. Nonetheless, they are still somewhat larger than my baseline result and, what is more, once we move to 2-year differences the discrepancy becomes very large. That discrepancies persist is not surprising, however, as this alternative method also relies on estimates in other years providing the counterfactual pattern for income growth across the income distribution.

The results for the sample of business owners is presented in Table 3. It can be seen that although the sample of individuals is much smaller than for the employment sample, the estimated business elasticity is still statistically significant at the 5% level. At 0.657, it is estimated at three times the value of the baseline estimate for the employment ETI. There is more variation in the magnitude of the placebo estimates from year to year than in the employment sample, which could be expected given the smaller sample size. However, these never produce an ETI estimate significantly different from zero, which is what we would expect if the identification assumption holds. Estimating the effects of the reform over a two-year difference between 2007 and 2009 increases the size of the estimate to 1.701, with an admittedly large 95% confidence interval between 0.55 and 2.84. While the estimate is greater than the elasticity of 1.099 found by Kopczuk (2015) for Polish business owners, the 95% confidence interval comfortably encompasses his estimates.

Table 2: Own-elasticity estimates for employment sample.

Years	2008-09	2005-06 (Placebo)	2006-07 (Placebo)	2010-11 (Placebo)	2008-09 (Gruber-Saez)	2008-09 (Weber)
A. 1-year differences						
ε_L	0.218	-0.041	0.003	-0.008	0.338	0.286
	(0.025)	(0.033)	(0.037)	(0.098)	(0.031)	(0.069)
Number of individuals	100,248	69,461	78,911	100,648	88,349	66,362
Years	2007-09	2004-06 (Placebo)	2005-07 (Placebo)	2010-12 (Placebo)	2007-2009 (Gruber-Saez)	2007-09 (Weber)
B. 2-year differences						
ε_L	0.303	0.002	-.0076	0.064	0.551	0.748
	(0.036)	(0.058)	(0.037)	(0.029)	(0.059)	(0.240)
Number of individuals	65,139	49,280	61,618	62,237	70,152	43,764

Notes: All regressions include a 10-piece piecewise linear spline in the log of base-year income as a control. Instruments are as described in the text. The placebo estimates use the difference in net-of-tax rates from 2008-2009 for calculating the size of the effect in the placebo years (2005-06, 2006-07 and 2010-11). The sample consists of individuals who reported only employment income for the years spanned by each estimation, with all individuals who earned below 20,000zł in 2004 terms or above twice the first tax threshold minus 20,000zł.

Overall, these results seem to confirm that the ETI for business owners is likely to be significantly larger than that for employees.[22] Placebo estimates for 2-year differences are never statistically significantly significant for the business sample (although they are statistically significant for the 2010-12 difference for the employed), and in magnitude much closer to zero than the reform-year estimates.

The baseline specifications for 1 and 2-year differences for both workers and the self-employed are presented in Table 4. Column (2) uses a log-linear form to control for base-year income in the employment sample, which does not alter the result substantially. The employment sample is also highly robust to reducing changing the range of base-year income spanned in the sample (column (3)), and requiring the Treatment and Control groups to be further away from the threshold of moving to a different tax band (column (4)). The self-employment elasticity is more sensitive to changes along the latter two dimensions, perhaps unsurprisingly given the much smaller sample.

In Table 1 in the online appendix, I present estimation results with revenues and costs for business owners as the dependent variable. Arguably, since there is less incentive for business owners to underreport costs than it is for them to underreport revenue, a pure change in revenue could be suggestive of a reporting response. Here, however, I observe economically significant positive coefficients for both revenue and costs, although the coefficients are not statistically significant. The standard errors are similar to those for the coefficient for net income, so the loss of statistical significance relative to net income

[22] Kopczuk's estimates are based on a 2003 reform in Poland which introduced the option of the flat tax for business owners, using a very similar dataset to the present one. The estimate is from a specification involving three-year differences, and includes income effects.

Table 3: Own-elasticity estimates for the business sample.

Years	2008-09	2005-06 (Placebo)	2006-07 (Placebo)	2010-11 (Placebo)
A. 1-year differences				
ε_B	0.657	0.192	0.036	-0.069
	(0.287)	(0.564)	(0.352)	(0.201)
Number of observations	6,856	7,033	6,730	7,831
Years	2007-09	2004-06 (Placebo)	2005-07 (Placebo)	2010-12 (Placebo)
B. 2-year differences				
ε_B	1.701	0.251	0.249	0.018
	(0.586)	(0.779)	(0.529)	(0.273)
Number of observations	6,392	6,679	6,445	7,423

Notes: The sample consists of individuals who reported only business income, with all individuals who earned above 10,000zł in 2004 terms, or above twice the first tax threshold minus 5,000zł. The log of base-year income is used as a control in all regressions.

seems to arise due to the estimation coefficients being lower in magnitude. If interpreted as evidence of a response in costs, these results are consistent with at least some of the increases in self-employment income being driven by a rise in business activity. On the other hand, since the response of revenue appears to be smaller than that of overall income, self-employed may have responded to the reform by reporting higher margins, which would suggest that the response cannot stem from real activity alone.[23]

6 Extensive-Margin Responses

Having estimated the intensive-margin responses to the 2009 tax reform, I now turn to estimating the extensive-margin parameters identified in Section 3. The relevant dependent variable for estimating the size of the extensive-margin response to reform according to the model in Section 3 is the change in the probability of reporting any income in a given tax base. Based on a linear approximation around a starting threshold value of fixed costs $\widetilde{\phi}_\theta^{Old}$, we know that the change in the proportion reporting business income as a result of a tax reform is

$$\triangle G\left(\widetilde{\phi}_\theta\right)|_{\widetilde{\phi}_\theta=\widetilde{\phi}_\theta^{Old}}= G\left(\widetilde{\phi}_\theta^{New}\right) - G\left(\widetilde{\phi}_\theta^{Old}\right) \approx g\left(\widetilde{\phi}_\theta^{Old}\right)\left(\widetilde{\phi}_\theta^{New} - \widetilde{\phi}_\theta^{Old}\right)$$

[23]The same analysis cannot be conducted for employees, as there is a statutorily-set notional amount of costs a taxpayer is allowed to report in each year. In general, since employment income is subject to double reporting to tax authorities - from the individual and the employee - it could be hypothesized that the reporting response of self-employment comprises an upper bound for the degree of reporting responses for employment.

Table 4: Elasticity estimates for different parts of the income distribution.

	Income interval (zł)				
Employment	20k-25k	25k-30k	30k-35k	35k-40k	40k-45k
2008-09	0.218	0.207	0.252	0.251	0.207
	(0.044)	(0.040)	(0.068)	(0.084)	(0.168)
Self-employment	Bottom 50%			Top 50%	
2008-09	0.071			0.845	
	(0.670)			(0.293)	

Notes: The employment-sample regressions include a 10-piece linear spline in log base income, while the self-employment sample regressions include log base income. The self-employment sample is divided around the median level of income in the sample in the base year, and the elasticity estimates are given for the bottom 50% and the top 50%. The median level of income is 24,912zł in 2008 (in 2008 values).

From our model, I know that a local change in tax policy affects utility only through the direct effect on the change in the tax differential between tax bases. Hence $\widetilde{\phi}_\theta^{New} - \widetilde{\phi}_\theta^{Old} = d(T_\theta^L - T_\theta^B)$, i.e. the size of the change in relative taxation for type θ. An empirical analogue of this proposed here is the linear probability model

$$\triangle Pr(base_{it} = L) = \alpha \triangle \left[T_t^L(z_{it}^L) - T_t^B(z_{it}^B)\right] + \varsigma_{it} \tag{5}$$

where α is a coefficient corresponding to an estimate of the density $g\left(\widetilde{\phi}_\theta\right)$ around its original value, and ς_{it} is an innovation term capturing other factors which may induce an individual to change tax bases from year to year.

It is important to note how precisely the term α relates to the term $g\left(\widetilde{\phi}_\theta\right)$, which features in the formulae present in Section 3. In principle, any estimate of changes in the share of individuals reporting in the employment tax base is a *local* estimate, around the density of fixed costs at which individuals are indifferent. For the purpose of calculating the DWL of the 2009 reform, the change in the share of individuals reporting within the employment tax base is precisely what is required. The main approximation used here will be that, as with the intensive-margin elasticities, the parameter is constant regardless of the level of income declared by any individual. However, for purposes of calculating an optimal tax schedule, a departure from the local estimated density may be necessary. Ib Section 8, the precise parametric assumptions about this are made explicit.

Finding the simple change in the fraction of individuals reporting business income around the time of the 2009 reform may of course be inappropriate as an estimate of the effect of the reform, however, as the term ς_{it} may contain other factors driving switching which are correlated with change $\triangle \left[T_t^L(z_{it}^L) - T_t^B(z_{it}^B)\right]$. For instance, an individual who incurs a negative income shock in the employment tax base may face a lower tax liability, but may also be more likely to switch to self-employment. Instead, for identification of the effect, I focus on *transitions* occurring between the employment and business tax bases for which suitable treatment and control groups may be constructed. If I assume that an individual only has a choice between business and employment income, and there is no entry or exit into reporting any income at all, the following describes the relationship between the change in the fraction of employment income between t and $t+s$ and transitions between bases

$$\begin{aligned} \triangle Pr(base_{it} &= L) = Pr(trans_{it+s}^{B\to L}) \times Pr(base_{it} = B) - \\ &- Pr(trans_{it+s}^{L\to B}) \times Pr(base_{it} = L). \end{aligned} \tag{6}$$

where $base_{it} = K$ means belonging to tax base K at time t, and $trans_{it+s}^{B\to L}$ refers to a transition from the business tax base to employment. Accordingly, I will separately estimate the two components $Pr(trans_{it+s}^{B\to L})$ and $Pr(trans_{it+s}^{L\to B})$ and use them to back out the total change in reporting probability $\triangle Pr(base_{it} = L)$.

6.1 Empirical Strategy

In this section, I estimate an equation of the form (5) for both $Pr(trans_{it+s}^{B\to L})$ and $Pr(trans_{it+s}^{L\to B})$. If the self-employment tax schedule does not change significantly, the change in relative taxation term $\triangle[T_t^L(z_{it}^L) - T_t^B(z_{it}^B)]$ will be dominated by $\triangle T_t^L(z_{it}^L)$. This is true in the context of the 2009 reform, for instance if I focus on individuals who file under the linear schedule (and did not experience a change in the schedule at all) or individuals filing under the progressive schedule with a low-income spouse, thereby avoiding falling into a higher tax bracket pre-reform (and experiencing a 1% fall in the marginal tax rate in both employment and self-employment schedules).[24] Specifically, I estimate linear probability models

$$Pr(trans_{it+s}^{L\to B}) = \alpha_1 \left[T_{t+s}^L(z_{it+s}^{L,P}) - T_t^L(z_{it}^L)\right] + \gamma_1 \mathbf{1}\{t \geq 2009\} + \beta_1 \mathbf{1}\{i \in T\} + \varsigma_{it}^1$$

$$Pr(trans_{it+s}^{B\to L}) = \alpha_2 \left[T_{t+s}^L(z_{it+s}^{L,P}) - T_t^L(z_{it}^{L,P})\right] + \gamma_2 \mathbf{1}\{t \geq 2009\} + \beta_2 \mathbf{1}\{i \in T\} + \varsigma_{it}^2$$

where $z_{it+s}^{L,P}$ is predicted employment income in year $t+s$, $\mathbf{1}\{t \geq 2009\}$ is a dummy for a post-reform year and $\mathbf{1}\{i \in T\}$ is a dummy for belonging to a suitably defined treatment group. The difference s used here will be two years. This is to allow for individuals who

[24] The component $\triangle T_t^B(z_{it}^B)$ is accounted by predicting the change in income and applying the predicted (unchanged) tax schedule to this new level of income. The predicted income will be inflated by average income growth in the intervening period, as in Gruber and Saez (2002). Based on the results of the Heckman selection model presented in Section D of the online appendix, I do not expect there to be significant changes in the level of declared taxable income on switching between tax bases. Hence, for simplicity, I assume that the income declared on switching is the same in both tax bases, after having been inflated.

transition gradually between tax bases, and to take into account the one-year grace period required by Polish tax authorities between being employed and providing services to the former employer. A transition from employment to the business tax base is considered to have occurred if an individual has begun reporting some level of business income, where none had been reported previously. Symmetrically, a transition from the business tax base to the employment tax base is assumed to have occurred if an individual ceases to report any business income in the next period, but reports some employment income.[25]

Employment to business transitions. As has already been mentioned, the key variable of interest – the predicted change in the total tax liability in the employment tax base – is potentially endogenous and must therefore be instrumented. For the employment-to-business transitions, the identification strategy pursued here is analogous to the 'bracket creep' methodology of Saez (2003). In particular, I restrict attention to a band of income around the first kink in the tax-schedule pre-reform. The instrument for $T^L_{t+s}(z^{L,P}_{it+s}) - T^L_t(z^L_{it})$ is then a dummy for having income predicted to fall above the threshold in year $t+s$. This is equivalent to assigning such individuals to a treatment group, while individuals whose income is predicted to fall below the threshold are assigned to the control group. The intuition behind this approach is that individuals below the cutoff are less likely to be significantly affected by the 2009 reform since their predicted income will not cross the first tax threshold beyond which there is a large change in total taxation. They are therefore unlikely to experience a rise in disposable income which could change the relative tax advantages of switching to the linear business schedule. However, they are assumed to be close enough in unobserved characteristics to the treated group that they constitute a suitable control group.[26]

To implement the strategy, the instrument for $T^L_{t+s}(z^{L,P}_{it+s}) - T^L_t(z^L_{it})$ is the interaction between the post-reform and treatment dummy $\mathbf{1}\{t \geq 2009\} \times \mathbf{1}\{i \in T\}$. It can be easily shown that this results in a Wald estimator of the form

$$\alpha = \frac{\left(\mathbb{E}[trans^{L\to B}_{it+s}|T] - \mathbb{E}[trans^{L\to B}_{it}|T]\right) - \left(\mathbb{E}[trans^{L\to B}_{it+s}|C] - \mathbb{E}[trans^{L\to B}_{it}|C]\right)}{\begin{array}{c}\Big(\mathbb{E}[T(z^{L,P}_{it+s}) - T^L_t(z^L_{it})|T] - \mathbb{E}[T^L_t(z^{L,P}_{it}) - T^L_t(z^L_{it-s})|T]\Big) - \\ - \Big(\mathbb{E}[T(z^{L,P}_{it+s}) - T^L_t(z^L_{it})|C] - \mathbb{E}[T^L_t(z^{L,P}_{it}) - T^L_t(z^L_{it-s})|C]\Big)\end{array}}.$$

The numerator compares the change in the rate of transitions in the treatment group to the change in the control group. The denominator, on the other hand, compares the change in the predicted levels of taxation between the two groups.

The identifying assumption behind the estimate is that, absent reform, the difference

[25] Thus, such an individual may continue to receive some employment income. The intention here is that such an individual must still bear the fixed cost of undertaking business activity.

[26] The individuals in the treatment group, i.e. above the tax kink, are very likely to experience a change in the marginal tax rate. Since the bands around the tax kink are quite wide, however, many of those individuals are also likely to experience a non-trivial change in the total tax rate and, consequently, a non-trivial change in the relative tax difference between employment and self-employment.

in rates of transition between individuals immediately below and above the cutoff is stable. This assumption can be verified by examining the relative differences in transition rates for the treatment and control groups in years which did not involve tax reform. These patterns of transition over a two-year lag are demonstrated in the online appendix. In the post-reform years (2010 and 2012), the patterns of transition are fairly stable around the first tax kink (although there appears to be a small surge in transitions in 2010 mid-way between the first and second tax kinks). Between 2008 and 2010, around the year of the introduction of the reform, however, I observe a significant fall in the level of transitions around the first kink, and this fall appears to persist between 2010 and 2012.

Business to employment transitions. For transitions from business to employment, the instrument for an exogenous change in the difference in taxation between the tax bases is filing under a linear schedule, when compared to the population of individuals filing under a progressive schedule but with a sufficiently low-income spouse to fall into the bottom tax bracket. The intuition is that for those filing business income under a linear schedule, the differential between employment and business tax levels changed more as a result of the 2009 reform than for those who were already filing business income under a progressive schedule (indeed, with a low-income spouse, there occurred only a 1% change in the marginal tax rate). The control group here consists of individual filing self-employment income in the base year in the first tax bracket, such as that there is almost no change in relative incentives to file as self-employed for these individuals in the years 2008-2009. On the other hand, if at least some of the individuals filing self-employment income under the linear schedule have high-income spouses (although we cannot observe this due to data availability), or if they are single but would otherwise fall in the higher tax bracket in the base year under employment income, there will be at least some individuals in this treatment group for whom the relative gains from self-employment fell due to the flatting of the progressive schedule in 2009. Thus, the relative tax incentives to file self-employment income would have changed significantly only for individuals in the treatment, and not the control, group. In Figure 6.1, it is shown graphically that the parallel trends assumption between two groups seems to hold in the pre-reform years 2006-2008 as far as transitions are concerned.

6.2 Results

The results for the extensive-margin regressions are presented in Table 5. Panel A presents the estimates for the employment-to-business transitions, while panel B presents estimates for transitions from business to employment.

Employment to self-employment transitions. The first line reports estimates based on comparing the years 2008 and 2010. The baseline estimate of $\alpha\times$ 10,000 zł at 0.0196 implies that a 10,000 zł increase in the tax burden induces 1.96 percentage points of

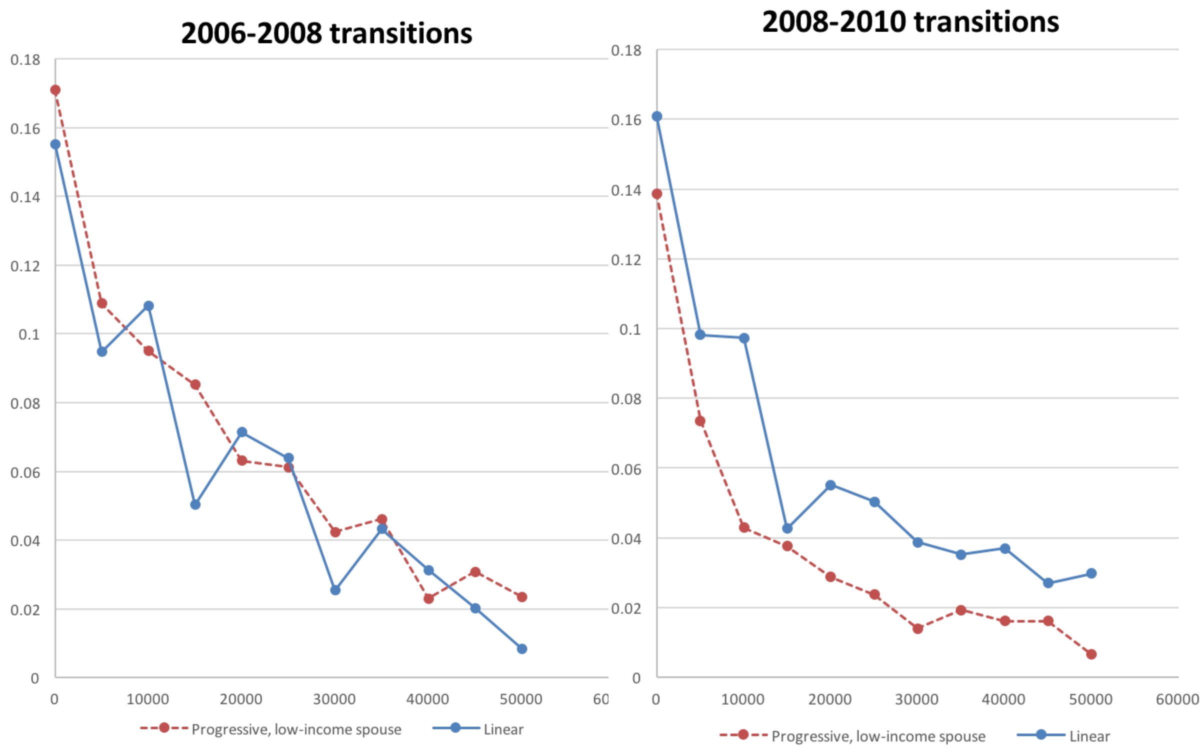


Figure 4: Transitions from business to employment between base year t and year $t+2$ for individuals who file under the linear schedule in the base year, and individuals who file under the progressive schedule but report with a spouse whose income falls below the first tax threshold.

those reporting employment income to transition to business income. A wider band around the second kink, which widens the definitions of treatment and control groups, lowers the estimate somewhat to 0.0162, or 1.62 percentage points in response to a 10,000 zł increase in the tax burden. The placebo estimates from comparing the years 2010 and 2012 are statistically insignificant, and thus do not contradict the identifying assumptions. The second column leaves out individuals closest to the kink as a robustness check: in the region of 5k zł above and below for the 30k zł-band estimates and 10k zł above and below for the 60k zł-band estimates. The baseline estimates in the 30k zł-band are not significantly altered, although the 60k

Self-employment to employment transitions. The estimate of $\alpha\times$ 10,000 zł of -0.0554 for transitions from business to employment shows that, for those who began in the business tax base, a 10,000 zł decrease in the tax burden in employment relative to business would induce 5.54% of those in business income to transition to employment. Again, the placebo estimates from the period from 2010 to 2012 are statistically insignificant, and thus support the identification strategy.

We may combine the estimated responses on transitions to and from self-employment, using formula (6). For simplicity, we may take the level of income of 150,000 zł, where there is roughly an equal share of individuals in employment and in self-employment. In

Table 5: Extensive-margin responses.

	2008 and 2010 (1)	2008 and 2010 (2-robustness)	2010 and 2012 (3-placebo)
A. Employment to business transitions			
30k zł band around first kink			
$\alpha^{\text{L to B}} \times$ 10,000 zł	0.0196	0.0169	-0.0081
	(0.00972)	(0.00887)	(0.0103)
Number of observations	170,287	101,951	177,292
60k zł band around first kink			
$\alpha^{\text{L to B}} \times$ 10,000 zł	0.0162	0.0079	-0.0005
	(0.00464)	(0.00247)	(0.0042)
Number of observations	375,771	273,631	388,985
B. Business to employment transitions			
$\alpha^{\text{B to L}} \times$ 10,000 zł	-0.0554	-	-0.0071
	(0.0084)	-	(0.0051)
Number of observations	26,639	-	18,301

Notes: The bands concern the width of the tax kink at the entry to the second tax bracket, e.g. at 30k zł band means the sample is restricted to those 15k zł below the cutoff and above (the control group) and those 15k zł above the cutoff and below (the treatment group). Column (2) excludes individuals immediately around the first tax threshold (earnings within 5k zł of threshold for the 30k zł band and within 10k zł for the 60k zł band). Column (3) contains the placebo estimates using the years 2010-2012; otherwise specification is as in column (1).

this case, the total estimate of $g\left(\widetilde{\phi_\theta}\right)$ would be $0.5 \times (-0.0554) - 0.5 \times 0.0196$, i.e. 0.0375, based on the above estimates. In other words, a decrease in the differential between employment and self-employment of 10,000zł at this income level would result in a total reduction of 3.75% of individuals transitioning to employment, partly due to increased transitions from self-employment to employment, and partly because fewer individuals would transition from employment to self-employment. These results may also be framed in terms of an elasticity, $\xi_\theta(\widetilde{\phi}) = \frac{\Delta T^{L,B}(\theta)}{G_\theta(\widetilde{\phi})} \frac{\partial G_\theta(\widetilde{\phi})}{\partial(\Delta T^{L,B}(\theta))}$. For instance, at the 150,000 zł income level where the difference in tax rates between employment and self-employment is around 6,500 zł, and the fraction of those on self-employment is about 0.5, the implied elasticity on the extensive margin is $\xi_\theta(\widetilde{\phi}) = \frac{6500}{0.5} \times 3.75 \times 10^{-6} = 0.049$. In other words, a 100 percent increase in the tax wedge between employment and self-employment results in a 4.9 percent increase in those declaring self-employment, i.e. increasing the share in self-employment from 0.500 to 0.525.

7 Deadweight Losses of the 2009 Reform

This section evaluates the DWL resulting from the reform, as well as the direct effect of the tax changes which occurred in Poland in 2009 on revenue. The results are presented in Table 6, and are based on the formulae for calculating DWL presented in Section 3 and extended to a case with multiple tax brackets in the online appendix. On the

Table 6: Calculations of DWL of reform.

Tax Base	Employment	Business
Mechanical effect		
1st tax band	-20,764,139	-6,755,519
2nd tax band	-286,883,745	-48,800,616
3rd tax band	-99,908,035	-16,905,894
Total	-407,555,918	-72,462,029
Behavioural effect		
Intensive Margin		
1st tax band	3,653,246	1,030,008
2nd tax band	29,700,905	14,927,247
3rd tax band	14,653,178	7,325,887
Total	48,007,329	23,283,143
Extensive Margin		
B to E	5,373,097	N/A
E to B	246,611	N/A
Total	5,619,708	

intensive-margin, this involves calculating the total income levels in each bracket, and adjusting it by the relevant tax parameters and elasticity estimates. The calculation on the extensive-margin is more involved. Specifically, at each income level I calculate the tax differential between employment and self-employment before and after reform, and use the extensive-margin parameter estimates to predict the share of individuals at each income level who would transition as a result of the reform.[27]

As can be seen in Table 6, the mechanical effects of losses in revenue due to the tax cuts are 407.5 million zł for employment and 72.5 million zł for the self-employment tax base. In comparison, the reduction in DWL on the intensive margin are 14.6 million zł for employment and 7.3 million zł for self-employment. It is interesting to note that the contribution of the behavioural responses along the intensive margin for self-employment is magnified by the higher elasticities in this tax base, so that they account for almost exactly half of recouped DWL but less than a quarter of the lost revenue.

The table shows that on the extensive margin, switching serves to further offset the mechanical losses of tax revenue. In particular, switching from the business tax base to employment results in an increase in tax revenue of around 5.4 million zł. A smaller amount of around 0.25 million zł was also predicted to have stemmed from reduced switching from the employment to the business tax base. These quantities would imply that, in the immediate aftermath of the reform, 7.9% of the revenue gains and the resultant

[27]Specifically, I do this based on pre-reform numbers of taxpayers in each tax base, in 10,000zł bands. Furthermore, for the calculation of DWL on the extensive margin, I assume that on switching taxable income does not change. This seems supported by the results of a Heckman-selection model presented in the online appendix, in which I do not observe large changes in income on transition from self-employment to employment.

fall in DWL were accounted for by the extensive-margin response.

8 Optimal Non-linear Taxation

I now examine how the presence of extensive-margin responses and different ETIs affect fully non-linear socially optimal tax schedules for employment and self-employment, in a framework in the spirit of Diamond (1998) and Saez (2001). I here closely follow the extensions of this framework by Kleven et al. (2009) and Scheuer (2014). Here, the question is: what nonlinear tax schedules the government should impose respectively on the self-employed and the workers in order to maximize social welfare, subject to not observing individuals' type, given different elasticities between workers and the self-employed, and given fixed costs of self-employment?

The assumptions regarding individual's characteristics and behaviour are here the same as in Section 3. In contrast to the model in Section 3, however, I now allow individuals to be subjected to non-linear taxation, with the possibility of employment and self-employment income being subject to different non-linear tax schedules. The budget constraint for each individual is then $c \leq l - T^L(l)$ if the individual earns employment income, and $c \leq b - T^B(b)$ if the individual earns business income. The terms $T^L(l)$ and $T^B(b)$ denote the non-linear tax schedules in employment and self-employment respectively, and are permitted to be either positive or negative (i.e. a transfer).

As is well-known since Saez (2001), the precise form of optimal tax schedules will depend on the empirical 'productivity' distributions of individuals, as well as on distributive preferences, in addition to parameters describing the behavioural responses to taxation. Here, the distribution of productivity corresponds to the distribution of θ and ϕ types in employment and self-employment. It is also necessary to augment efficiency considerations with social preferences. For this purpose, I introduce social welfare weights $\widetilde{g}_\theta(\phi)$ and $\widetilde{f}(\theta)$ to the model presented in Section 3, which may differ from the actual densities in the population and can be chosen to represent arbitrary preferences for redistribution between individuals of type θ and ϕ. As before, the cumulative distribution functions in the population are presented in capital letters $\tilde{F}(\theta)$ and $\tilde{G}_\theta(\phi) = \tilde{G}(\phi|\theta)$.

Individuals are assumed to have a quasi-linear utility of the form as in Section 3. Here, however, the cost functions $\psi^L(.)$ and $\psi^B(.)$ are given a specific form, namely $\psi^K(x) = \left(\frac{x}{1+\frac{1}{\varepsilon_K}}\right)^{1+\frac{1}{\varepsilon_K}}$. This functional form implies that the parameters ε_L and ε_B have a ready interpretation as elasticities of taxable income with respect to the marginal net-of-tax rate. As before, the fixed cost ϕ can be interpreted as the cost of setting up a business, attitudes towards risk, and preferences towards the flexibility of self-employment (or, alternatively, easier access to evasion opportunities). The parameters θ and $\tilde{\omega}\theta$ can be interpreted as the hourly wage rates for an hour worked as an employee or as a self-employed individual, respectively, in a standard model of labour supply. In this case, the parameter $\tilde{\omega}$ is a premium (positive or negative) to engaging in self-employment, per hour worked.

As before, the model allows us to separate the intensive margin decision (how much income to declare), from the extensive margin decision (what type of income to declare). The tax base choice for the individual is therefore determined by whether or not their fixed costs exceed the following threshold $\widetilde{\phi}(\theta) = v^B(\theta) - v^L(\theta)$. Now, however

$$v^L(\theta) = l(\theta) - T^L(l(\theta)) - \psi^L(l(\theta)/\theta) \quad (7)$$

$$v^B(\theta) = b(\theta) - T^B(b(\theta)) - \psi^B(b(\theta)/(\tilde{\omega}\theta)) - \phi \quad (8)$$

which are a function of the *non-linear* tax schedules faced in each tax base.

8.1 Optimal Tax Formulae

The social planner sets the parameters of the tax and transfer system, $T^L(l)$ and $T^B(b)$, in order to maximize aggregate utility, with each individual's utility using using the social planner's social welfare weights. The framing of the social planner's problem follows the classic approach of mechanism design – we require that each individual truthfully reports their type subject to the tax schedule (i.e. incentive compatibility holds), and that the total quantity of taxes and transfers balance (i.e. the resource constraint is satisfied).

The precise statement of the social planner's problem and its solution is shown in detail in the online appendix. Here, I only state the proposition summarizing the features of an optimal tax system thus derived:

Proposition 8.1. *At the social optimum, the tax schedule for workers, $T^L(\theta)$, and the tax schedule for the self-employed, $T^B(\theta)$, satisfy*

$$\frac{T'^L(\theta)}{1-T'^L(\theta)} = \left[\frac{1+\frac{1}{\varepsilon_L}}{\theta f(\theta)(1-G_\theta(\widetilde{\phi}_\theta))}\right] \int_{\underline{\theta}}^{\theta} \left[\left\{\widetilde{f}(\hat{\theta})(1-\widetilde{G}(\widetilde{\phi}_{\hat{\theta}})) - f(\hat{\theta})(1-G(\widetilde{\phi}_{\hat{\theta}}))\right\} - A(\hat{\theta};\widetilde{\phi}_{\hat{\theta}})\right]$$

$$\frac{T'^B(\theta)}{1-T'^B(\theta)} = \left[\frac{1+\frac{1}{\varepsilon_B}}{\theta f(\theta)G_\theta(\widetilde{\phi}_\theta)}\right] \int_{\underline{\theta}}^{\theta} \left[\left\{\widetilde{f}(\hat{\theta})\widetilde{G}(\widetilde{\phi}_{\hat{\theta}}) - f(\hat{\theta})G(\widetilde{\phi}_{\hat{\theta}})\right\} + A(\hat{\theta};\widetilde{\phi}_{\hat{\theta}})\right]$$

where $\Delta T^{L,B}(\hat{\theta}) = T^L(\hat{\theta}) - T^B(\hat{\theta})$ and $A(\hat{\theta};\widetilde{\phi}_{\hat{\theta}}) = f(\hat{\theta})\left(g(\widetilde{\phi}_{\hat{\theta}})\Delta T^{L,B}(\hat{\theta})\right) d\hat{\theta}$ is the extensive-margin term. Moreover, at the top of the θ-distribution, it can be shown that $T'^L(\bar{\theta}) = T'^B(\bar{\theta})=0$, and at the bottom, I have $T'^L(\underline{\theta}) = T'^B(\underline{\theta})=0$.[28]

Proof. See the online appendix. □

[28] With a bounded distribution of skills, a top marginal tax rate of zero is a standard result. Another classic result, namely zero marginal tax rates at the bottom, does not hold in the numerical implementation of the formula here as there is a positive mass of individuals not working at the bottom of the skill distribution.

The intuition outlined at the beginning of this paper regarding the interplay of intensive- and extensive-margin distortions is borne out in Proposition 8.1. The optimal marginal tax rate facing workers and the self-employed of type θ is negatively related to the parameters ε_L and ε_B, respectively, which turn out to be equivalent elasticities of employment (and self-employment) with respect to the net-of-tax rate:

$$\varepsilon_K = \frac{1 - T'^K(k(\theta))}{k(\theta)} \frac{\partial k(\theta)}{\partial(1 - T'^K(k(\theta)))}$$

where $k \in \{l, b\}$ deotes the level of income declared in a tax base and $K \in \{L, B\}$ indicates the tax base to which the elasticity or tax schedule applies. The optimal marginal tax is also negatively related to the mass of workers $f(\theta)(1 - G_\theta(\widetilde{\phi}_\theta))$ and the self-employed $f(\theta)G(\widetilde{\phi}_\theta)$. These terms together represent the mass of individuals whose declared income is distorted as a result of the marginal tax rates, and they scale the size of the distortion as captured by the elasticities of taxable income. The terms in curled brackets, which compare the social-welfare weights for all skill types $\hat{\theta}$ below θ with their respective population densities, capture the redistributive effects of the tax scheme. These terms are standard in the literature on optimal non-linear taxation (Diamond, 1998; Saez, 2001). The term $A(\hat{\theta}; \widetilde{\phi}_{\hat{\theta}})$, however, only appears as a result of the addition of the extensive margin. It captures the effect of a small change in tax rates T^L and T^B on the decision of taxpayers about whether to declare themselves employed or self-employed. A small increase in the tax rate on workers, T^L, for instance, has the effect of inducing $f(\theta)g(\widetilde{\phi}_\theta)$ to switch from employment to self-employment. The effect of this switching for each individual of type θ on the social planner's budget is the differential total tax rate between employment and self-employment, $\Delta T^{L,B}(\theta)$.[29]

From Proposition (8.1), it can be seen that the intensive-margin elasticities matter for both the DWL and optimal tax calculations. It can also be seen that, for the extensive margin, the relevant estimable parameters for the DWL and the optimal tax formulae are somewhat different. In particular, for the DWL, I require an estimate of all of the transitions to the business tax base as a result of the tax reform, weighted by the tax differential between the bases. For the optimal tax calculations, however, what is required is an estimate of the probability of transitions at each point in the θ-distribution as a result of a small change in the tax schedule. The goal of the remaining sections of this paper is to use the optimal tax formulae in Proposition 8.1 to make predictions about the optimal tax schedules for employment and self-employment for empirically plausible

[29]These results replicate the findings of Scheuer (2014), Proposition 4, those in Kleven et al. (2009), Proposition 8.1, and are also analogous to the intuitive derivation of the DWL for the Polish tax reform analyzed in Section 3. In contrast to the DWL formulae in Section 3, the optimal tax formulae require the entire set of probability densities, $g(.)$, and not just estimates locally around the point where the reform is happening. The precise way in which these densities are calibrated in this paper is outlined in Section 8.2.

values of the parameters.

8.2 Calibration

I now in turn present the calibration of each parameter required to make quantitative predictions using Polish tax data and estimates from Sections 5 and 6. For clarity, the complete set of parameter values chosen is summarized in Table 10 in the online appendix.

Elasticities of taxable income ε_L *and* ε_B. The elasticities of taxable income for workers and the self-employed are set to the baseline estimates from Section 5. I therefore set ε_L to be 0.2, and ε_B to be significantly higher at 0.6.

Distribution of 'skill' parameter θ *and* $\log \tilde{\omega}$. Following the method introduced by Saez (2001), I identify the distribution of θ from the observed income distribution for workers and the self-employed using tax data for the year 2007 in Poland. The method involves calibrating the skill distribution such that, given the utility function chosen and the actual tax schedule, the resulting income distribution replicates the observed earnings distributions in the two tax bases. From the first-order conditions for taxable income in both tax-bases, it is easy to show that, conditional on an individual of type θ being in either the employment and self-employment tax base, $\log(\theta)$ can be related to taxable income and marginal tax rates in the following way:

$$\log(\theta_i) = \frac{1}{1+\varepsilon_L}\log(l_{i,2007}) - \frac{\varepsilon_L}{1+\varepsilon_L}\log(1-\tau^L_{i,2007})$$

$$\log(\theta_i) = \frac{1}{1+\varepsilon_B}\log(b_{i,2007}) - \frac{\varepsilon_B}{1+\varepsilon_B}\log(1-\tau^B_{i,2007}) - \log(\tilde{\omega})$$

In other words, in order for the distribution of θ to replicate the observed income distributions, I can simply calculate θ for each individuals by using the formulae above.

Estimating θ for each employed individual is straightforward given the elasticity estimate for workers, observed income and marginal tax rates. For the self-employed, however, this is complicated by the term $\tilde{\omega}$, which I have noted could be interpreted as a form of a 'wage premium' on self-employment income, and which has not so far been estimated. Here, I examine several plausible values for $\tilde{\omega}$. The precise calibration has a large impact on the predicted self-employment incomes for each θ. For instance, a value of -2.0 tends to predict higher levels of self-employment income than employment income, especially at the higher values of θ in the range considered. On the other hand, -2.5 predicts self-employment income to be significantly lower for most of the distribution of θ than $\tilde{\omega} = -2.0$, while being slightly higher at the very top. A value of $\tilde{\omega} = -3.0$ predicts significantly lower self-employment income for individuals on transitioning. In the online appendix, I show that transitioning from employment to self-employment does not result in a dramatic change in income, or result in an increase in reported income, suggesting that -2.0 is the most plausible parameterisations.[30]

Distribution of fixed costs ϕ. The fixed costs of declaring self-employment income are assumed to be Pareto distributed. In particular, following similar previous quantitative explorations of models of optimal taxation with extensive-margin responses[31], the functional form of the cumulative density function is assumed to have a constant elasticity of the fraction of individuals in self-employment with respect to the fixed cost of transition. The cumulative density of fixed costs in the population is assumed to take the form:

$$G(\phi) = \begin{cases} 0 & \text{if} \quad \phi \leq \underline{\phi} \\ \left(\frac{\phi - \underline{\phi}}{\overline{\phi}}\right)^{\eta} & \text{if} \quad \phi \in \{\underline{\phi}, \overline{\phi}\} \\ 1 & \text{if} \quad \phi > \overline{\phi} \end{cases} \tag{9}$$

The distribution has three parameters, all of which require calibration: the elasticity η, as well as the bottom and top limits of the distribution of fixed costs $\{\underline{\phi}, \overline{\phi}\}$.

The bottom cut-off for the distribution of the fixed cost is arbitrarily set to $-10,000z\l$. This allows there to be a benefit to being self-employed for some individuals in the lower tail of the distribution of fixed costs. As a result, it helps the model to account for there being a net utility loss (absent fixed costs) on transitioning to self-employment at the bottom of the distribution (something which occurs in the parameterization used here), while also having a non-trivial proportion of individuals reporting to be self-employed.

The top limit of fixed costs $\overline{\phi}$ and the elasticity η are then calibrated together, separately for each parameterization of $\log \tilde{\omega}$. For the baseline calibration $\log \tilde{\omega} = -2.0$, the value of η is set to 0.45, while the value of the maximum is fixed at 900,000zł. These values guarantees an estimate of the marginal response of transitions to a change in fixed costs, $g(\phi)$, close to the estimate of this parameter in Section 6 (approx. 0.375×10^{-6}), at a level of business income equivalent to 80,000zł. This is roughly the mean of the treatment group in terms of income. A plot of the predicted fraction of individuals in self-employment for each level of fixed costs ϕ under the above parameterization is shown in Figure 9. Interestingly, the parameterization has the capacity to explain the level and rising share of the self-employed observed in the data in the top half of the θ distribution.

Social welfare weights $\tilde{G}(.), \tilde{g}(.), \tilde{F}(.), \tilde{f}(.)$. As Scheuer (2014) points out, a combination of weights $\tilde{G}(.)$ and $\tilde{F}(.)$, can be used to replicate a concave social welfare function applied to the utilities of each individual, $c - \psi^k(k/\theta) - \phi$. The presence of fixed costs of self-employment has implications for the type of weighting which should be used. In

[30]I use the Polish tax data from 2007 to find empirical marginal tax rates faced by individuals at each income level, for each tax base. All individuals in the sample who reported positive employment or self-employment income were used, with individuals reporting any self-employment income being classified as self-employed. The marginal tax rates were calculated based on an individual's reported taxable income and any other information about them, including whether or not they were reporting with a spouse, had any deductions or whether they were reporting according to the preferential linear rate for the self-employed.

[31]These include Kleven et al. (2009) and Scheuer (2014).

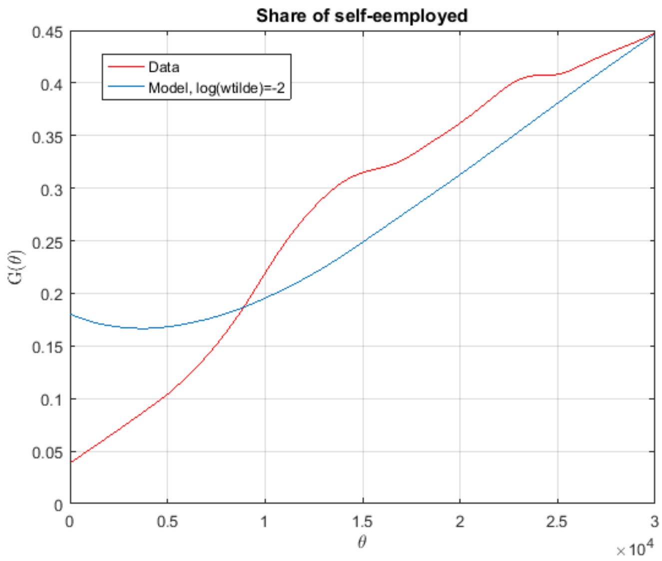


Figure 5: The shares of self-employed for each θ in the model under this parameterization (blue line) and in the data (red line).

particular, if the fixed costs are positive, the weights $\tilde{G}(.)$ should be below the observed distribution, $\tilde{G}(.) \leq G(.)$, whereas if they are negative (and there are net welfare benefits to self-employment), the reverse should be the case. For simplicity, since for part of the range of the distribution ϕ is negative, and for another it is positive, I assume that the social planner cares only about redistribution along the θ-dimension. Consequently, I set $G(.) = \tilde{G}(.)$, and model preferences for redistribution purely in terms of θ. Specifically, for assigning weights to individual of different θ-types, I use the weights $\tilde{F}(\theta) = 1 - (1 - F(\theta))^{1+\rho}$ which are also used by Scheuer (2012).

The welfare weights can be simply interpreted as the weight assigned to the indirect utility of each type in the social welfare function and fixed costs faced by individuals. If the welfare weights $\tilde{F}(.)$ are the same as the observed density $F(.)$, then because of the quasi-linear utility there are no benefits from redistribution. Thus, if $\rho = 0$, the welfare weights will be the same and it follows from Proposition 8.1 that optimal marginal tax rates will be zero. Following Scheuer (2014), the value of ρ is set to 0.1 in the baseline parameterization, implying a small preference for redistribution. As a robustness check, however, in the online appendix I also examine the solution to the optimal tax problem where the vale of ρ is set to higher values, namely 0.5 and 1.0. The basic results regarding the differential in marginal and total tax rates are largely unchanged.

8.3 Analysis

Having set out the calibrations for the model, Figure 6 plots against each other the marginal and total tax rates for each skill type for the case in which there is an extensive-margin response (shown in blue), and for the case where this response has been switched off (shown in red), and where the fraction of self-employed individuals is kept at 0.15 for each θ. The left-hand side panels show the tax schedules for the self-employed, while the right-hand side panels show the tax schedule for workers. The tax schedules are represented in terms of marginal and total rates: the top panels illustrate marginal tax rates and the bottom panels illustrate total tax rates.

The first striking feature is that marginal tax rates on self-employment are not seen to vary much across the distribution for the income range without an extensive margin, falling to around 15% at the mode, and thereafter rising to around 20-22% in the 100,000-200,000zł range. This is surprisingly close to the 19% preferential tax rate observed in practice in the Polish tax system. For employment income, on the other hand, marginal tax rates at the mode are higher at around 27%, and subsequently rise to around 40%.

Secondly, the introduction of the extensive margin does not radically alter the optimal tax schedules in the calibrations used here. In line with what would be expected, without an extensive-margin response the marginal tax rates on self-employment are permitted to be slightly lower than with an extensive-margin response. The social planner can reduce lower marginal tax rates in the case without an extensive margin and not worry about inducing switching on the extensive margin. The degree to which marginal tax rates can be lowered with an extensive-margin response present is reduced. This can be seen in the top-left panel of Figure 6. However, this effect appears to be quantitatively small at about 3-4% around the 300,000zł mark. Moreover, it only manifests itself in differences for the self-employment schedule – the introduction of extensive-margin responses barely makes any difference to tax schedules for workers. The differences in marginal tax rates for self-employed individuals manifest themselves in a slightly higher level of total tax rates on self-employment. In terms of total tax rates, the differential total tax burden at 200,000zł is about 4,000zł lower with the presence of an extensive-margin response.

9 Conclusion

This paper addresses two key outstanding questions surrounding the optimal taxation of employment and self-employment. Firstly, it estimates the degree of switching in response to differential taxation between the two tax bases, while also contributing to the literature on the intensive-margin elasticities within each tax base. Secondly, it applies these estimates to an optimal taxation theoretical framework to ask whether there is a case for self-employment income to be taxed at preferential rates relative to employment.

The magnitude of the extensive-margin response is non-trivial, with a 10,000 zł ($ 2,670 as of May 2017) change in the relative tax burden implying a change in the probability of a taxpayer filing business income by 3.75 percentage points. This is a sizeable response considering that the magnitude of taxpayers reporting self-employment income at median income levels is small (around 10%, which rises to around 50% in the top

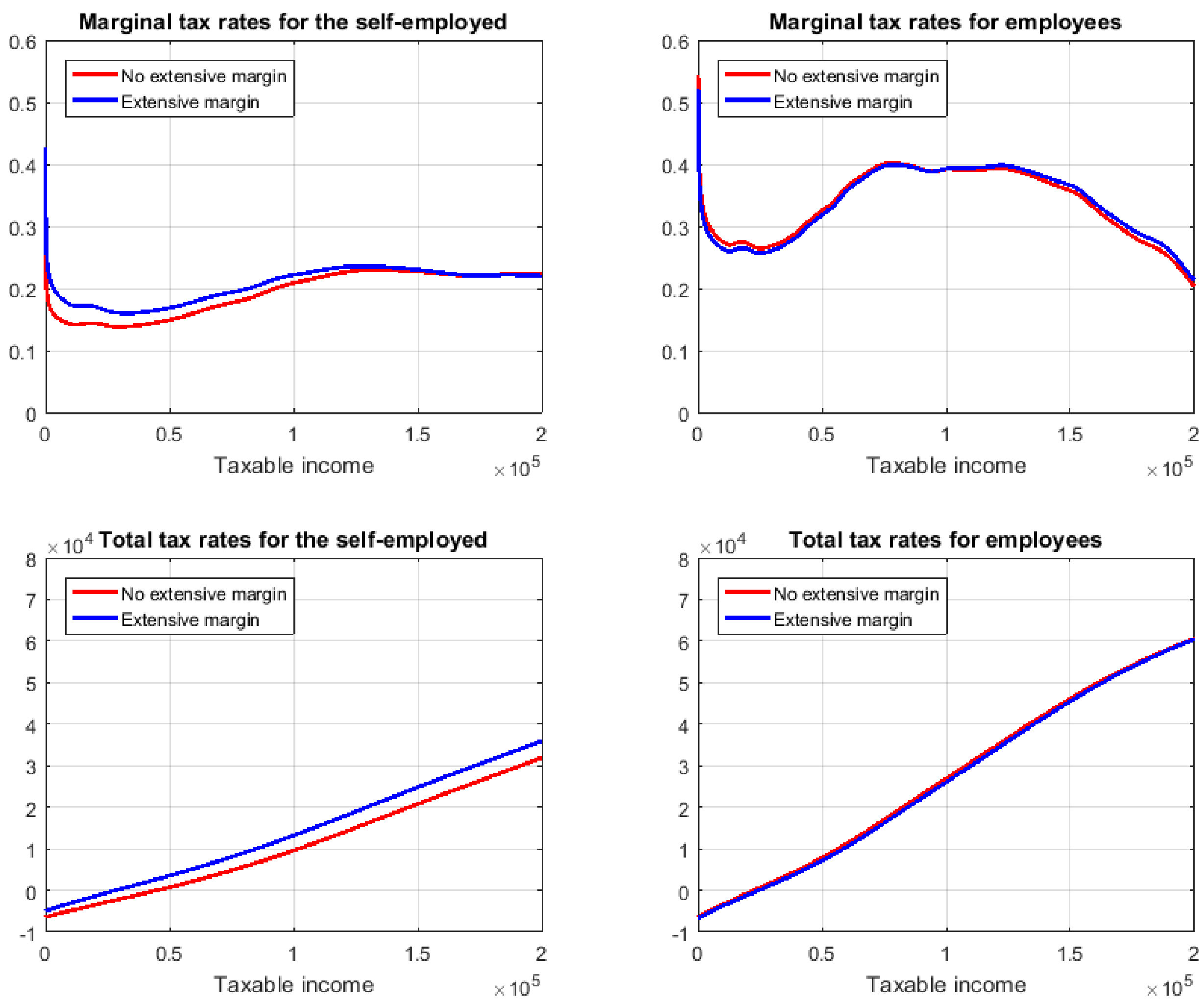


Figure 6: Comparison of optimal tax schedules with and without the extensive margin.

percentile of the income distribution). Using a methodology which allows controlling for time-varying income-growth in the income distribution, intensive-margin elasticities for the employed were found to be in the 0.21-0.25 region across most of the income distribution. The self-employed elasticities were found to be roughly three times larger than elasticities for the employed. The higher elasticities for self-employment were found to make a disproportionately large contribution of self-employment to the DWL from the tax schedule present in Poland, and hence accounted for a disproportionately large contribution to the change in DWL as a result of the 2009 Polish tax cuts.

For the calibrations used here, higher elasticities of taxable income for self-employment do indeed justify significantly lower marginal tax rates in this tax base. The presence of an extensive-margin switching response tempers the rationale for preferential taxation only slightly, at least according to a parameterisation which fits the estimated size of extensive-margin responses. In this sense, the model appears to rationalize the preferential tax treatment observed in many countries. In particular, this would seem to provide one justification for the Polish tax system, which offers the self-employed a preferential linear tax rate of 19%, with higher marginal tax rates on employment income at the top.

The applicability of these results to other contexts crucially depends on two questions: 1) similarities in intensive-margin elasticities and 2) similarities of the institutional framework around self-employment, which is likely to determine the non-tax related costs of self-employment. It may be said that the estimated elasticity of self-employment income here, in Kopczuk (2015) and in Saez (2003) all point to a magnitude of elasticities among the self-employed closer to 1. These studies focus on quite different countries – Poland and the United States – yet the elasticity estimates are similar and higher than other studies that focus on the whole population of taxpayers. Regarding the institutional frameworks, on the other hand it would appear that in the United States, at least, there is an apparently more stringent 20-part test to determine whether someone should be classified as an employee or contractor. While this partly overlaps with the formal test for Poland outlined in the Introduction, it would appear more stringent. As a result, it may be reasoned that the non-tax costs of switching in Poland may be lower.

At the same time, the results in this paper do not necessarily entail that the present tax system in Poland approaches optimality. Different levels of redistributive preferences would naturally result in different levels of overall taxation and redistribution to those observed in practice, and there is considerable variation in marginal tax rates over income implied by the model which are not reflected in the existing Polish system. Finally, as has been noted in the existing literature on intensive-margin responses to taxation, the size of behavioural responses cannot really be treated as an exogenous parameter, but is a function of the rules governing the tax system, as well as the regulation of self-employment. Thus, improving tax enforcement for self-employment might well be a further policy tool for governments not considered here. Changes in responses of self-employment to taxation as a result of changes in reporting requirements is an important follow-up to the current project.

References

Alstadster, Annette and Martin Jacob, "Dividend Taxes and Income Shifting," *Scandinavian Journal of Economics*, 2016, *118* (4), 693–717.

Burns, Sarah K. and James P. Ziliak, "Identifying the Elasticity of Taxable Income," *Economic Journal*, March 2017, *127* (600), 297–329.

Chetty, Raj, "Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance," *American Economic Journal: Economic Policy*, August 2009, *1* (2), 31–52.

Crawford, Claire and Judith Freedman, "Small Business Taxation," in "The Mirrlees Review – Dimensions of Tax Design," Institute for Fiscal Studies, 2010.

Diamond, Peter A, "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, March 1998, *88* (1), 83–95.

Feldstein, Martin, "Behavioral Responses to Tax Rates: Evidence from the Tax Reform Act of 1986," *American Economic Review*, May 1995, *85* (2), 170–74.

Fudenberg, Drew and Jean Tirole, *Game Theory*, MIT Press, Cambridge, 1991.

Gordon, Roger H and Joel B Slemrod, "Are" Real" Responses to Taxes Simply Income Shifting Between Corporate and Personal Tax Bases?," *Does Atlas Shrug?: The Economic Consequences of Taxing the Rich*, 2000, p. 240.

Gruber, Jon and Emmanuel Saez, "The elasticity of taxable income: evidence and implications," *Journal of Public Economics*, April 2002, *84* (1), 1–32.

Harju, Jarkko and Tuomas Matikka, "Business owners and income-shifting: evidence from Finland," *Small Business Economics*, January 2016, *46* (1), 115–136.

_ **and** _ , "The elasticity of taxable income and income-shifting: what is real and what is not?," *International Tax and Public Finance*, August 2016, *23* (4), 640–669.

Kleven, Henrik Jacobsen and Esben Anton Schultz, "Estimating Taxable Income Responses Using Danish Tax Reforms," *American Economic Journal: Economic Policy*, November 2014, *6* (4), 271–301.

_ **, Claus Thustrup Kreiner, and Emmanuel Saez**, "The Optimal Income Taxation of Couples," *Econometrica*, 03 2009, *77* (2), 537–560.

Kopczuk, Wojciech, "Tax bases, tax rates and the elasticity of reported income," *Journal of Public Economics*, December 2005, *89* (11-12), 2093–2119.

_ , "The Polish business 'flat' tax and its effect on reported incomes: a Pareto improving tax reform?," *Mimeo.*, 2015.

Mirrlees, J. A., "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 1971, *38* (2), 175–208.

Piketty, Thomas and Emmanuel Saez, "Optimal Labor Income Taxation," NBER Working Papers 18521, National Bureau of Economic Research, Inc November 2012.

Saez, Emmanuel, "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, Jan 2001, *234* (68), 205–229.

_ , "The effect of marginal tax rates on income: a panel study of 'bracket creep'," *Journal of Public Economics*, May 2003, *87* (5-6), 1231–1258.

_ , "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, August 2010, *2* (3), 180–212.

_ **and Stefanie Stantcheva**, "Generalized Social Marginal Welfare Weights for Optimal Tax Theory," *American Economic Review*, January 2016, *106* (1), 24–45.

_ **, Joel Slemrod, and Seth H. Giertz**, "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, March 2012, *50* (1), 3–50.

Scheuer, Florian, "Entrepreneurial Taxation and Occupational Choice," in "Business Taxation (Trans-Atlantic Public Economics Seminar)" NBER Chapters, National Bureau of Economic Research, Inc, 02 2012.

_ , "Entrepreneurial Taxation with Endogenous Entry," *American Economic Journal: Economic Policy*, May 2014, *6* (2), 126–63.

Slemrod, Joel and Wojciech Kopczuk, "The optimal elasticity of taxable income," *Journal of Public Economics*, April 2002, *84* (1), 91–112.

Weber, Caroline E., "Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences," *Journal of Public Economics*, 2014, *117* (C), 90–103.

Online Appendix Not For Publication

A DWL of Polish 2009 Reform

In the main text, I show how the extensive-margin transitions affect the DWL of taxation under a simplified linear tax schedules. Here, I can use an analogous method to evaluate actual tax reforms such as the 2009 Polish tax reform. These are the formula used to calculate the DWL in Section 7 of the paper.

To calculate the DWL, I take into account the piecewise linear nature of the initial tax schedule, as captured by the presence of tax bands. Let $\bar{l}^1$and $\bar{l}^2$ be the threshold values for entering the middle and top tax bracket respectively, and $d\tau_L^1$ and $d\tau_L^2$ are changes to the marginal tax rate in these tax brackets. If I assume that the elasticity of taxable income for the employed is constant across the population, such that $\bar{\varepsilon}_L = \int_\Theta \left[(1-G_\theta(\widetilde{\phi}))l(\theta)\varepsilon_L(\theta)\right] dF\theta = \varepsilon_L \times \int_\Theta \left[(1-G_\theta(\widetilde{\phi}))l(\theta)\right] dF\theta$, the DWL associated with the tax reform now is given by:[32]

$$
\begin{aligned}
DWL = \varepsilon_L & \left[\left(\frac{\tau_L^1}{1-\tau_L^1}\right)\bar{l}^1 \times d\tau_L^1 + \left(\frac{\tau_L^2}{1-\tau_L^2}\right)\bar{l}^2 \times d\tau_L^2\right] + \\
& + \int_\Theta \left[\mathbf{1}\{l(\theta) \in [\bar{l}^1, \bar{l}^2]\} \times g_\theta\left(\widetilde{\phi}\right)\left((l(\theta)-\bar{l}^1)\times d\tau_L^1\right)\Delta T^{L,B}(\theta)\right] dF\theta + \\
& + \int_\Theta \left[\mathbf{1}\{l(\theta) \in [\bar{l}^2, +\infty)\} \times g_\theta\left(\widetilde{\phi}\right)\left((\bar{l}^2-\bar{l}^1)\times d\tau_L^1\right)\Delta T^{L,B}(\theta)\right] dF\theta + \\
& + \int_\Theta \left[\mathbf{1}\{l(\theta) \in [\bar{l}^2, +\infty)\} \times g_\theta\left(\widetilde{\phi}\right)\left((l(\theta)-\bar{l}^2)\times d\tau_L^2\right)\Delta T^{L,B}(\theta)\right] dF\theta. \quad (10)
\end{aligned}
$$

Here, $\mathbf{1}\{l(\theta) \in [\bar{l}^1, \bar{l}^2]\}$ is an indicator for an individual falling into the middle tax bracket, $\mathbf{1}\{l(\theta) \in [\bar{l}^2, +\infty)\}$ is an indicator for an individual falling into the top tax bracket, and τ_L^1 and τ_L^2 are the marginal tax rates in the middle and top tax brackets respectively. The expression $\bar{h}^k$ is the average income in bracket k, multiplied by the number of individuals in that tax bracket. The expressions for the extensive-margin response now take into account how the reforms affect the probability of switching in a non-linear fashion, depending on whether an individual falls in the middle or higher tax bracket. The probability of switching is simply the density of individuals at the point of indifference between the two tax bases, multiplied by the change in the income differential between the employment and business tax base. This is just $g_\theta\left(\widetilde{\phi}\right)\left((l(\theta)-\bar{l}^1)\times d\tau_L^1\right)$

[32] In the main text, it is shown that the assumption of constant elasticities across the income-distribution appears to be in fact plausible.

Table 7: Own-elasticity estimates - robustness checks.

Years	2008-09 (1)	2008-09 (2)	2008-09 (3)	2008-09 (4)
A. Employment sample				
ε_L	0.218	0.203	0.229	0.230
	(0.025)	(0.025)	(0.026)	(0.031)
Base-year income controls	Spline	Log-linear	Spline	Spline
First-stage F-stat	35,345.0	36,237.3	41,053.7	50,254.9
Number of individuals	100,248	100,181	149,006	69,680

Years	2008-09 (1)	2008-09 (2)	2008-09 (3)	2008-09 (4)
B. Business sample				
ε_B	0.657	-	0.834	0.400
	(0.287)	-	(0.286)	(0.311)
Base-year income controls	Log-linear	-	Log-linear	Log-linear
First-stage F-stat	1,503.1	-	1,572.7	1,738.5
Number of individuals	6,856	-	7,172	4,943

Notes: (1) is the baseline specification controlling for base-year income using a 10-piece spline in the employment sample and a log-linear specification for the business sample (2) controls for base-year income using a log-linear specification (3) introduces a lower 10,000 zł cut-off above 0 and below twice the first tax threshold (previously 20,000) and 2,500zł for the business sample (previously 5,000zł) (4) requires the join income of the taxpayer and spouse to be 20,000 zł above the threshold for the Treatment group and below the threshold for the Control group.

and $g_\theta\left(\widetilde{\phi}\right)\left((l(\theta)-\bar{l}^2)\times d\tau_L^2+(\bar{l}^2-\bar{l}^1)\times d\tau_L^1\right)$ for each type θ, for the middle and top bracket, respectively.

It is well-known that intensive-margin elasticities ε_L and ε_B, as well as the average income levels in the tax bands, matter for DWL. It is important to note that now what also matters is the change in the probability of reporting employment income captured in the integral terms in (10). This is a statistic which can be estimated empirically. The formula in (10) also illustrates that in calculating the switches as a result of the reform, we ought to weight observations by the predicted differential in tax rates between the employment and business for each individual.

B Intensive-Margin Results: Robustness

See Table 7.

C Impact of Reform on Self-Employment Revenue and Costs

Table 8 provides the results for estimating the impact of the 2009 reform on the amount of business revenue and business costs among the self-employed. A discussion of these results is presented in the main text of the paper.

Table 8: Responses of self-employment net income, revenue and costs.

Years	2008-09	2008-09	2008-09
Dependent variable	(Net income)	(Revenue)	(Costs)
ε_B	0.657	0.236	0.289
	(0.287)	(0.195)	(0.211)
Number of observations	6,856	7,227	7,191

Notes: The sample consists of individuals who reported only business income, with all individuals who earned below 5,000zł in 2004 terms, or above twice the first tax threshold minus 5,000zł. The log of base-year income is used as a control in all regressions. The instruments are as indicated in the main text.

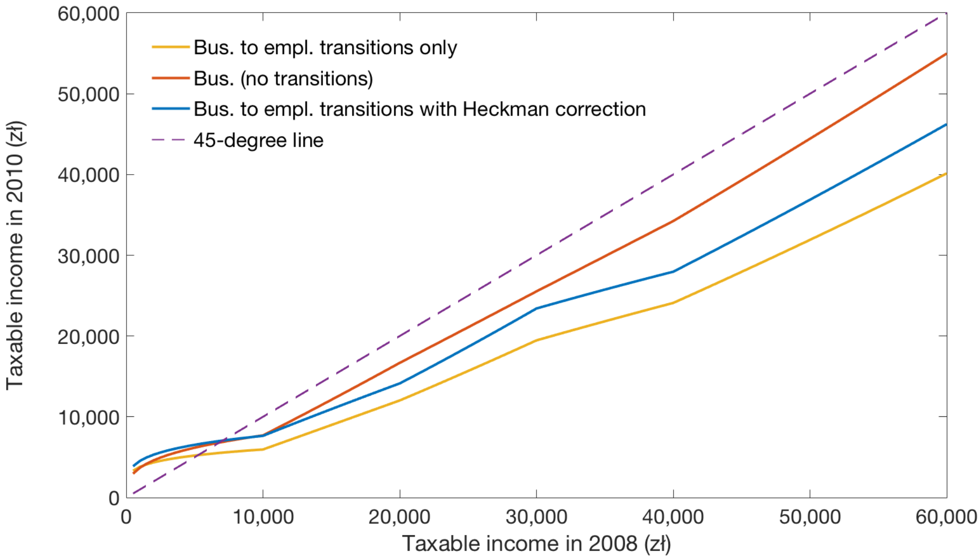
D Effect of Switching on Income

To identify the impact of moving from the business to the employment tax base on taxable income, the Heckman selection-correction model presents a suitable approach. In practice, the difference in income between employment and self-employment is only observed for those who switch tax bases. To the extent that switching is driven by unobservable factors which are also correlated with income growth, any estimate relying on those who do actually switch will be subject to selection bias. However, a Heckman selection-correction model with a suitable exclusion restriction may allow us to adjust for a term in the estimate of the income change on switching.

Here, we focus on the sample of individuals who were in self-employment in 2008 who (i) either filed under the linear schedule or (ii) filed under the progressive schedule with spouses whose spousal income was low enough to guarantee that they were in the lower tax bracket in 2008. The former group is understood as the treatment group, while the latter group is understood to be the control group. This is analogous to the methodology for estimating the effect of the reform on switching from self-employment to employment presented above.

As in a classic Heckman model, the transition from the business to the employment tax base is modelled as the Probit model

Figure 7: Income change on transition from business to employment.



$$Pr(trans_{it+s}^{B\to L}) = \Phi\left(\gamma + g(z_{it}) + \delta \cdot 1(i \in T)\right)$$

where the group T is defined as the individuals reporting income according to the linear tax schedule. The interaction term $1(i \in T) \cdot 1(t \geq 2009)$ will here be treated as an exclusion restriction. The intuition is that an individual owning a business under linear reporting is more likely to switch to employment than a business owner reporting under the progressive tax base with a low-income spouse. This is because they are more likely to fall into the higher tax brackets on switching than an individual sharing tax liability with a spouse, all other things being equal.

The model of the income process for the switchers is

$$\Delta \log(z_{i,t+1}) = \alpha + \beta\lambda\left(\hat{\gamma} + \hat{g}(z_{it}) + \hat{\delta} \cdot 1(i \in T)\right) + f(z_{it}) + \xi_{i,t+1}$$

where $\lambda(.)$ is the estimated Mills ratio, $\lambda(.) = \phi(.)/\Phi(.)$ obtained from the estimated first-stage regression, $f(.)$ is a flexible spline in base-year income. The results to this estimation procedure are presented in Table 9, and the predicted levels of income on transition are shown in Figure 7.

E Social Planner's Problem

I here present the social planner's problem regarding the optimal nonlinear tax policy, which is used to derive the optimal tax formulae in Proposition 8.1.

By the revelation principle, any tax mechanism can be replicated by a tax mechanism which induces truthful revelation of individuals' types. The choice of the tax schedule by

Table 9: Heckman model of changes in income on transition between tax bases.

	Heckman	OLS (non-switchers)	OLS (switchers)
Base-year (log) taxable income spline 1	0.202	0.319	0.195
	(0.0653)	(0.0150)	(0.0295)
Base-year (log) taxable income spline 2	0.860	1.115	1.014
	(0.122)	(0.0402)	(0.0875)
Base-year (log) taxable income spline 3	1.214	1.050	1.188
	(0.247)	(0.0814)	(0.188)
Base-year (log) taxable income spline 4	0.586	1.020	0.743
	(0.408)	(0.131)	(0.279)
Base-year (log) taxable income spline 5	1.204	1.168	1.258
	(0.188)	(0.0611)	(0.0900)
Base-year (log) business income spline 1	-0.00459	-0.0257	-0.0763
	(0.0983)	(0.00966)	(0.0140)
Base-year (log) business income spline 2	-0.381	-0.186	-0.468
	(0.177)	(0.0391)	(0.101)
Base-year (log) business income spline 3	-0.363	-0.159	-0.650
	(0.390)	(0.0915)	(0.302)
Base-year (log) business income spline 4	-0.930	-0.390	-0.323
	(0.780)	(0.166)	(0.459)
Base-year (log) business income spline 5	-0.777	-0.0333	-1.195
	(1.529)	(0.273)	(0.116)
Mill's ratio	-0.318		
	(0.448)		
Constant	7.738	6.305	7.750
	(0.596)	(0.124)	(0.247)
Observations	25,704	25,704	3,575

Standard errors in parentheses

the social planner can then be framed as an assignment of a level of consumption $c(.)$, business income $b(.)$, and employment income $l(.)$ to each individual of type $[\theta, \phi]$. There are three sets of incentive-compatibility constraints a social planner must take into account when setting the tax schedules, to ensure truthful reporting. Specifically, conditional on occupational choice, each type θ must prefer to select truthful consumption-income bundles in both tax bases

$$c^L(\theta) - \psi^B \left(\frac{l(\theta)}{\theta} \right) \geq c^L(\hat{\theta}) - \psi^L \left(\frac{l(\hat{\theta})}{\hat{\theta}} \right) \qquad \forall \theta, \hat{\theta} \in \Theta \tag{11}$$

$$c^B(\theta) - \psi^B \left(\frac{b\theta)}{\tilde{\omega}\theta} \right) \geq c^B(\hat{\theta}) - \psi^B \left(\frac{b(\hat{\theta})}{\tilde{\omega}\hat{\theta}} \right) \qquad \forall \theta, \hat{\theta} \in \Theta \tag{12}$$

At the same time, the threshold cost for entrepreneurship must satisfy the condition

$\widetilde{\phi}(\theta) = v^B(\theta) - v^L(\theta)$ discussed in the main text, for all $\theta \in \Theta$.

Differentiating the expressions for indirect utility in the main text (equations (7) and (8)) with respect to θ, and applying the envelope theorem, I get

$$v^{B\prime}(\theta) = \psi^B\left(\frac{e^B(\theta)}{\theta}\right)\frac{e^B(\theta)}{\theta^2} \qquad \forall\theta \in \Theta \tag{13}$$

$$v^{L\prime}(\theta) = \psi^L\left(\frac{l^L(\theta)}{\theta}\right)\frac{l^L(\theta)}{\theta^2} \qquad \forall\theta \in \Theta \tag{14}$$

It can be shown that the incentive constraints (11) and (12) are satisfied if and only if two conditions are satisfied: the above envelope conditions hold and the functions $b(\theta)$ and $l(\theta)$ are non-decreasing (see Fudenberg and Tirole, 1991, theorems 7.2 and 7.3).[33]

To incorporate the social planner's preferences for redistribution, I add social welfare weights expressed by the cumulative distribution functions $\widetilde{G}(\phi)$ and $\widetilde{F}(\theta)$, with corresponding marginals $\widetilde{g}(\phi)$ and $\widetilde{f}(\theta)$. Such welfare weights are an alternative to using social welfare functions to capture preferences for redistribution.[34] To create a rationale for taxation in the present model, with utilities additive in consumption, these must differ from the actual densities in the population. They can be chosen to represent arbitrary preferences for redistribution between individuals of type θ and/or ϕ.

We are now in a position to represent the social planner's problem. We can replace consumption as a control variable by using the indirect utilities, thus leading to the set of choice variables $\{b(\theta), l(\theta), v_L(\theta), v_B(\theta), \tilde{\phi}_\theta\}$. In this case, the social planner will seek to maximize utility using their social preference weights:

$$\max_{\{b(\theta), l(\theta), v_L(\theta), v_B(\theta), \tilde{\phi}_\theta\}} \int_\Theta \left[\tilde{G}(\tilde{\phi}_\theta)v^B(\theta) + (1 - \tilde{G}(\tilde{\phi}_\theta))v^L(\theta)\right] d\widetilde{F}\theta - \int_\Theta \left[\int_{\underline{\phi}}^{\widetilde{\phi}_\theta} \phi\widetilde{g}\left(\phi\right) d\phi\right] d\widetilde{F}\theta$$

such that the incentive-compatibility constraints (13), (14) and $\widetilde{\phi}(\theta) = v^B(\theta) - v^L(\theta)$ are satisfied $\forall\theta, \hat{\theta} \in \Theta$, and the following resource constraint holds:

$$\int_\Theta \left[(G(\tilde{\phi}_\theta)b(\theta) + (1 - G(\tilde{\phi}_\theta)l(\theta))\right] dF\theta - \int_\Theta \left[G(\tilde{\phi}_\theta)c^B(\theta) + (1 - G(\tilde{\phi}_\theta))c^L(\theta)\right] dF\theta \geq 0$$

The resource constraint simply ensures that the total level of consumption in the economy is less than or equal to the total resources produced by the self-employed and the workers.

Screening problems with more than one dimension have generally been recognized to

[33]Whether the functions $b(\theta)$ and $l(\theta)$ are in practice non-decreasing needs to be verified. The distribution of skills may then result in a distribution of incomes with either bunching or a gap in the distribution of incomes. As is common in the literature (see e.g. Scheuer (2014)), I proceed with the assumption of no bunching and show that it does not apply for a variety of parameter values.

[34]Saez and Stantcheva (2016) explain the relationship between such weights and the social welfare function approach in detail.

pose significant methodological problems, as the standard first-order approach used by Mirrlees (1971) and much of the subsequent literature may no longer be valid. This is because the incentive-compatibility constraints must account for the ability of agents to deviate along more than one dimension of type. Here, double-deviation (for instance, a change in tax base associated with a change on the intensive-margin) is not a concern, since at a given declared earnings are allowed to change optimally, within a given tax base.

E.1 Proof of Proposition 8.1

To recapitulate what is stated in the main text, the social planner's maximization problem is that of maximizing utility subject to truthful revelation by each type of individual, and subject to the resource constraint that all consumption is no greater than the output of the economy. This can be summarized as the following:

$$\max_{\{b(\theta),l(\theta),v_L(\theta),v_B(\theta),\tilde{\phi}_\theta\}} \int_\Theta \left[\tilde{G}(\tilde{\phi}_\theta)v^B(\theta) + (1-\tilde{G}(\tilde{\phi}_\theta))v^L(\theta)\right] d\widetilde{F}\theta - \int_\Theta \left[\int_{\underline{\phi}}^{\tilde{\phi}_\theta} \phi\widetilde{g}(\phi)\, d\phi\right] d\widetilde{F}\theta \tag{15}$$

such that

$$v^{B'}(\theta) = \psi^B\left(\frac{e^B(\theta)}{\theta}\right)\frac{e^B(\theta)}{\theta^2} \qquad \forall\theta,\hat{\theta}\in\Theta \tag{16}$$

$$v^{L'}(\theta) = \psi^L\left(\frac{l^L(\theta)}{\theta}\right)\frac{l^L(\theta)}{\theta^2} \qquad \forall\theta,\hat{\theta}\in\Theta \tag{17}$$

$$\widetilde{\phi}(\theta) = v^B(\theta) - v^L(\theta) \qquad \forall\theta,\hat{\theta}\in\Theta \tag{18}$$

$$\int_\Theta \left[(G(\tilde{\phi}_\theta)(\tilde{\omega}e(\theta)) + (1-G(\tilde{\phi}_\theta)l(\theta))\right] dF\theta - \int_\Theta \left[G(\tilde{\phi}_\theta)c^B(\theta) + (1-G(\tilde{\phi}_\theta))c^L(\theta)\right] dF\theta \geq 0 \tag{19}$$

with variables defined in the main text in Section 8. Here, we have also introduced the notion of an 'original wage' in self-employment $e(\theta)$, defined as $e(\theta) = b(\theta)/\bar{\omega}$. This can be interpreted as the additional level of income generated by an extra hour of self-employment.

Truthful revelation is ensured by three incentive-compatibility constraints, two on the intensive-margin for each tax base, (16) and (17), and one on the extensive margin, (18). The first two are derived from the first-order conditions of individuals optimally choosing the amount of taxable income to declare in each tax base, while the latter is derived from the threshold condition at which an individual is indifferent between the two tax bases.

The social planner's objective function involves purely the social welfare weights assigned to the indirect utilities, as well as to the fixed costs faced by individuals. It is important to note that in this derivation, we use the assumption that the level of taxable business income $b(\theta)$ is equal to the premium for self-employment multiplied by the 'original wage' in self-employment, $e(\theta)$. In other words, $b(\theta) = \tilde{\omega}e(\theta)$. This simplifies considerably the derivations.

Having applied integration by parts to (16) and (17), and incorporating the fact that $T^L(\theta) = v^L(\theta) + \psi^L(l(\theta)/\theta)$ and $T^B(\theta) = v^B(\theta) + \psi^B\left(e^B(\theta)/\theta\right)$, the resulting Lagrangian for the social planner's problem becomes

$$
\begin{aligned}
L = \int_{\Theta} \left[\tilde{G}(\tilde{\phi}_\theta)v^B(\theta) + (1-\tilde{G}(\tilde{\phi}_\theta))v^L(\theta)\right] d\widetilde{F}\theta - \int_{\Theta}\left[\int_{\underline{\phi}}^{\tilde{\phi}_\theta} \phi\widetilde{g}(\phi)\,d\phi\right] d\widetilde{F}\theta - \\
- \int_{\Theta}\left[\mu^{L'}(\theta)v^L(\theta) + \mu^L(\theta)\psi^L\left(\frac{l^L(\theta)}{\theta}\right)\frac{l^L(\theta)}{\theta^2}\right] dF\theta - \\
- \int_{\Theta}\left[\mu^{B'}(\theta)v^B(\theta) + \mu^B(\theta)\psi^B\left(\frac{e^B(\theta)}{\theta}\right)\frac{e^B(\theta)}{\theta^2}\right] dF\theta - \\
+ \lambda^{RC}\int_{\Theta}\left[G(\tilde{\phi}_\theta)\left(\tilde{\omega}e(\theta) - v^B(\theta) - \psi^B\left(\frac{e^B(\theta)}{\theta}\right)\right)\right] dF\theta - \\
- \lambda^{RC}\int_{\Theta}\left[(1-G(\tilde{\phi}_\theta))\left(l(\theta) - v^L(\theta) - \psi^L\left(\frac{l^L(\theta)}{\theta}\right)\right)\right] dF\theta
\end{aligned}
$$

The social planner is assumed to be maximizing the Lagrangian by assigning levels of taxable income to individuals of each type, subject to truthful revelation and optimization on the extensive margin, while also assigning levels of indirect utility to each type. This procedure yields the same result as if the social planner simply set appropriate taxes and transfers, and allowed individuals to respond to them optimally on the intensive and extensive margins. Thus, the *control variables* of the problem are $\{b(\theta), l(\theta), v_L(\theta), v_B(\theta), \tilde{\phi}_\theta\}$. We derive the first-order conditions for each of these accordingly.

The first-order conditions for $v^B(\theta)$ and $v^L(\theta)$ are, respectively:

$$
\mu^{B'}(\theta) = \tilde{G}(\tilde{\phi}_\theta)\widetilde{f}(\theta) - \lambda^{RC}G(\tilde{\phi}_\theta)f(\theta) + \lambda^{RC}g(\tilde{\phi}_\theta)f(\theta)\left[b(\theta) - c^B(\theta) - (l(\theta) - c^L(\theta))\right]
$$

$$
\mu^{L'}(\theta) = \left(1-\tilde{G}(\tilde{\phi}_\theta)\right)\widetilde{f}(\theta) - \lambda^{RC}\left(1-G(\tilde{\phi}_\theta)\right)f(\theta) - \lambda^{RC}g(\tilde{\phi}_\theta)f(\theta)\left[b(\theta) - c^B(\theta) - (l(\theta) - c^L(\theta))\right]
$$

We define the term $\Delta T(\theta) = T^B(\theta) - T^L(\theta) = b(\theta) - c^B(\theta) - (l(\theta) - c^L(\theta))$ as the tax differential between self-employment and employment. Integrating the above first-order conditions, the definition for $\Delta T(\theta)$ and the transversality conditions to maximization problem, $\mu^B(\underline{\theta}) = \mu^L(\underline{\theta}) = 0$ and $\mu^B(\bar{\theta}) = \mu^L(\bar{\theta}) = 0$, we obtain:

$$0 = \int_{\Theta} \left[\tilde{G}(\tilde{\phi}_\theta) f(\theta) - \lambda^{RC} G(\tilde{\phi}_\theta) f(\theta) + \lambda^{RC} g(\tilde{\phi}_\theta) f(\theta) \Delta T(\theta) \right] d\widetilde{F}\theta$$

$$0 = \int_{\Theta} \left[\left(1 - \tilde{G}(\tilde{\phi}_\theta)\right) f(\theta) - \lambda^{RC} \left(1 - G(\tilde{\phi}_\theta)\right) f(\theta) - \lambda^{RC} g(\tilde{\phi}_\theta) f(\theta) \Delta T(\theta) \right] d\widetilde{F}\theta$$

Adding the two conditions yields the result that $\lambda^{RC} = 1$. We incorporate this insight into the first-order conditions for $v^B(\theta)$ and $v^L(\theta)$. Integrating, we see that

$$\mu^B(\theta) = \int_{\Theta} \left[\tilde{G}(\tilde{\phi}_\theta) f(\theta) - G(\tilde{\phi}_\theta) f(\theta) + g(\tilde{\phi}_\theta) f(\theta) \Delta T(\theta) \right] d\widetilde{F}\theta$$

$$\mu^L(\theta) = \int_{\Theta} \left[\left(1 - \tilde{G}(\tilde{\phi}_\theta)\right) f(\theta) - \left(1 - G(\tilde{\phi}_\theta)\right) f(\theta) - g(\tilde{\phi}_\theta) f(\theta) \Delta T(\theta) \right] d\widetilde{F}\theta$$

Now, the first-order conditions for $e(\theta)$ and $l(\theta)$ are, respectively:

$$G(\tilde{\phi}_\theta) f(\theta) \left[\tilde{\omega} - \frac{1}{\theta} \psi^{B'} \left(\frac{e(\theta)}{\theta} \right) \right] = \frac{\mu^B(\theta)}{\theta} \left[\frac{1}{\theta} \psi^{B'} \left(\frac{e(\theta)}{\theta} \right) + \frac{e(\theta)}{\theta^2} \psi^{B''} \left(\frac{e(\theta)}{\theta} \right) \right]$$

$$\left(1 - G(\tilde{\phi}_\theta)\right) f(\theta) \left[1 - \frac{1}{\theta} \psi^{L'} \left(\frac{l(\theta)}{\theta} \right) \right] = \frac{\mu^L(\theta)}{\theta} \left[\frac{1}{\theta} \psi^{L'} \left(\frac{l(\theta)}{\theta} \right) + \frac{l(\theta)}{\theta^2} \psi^{L''} \left(\frac{l(\theta)}{\theta} \right) \right]$$

It is easy to rearrange this first-order conditions and introduce terms representing the elasticities of taxable income for the employment and self-employment tax bases. Dividing through by $\psi^{B'}(e(\theta)/\theta)/\theta$ and $\psi^{L'}(l(\theta)/\theta)/\theta$, respectively, we get:

$$\frac{\tilde{\omega} - \psi^{B'} \left(e(\theta)/\theta \right)/\theta}{\psi^{B'} \left(e(\theta)/\theta \right)/\theta} = \frac{\mu^B(\theta)}{\theta f(\theta) G(\tilde{\phi}_\theta)} \left(\frac{1 + \psi^{B''}(e(\theta)/\theta)(e(\theta)/\theta^2)}{\psi^{B'} \left(e(\theta)/\theta \right)/\theta} \right)$$

$$\frac{1 - \psi^{L'} \left(l(\theta)/\theta \right)/\theta}{\psi^{L'} \left(l(\theta)/\theta \right)/\theta} = \frac{\mu^L(\theta)}{\theta f(\theta) \left(1 - G(\tilde{\phi}_\theta)\right)} \left(\frac{1 + \psi^{L''}(l(\theta)/\theta)(l(\theta)/\theta^2)}{\psi^{L'} \left(l(\theta)/\theta \right)/\theta} \right)$$

Since the elasticities of self-employment income with respect to the net-of-tax rate, and the elasticity of employment income with respect to the net-of-tax rate, are:

$$\varepsilon^B(\theta) = \frac{\psi^{B'} \left(b(\theta)/\tilde{\omega}\theta \right)/(\tilde{\omega}\theta)}{\psi^{B''} \left(b(\theta)/\tilde{\omega}\theta \right) \left(b(\theta)/(\tilde{\omega}\theta)^2 \right)}$$

$$\varepsilon^L(\theta) = \frac{\psi^{L'} \left(l(\theta)/\theta \right)/\theta}{\psi^{L''} \left(l(\theta)/\theta \right) \left(l(\theta)/\theta^2 \right)}$$

respectively, we can immediately substitute them into the earlier expressions.

Finally, we may substitute into the first-order conditions for $e(\theta)$ and $l(\theta)$ the terms for the Lagrange multipliers on the incentive-compatibility constraints from the first-order conditions for the levels of indirect utility $v^B(\theta)$ and $v^L(\theta)$. This gives the conditions in Proposition 8.1:

$$\frac{T'^B(\theta)}{1-T'^B(\theta)} = \left[\frac{1+\frac{1}{\varepsilon_B}}{\theta f(\theta) G_\theta(\widetilde{\phi_\theta})}\right] \times \int_{\underline{\theta}}^{\theta} \left[\left\{\widetilde{f}(\hat{\theta})\widetilde{G}(\widetilde{\phi_{\hat{\theta}}}) - f(\hat{\theta})G(\widetilde{\phi_{\hat{\theta}}})\right\} + f(\hat{\theta})\left(g(\widetilde{\phi_{\hat{\theta}}})\Delta T^{L,B}(\hat{\theta})\right)\right] d\hat{\theta}$$

$$\frac{T'^L(\theta)}{1-T'^L(\theta)} = \left[\frac{1+\frac{1}{\varepsilon_L}}{\theta f(\theta)(1-G_\theta(\widetilde{\phi_\theta}))}\right] \times$$
$$\times \int_{\underline{\theta}}^{\theta} \left[\left\{\widetilde{f}(\hat{\theta})(1-\widetilde{G}(\widetilde{\phi_{\hat{\theta}}})) - f(\hat{\theta})(1-G(\widetilde{\phi_{\hat{\theta}}}))\right\} - f(\hat{\theta})\left(g(\widetilde{\phi_{\hat{\theta}}})\Delta T^{L,B}(\hat{\theta})\right)\right] d\hat{\theta}$$

The limit conditions $T^B(\underline{\theta}) = T^L(\underline{\theta}) = 0$ and $T^B(\bar{\theta}) = T^L(\bar{\theta}) = 0$ are a consequence of evaluating at $\underline{\theta}$ and $\bar{\theta}$, and applying the transversality conditions $\mu^B(\underline{\theta}) = \mu^L(\underline{\theta}) = 0$ and $\mu^B(\bar{\theta}) = \mu^L(\bar{\theta}) = 0$. ■

E.2 Incorporating an intensive switching margin

An extension of the above model introduces a third category of reporting, namely reporting both positive employment and self-employment income. Here, I consider two types of fixed cost: one for reporting positive employment income with c.d.f. $G^L_{\hat{\theta}}(\widetilde{\phi^L_{\hat{\theta}}})$, and one for reporting positive self-employment income with c.d.f. $G^B_{\hat{\theta}}(\widetilde{\phi^B_{\hat{\theta}}})$. The model here assumes that such an individual pays fixed costs of reporting both types of income if they wish to report positive amount of each, and having paid these may alter both levels of income on the intensive margin in response to changes in the relative tax treatment of these types of income.

Following an analogous derivation to the one above, for those who report both types of income, FOCs of Lagrangean can be shown to be:

$$\frac{T_b^{L,B}(\theta)}{1-T_b^{L,B}(\theta)} = \left[1+\frac{\varepsilon_{L,1-T_l^{L,B}} - \varepsilon_{B,1-T_l^{L,B}}}{\varepsilon_{B,1-T_b^{L,B}}\cdot\varepsilon_{L,1-T_l^{L,B}} - \varepsilon_{L,1-T_b^{L,B}}\cdot\varepsilon_{B,1-T_l^{L,B}}}\right] \times A(\theta)\times B(\theta)$$
$$\frac{T_l^{L,B}(\theta)}{1-T_l^{L,B}(\theta)} = \left[1+\frac{\varepsilon_{B,1-T_l^{L,B}} - \varepsilon_{L,1-T_l^{L,B}}}{\varepsilon_{L,1-T_l^{L,B}}\cdot\varepsilon_{B,1-T_b^{L,B}} - \varepsilon_{L,1-T_b^{L,B}}\cdot\varepsilon_{B,1-T_l^{L,B}}}\right] \times A(\theta)\times B(\theta)$$

where $\varepsilon_{k,1-T_j^{L,B}}$ is the intensive-margin elasticity of income in the tax base k with respect to the net-of-tax rate in the tax base j. Also

$$A(\theta) = \left[\theta f(\theta) S^{L,B}(\theta)\right]^{-1}$$

$$B(\theta) = \int_{\underline{\theta}}^{\theta} [\Big(\widetilde{f}(\hat{\theta})\widetilde{S}^{L,B}(\theta) - f(\hat{\theta})S^{L,B}(\theta)\Big) + \\ + \Big(g^B_{\hat{\theta}}(\widetilde{\phi^B_{\hat{\theta}}})G^L_{\hat{\theta}}(\widetilde{\phi^L_{\hat{\theta}}})f(\hat{\theta})\Delta T^{LB,L}(\hat{\theta}) + g^L_{\hat{\theta}}(\widetilde{\phi^L_{\hat{\theta}}})G^B_{\hat{\theta}}(\widetilde{\phi^B_{\hat{\theta}}})f(\hat{\theta})\Delta T^{LB,B}(\hat{\theta})\Big)]d\hat{\theta}$$

$\Delta T^{j,k}(\hat{\theta})$ represent the difference in total taxation of type θ between tax base j and k.

It can be seen that with a 0 cross-elasticity between the two tax bases, the formulae above reduce to the formulae in the previous section. For those who report only business and only labour income, the optimal tax formulae are analogous to those in the previous section, where only own-elasticities matter on the intensive-margin:

$$\frac{T^B_b(\theta)}{1 - T^B_b(\theta)} = \left[1 + \frac{1}{\bar{\varepsilon}_{B,1-T^B_b}}\right] / \left[\theta f(\theta) S^B(\theta)\right] \times \int_{\underline{\theta}}^{\theta} [\Big(\widetilde{f}(\hat{\theta})\widetilde{S}^B(\theta) - f(\hat{\theta})S^B(\theta)\Big) + \\ + f(\hat{\theta})\left(\left[\int_{\widetilde{\phi}^B}^{\bar{\phi}^B} g^B_\theta\left(\phi^B\right) g^L_\theta\left((\widetilde{\phi}^L - \widetilde{\phi}^B) + \phi^B\right) d\phi^B\right] \Delta T^{B,L}(\hat{\theta}) - \\ - g^L_{\hat{\theta}}(\widetilde{\phi^L_{\hat{\theta}}})G^B_{\hat{\theta}}(\widetilde{\phi^B_{\hat{\theta}}})\Delta T^{LB,B}(\hat{\theta})]d\hat{\theta}$$

$$\frac{T^L_l(\theta)}{1 - T^L_l(\theta)} = \left[1 + \frac{1}{\bar{\varepsilon}_{L,1-T^L_l}}\right] / \left[\theta f(\theta) S^L(\theta)\right] \times \int_{\underline{\theta}}^{\theta} [\Big(\widetilde{f}(\hat{\theta})\widetilde{S}^L(\hat{\theta}) - f(\hat{\theta})S^L(\hat{\theta})\Big) - \\ - f(\hat{\theta})\left(\left[\int_{\widetilde{\phi}^B}^{\bar{\phi}^B} g^B_{\hat{\theta}}\left(\phi^B\right) g^L_{\hat{\theta}}\left((\widetilde{\phi}^L_{\hat{\theta}} - \widetilde{\phi}_{\hat{\theta}}^{B}) + \phi^B\right) d\phi^B\right] \Delta T^{B,L}(\hat{\theta}) + \\ + g^B_{\hat{\theta}}(\widetilde{\phi^B_{\hat{\theta}}})G^L_{\hat{\theta}}(\widetilde{\phi^L_{\hat{\theta}}})\Delta T^{LB,L}(\hat{\theta})]d\hat{\theta}$$

where, $S^B(\theta)$ and $S^L(\theta)$ represent shares reporting self-employment and employment, respectively. The only additional complication is that now there are three 'bases' between which there is switching: the employment tax base, the self-employment tax base, and the tax base of those who report both types of income.

F Robustness: Social Welfare Weights

Here, I explore the impact of varying the core parameter for preferences for redistribution, ρ. The baseline parameterisation, setting $\rho = 0.1$, is somewhat arbitrary, and in practice a social planner may have greater preferences for redistribution. Figure 8 illustrates the optimal marginal and total tax rates for the two tax bases for the baseline parameterisation, as well as for ρ set to 0.5 and 1.0.

Table 10: Calibration of key parameters.

Parameter	Value	Source
ε_L	0.2	Table 2.
ε_B	0.6	Table 3.
$\log\tilde{\omega}$	-2	See discussion in text.
$f(\theta)$, $F(\theta)$	See online appendix	Polish tax data.
η	0.45	Table 5, business shares.
$\bar{\phi}$	900,000 zł	Table 5, business shares.
ϕ^{MIN}	-20,000 zł	Arbitrary bottom cut-off.
ρ	0.1	Scheuer (2014)

Interestingly, the implication of increasing preferences for redistribution has an effect of increasing marginal tax rates relative to the baseline for both the self-employment and employment schedule, but to a large extent the wedge between the employment and self-employment tax base is preserved. The exception to this pattern can be seen in the middle panel, where I observe that the difference in marginal tax rates between employment and self-employment is increased relative to the baseline for values of θ between 0 and 5,000. While the marginal tax rates for $\rho = 0.5$ preserve something of the U-shape of the baseline regression, for $\rho = 1.0$, preference for redistribution are such that marginal tax rates are declining almost completely along the θ-distribution. Analogously to the case of marginal tax rates, the wedge in total tax rates between employment and self-employment also appears to be to a large extent similar for all three parameterisations. Naturally of course, total tax rates increase for both bases to finance greater positive transfers at the very bottom of the θ-distribution (total tax rates below 0 here of course represent positive transfers).

G Calibration and Optimal Tax Formulae

G.1 Calibration

Here, I provide additional illustration of the calibrations used in the paper. Figure 9 illustrates the cumulative distribution function of fixed cost of self-employment under the baseline calibration. It should be clear from the graph that a certain fraction of individuals would choose to be self-employed even if there were no net utility gains from self-employment (absent fixed costs), since they derive a positive gain from being in self-employment. Figure 10 illustrates the predicted density of skills θ for different values of the parameter $\log\tilde{\omega}$. Under all calibrations, the mode of the skill distribution is located at around 300k zł. Figure 11 illustrates how predicted income for self-employment and employment changes for each skill type at each calibration of $\log\tilde{\omega}$, and also how the shares of self-employed change for each value of the parameter θ. Finally, Figure 12 provides a graphical representation of the welfare densities used by the social planner

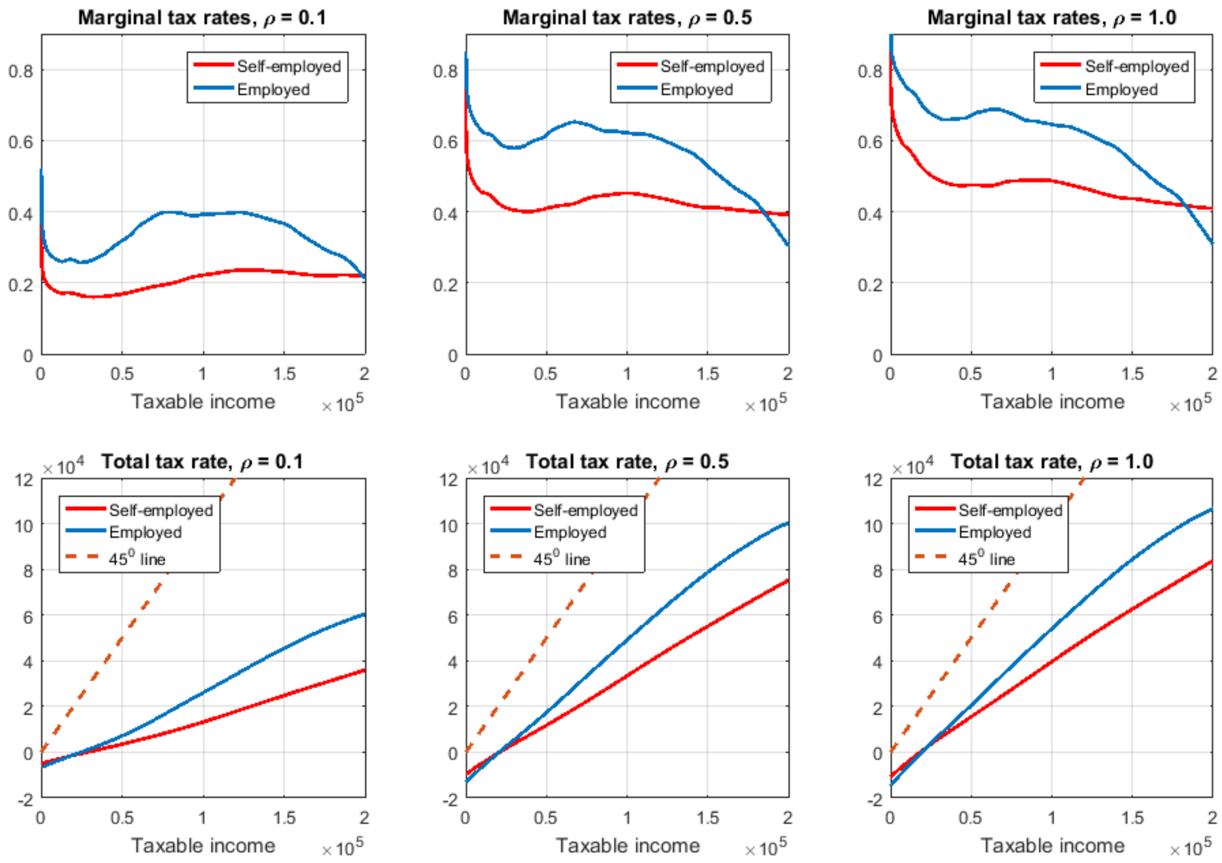


Figure 8: Optimal tax schedules for different values of the redistribution parameter, ρ.

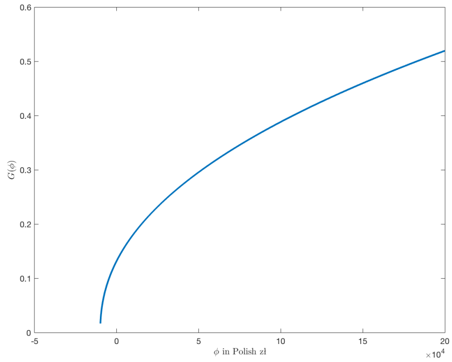


Figure 9: The CDF of fixed costs under the baseline parameterization $\log \tilde{\omega} = -2.0$.

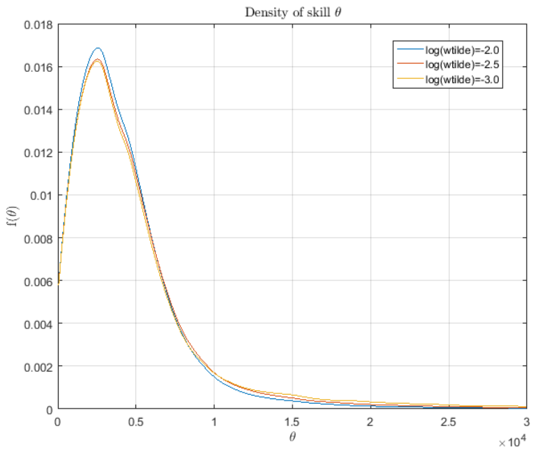


Figure 10: The predicted distribution of θ for each calibration.

under different values of the redistribution parameter ρ.

G.2 Comparison of Optimal Tax Rate Schedules with Existing Literature

It is instructive to compare the profile of marginal and total tax rates found in the paper with the profiles usually found in optimal tax simulations. Similarly to the simulations of Saez (2001), the zero-marginal tax rate at the bottom result does not obtain, as even at the bottom there is a positive mass of individuals supplying zero income at the bottom of the distribution. However, since the distribution of θ is bounded, the marginal tax rates are zero at the top of the distribution of θ (this is not visible in the top-left panel, as the x-axis is restricted to a subset of the income range spanning all individuals). Additionally, marginal tax rates are seen to decline from a relatively high level for both employment and self-employment until they reach the mode of the skill distribution, and increase thereafter. Such a U-shaped pattern of marginal tax rates has been discussed at length by Diamond (1998), and is also a feature of the simulations in Saez (2001). A high density

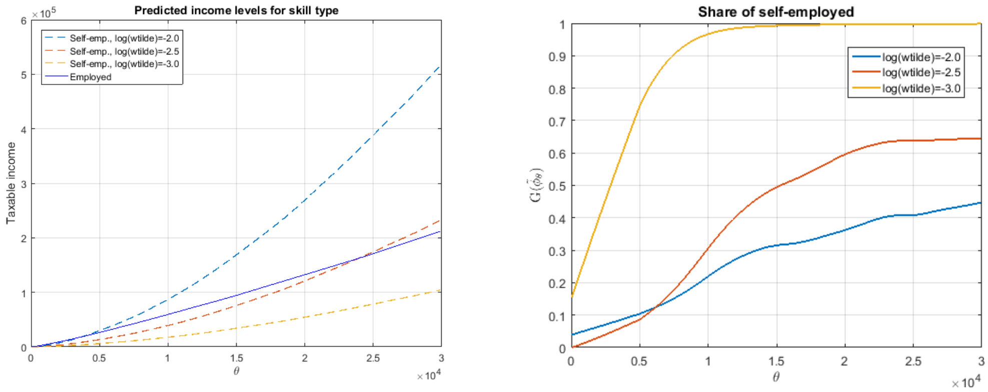


Figure 11: The left panel shows the predicted income levels in self-employment and in employment for individuals of type $\theta \in [0, 30000]$. Calculations for self-employment use three different calibrations for $\log \tilde{\omega}$, namely -2.0, -2.5 and -3.0. The panel on the right shows the implied shares of individuals of each type declaring self-employment for each calibration.

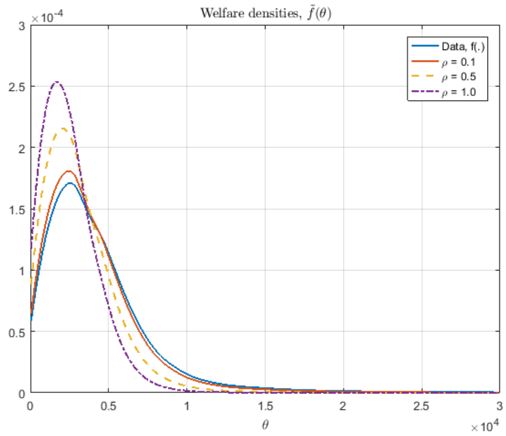


Figure 12: Comparison of actual density of the productivity parameter, θ, with welfare weights assigned by a social planner with increasing preferences for redistribution.

of individuals around the mode implies large welfare losses of high marginal tax rates due to intensive-margin responses, creating a rationale for increasing marginal tax rates lower down the income distribution (and effecting high total tax rates for individuals higher up the distribution, without at the same time creating intensive-margin distortions). On the other hand, as the modal skill is passed, there is a declining ratio of individuals at the point of distortion relative to the mass of individuals higher up, creating a rationale for increasing marginal tax rates. At the very top, this rationale disappears, and I arrive at the zero top marginal tax rate result.

H Deadweight losses under evasion

Suppose now individuals have an additional opportunity to either hide some of their income x or to report it as business costs. As a result, if they are reporting business income their utility now takes the form:

$$u(c, b; \theta) = b - \tau_B(b - x) - \psi^B(b/(\tilde{\omega}\theta)) - \lambda(x) - \phi \tag{20}$$

The amount of taxed income is now $z = b - x$. However, this evasion behaviour incurs a cost given by the convex function $\lambda(.)$. We assume that the utility function for an individual who remains in the employment sample is unchanged from before.

We now explore the implications this availability of avoidance behaviour has for the DWL of taxation. The agent optimally sets the marginal cost of not reporting 1zł less to the tax authority or reporting it as costs, namely $\lambda'(x)$, equal to the marginal private value of doing so, namely τ_B, the amount of tax not paid. At the same time, the agent also works to the point where their marginal disutility of earning another zł, $1/(\tilde{\omega}\theta) \times \psi^{B\prime}(b/\tilde{\omega}\theta))$ equals the marginal cost of doing so, $1 - \tau_B$.

Since the agent is already optimising, a change in marginal tax rates has no impact on the welfare of the agent. This is a result of the envelope theorem. The marginal social cost of reducing earnings and reporting less income is given purely by the marginal social value of the tax revenue earned. The relevant parameter to estimate this is not the elasticity of all income, but only the elasticity of taxable income, $z = b - x$. In the above case, this is $\frac{\partial z(\theta)}{\partial \tau_B} d\tau_B$, which can be expressed in terms of the taxable income elasticity:

$$\varepsilon_B(\theta) = \frac{1 - \tau_B}{z(\theta)} \frac{\partial(z(\theta))}{\partial(1 - \tau_B)}. \tag{21}$$

as $-\frac{z(\theta)}{(1-\tau_B)}\varepsilon_B(\theta) d\tau_B$

It may be noted that this is the case either when the individual is reporting income as costs, or if they are hiding their income completely.