Dingquan Miao

Two Essays in Public Economics

Licentiate Thesis
Economics
2020
Acknowledgments

I am greatly indebted to my Ph.D. advisors, Spencer Bastani and Thomas Giebe, for all the support they gave me since the first day of the Ph.D. project. Spencer provided me with every bit of guidance, assistance, and expertise that I needed throughout the research work. Thomas gave me valuable academic and health advice. I am honored to be their first Ph.D. student at Linnaeus University.

I am very grateful to my co-authors, Håkan Selin and Martin Söderström, who guide me to a whole new world of writing a joint paper. Håkan inspired me and encouraged me in every conversation we had. Martin is a very kind and smart friend. I thank Håkan and Martin for being great teachers and team members. It was a great experience to meet Håkan and Martin at IFAU and to have an anniversary dinner at Holmes’ English Pub.

I would like to thank Tobias König for his valuable comments in my final seminar of Licentiate. It was a lot of fun to teach public economics with Tobias. I thank Gazi Salah Uddin, Malte Zoubek, and seminar participants at Linnaeus University, UCFS Ph.D. workshop in Public Economics in Uppsala, SWEGPEC workshop in Blekinge, IFAU, and Umeå School of Business, Economics and Statistics for comments.

I appreciate the funding sources that made my Ph.D. work possible. I was funded by the Department of Economics and Statistics at Linnaeus University. I also acknowledge funding from Linnaeus University Centre for Discrimination and Integration Studies. The funding allows me to take courses and attend international seminars. I would like to thank Magnus Carlsson for answering all my practical questions about the Ph.D. program. I am also very grateful to my examiner, Lars Behrenz, who has helped me with the courses and thesis. I would like to express my gratitude and appreciation to all my colleagues at the department.

Lastly, I would like to thank my family for all their love, support, and encouragement. I am truly grateful for my parents’ unconditional support and encouragement. I want to thank my wife Shujun who has been my best friend and a great companion during my Ph.D. studies.
Abstract

This thesis consists of two self-contained essays.

Essay I.
I compute how optimal income tax schedules optimally should be differentiated between immigrants and natives. I use a calibrated optimal tax model with heterogeneous labor supply elasticities across and within groups and employ Utilitarian and Rawlsian social welfare functions. As compared to an optimal tax system that treats both groups the same, the optimal differentiated tax system increases marginal tax rates for the majority of natives, with a decrease in the marginal tax rates of immigrants. However, there is not much redistribution between the groups.

Essay II. (with Håkan Selin and Martin Söderström)
Sweden introduced a phase-out of the earned income tax credit in 2016. As a consequence, taxpayers belonging to the top 5 percent of the earnings distribution, already facing high taxes, experienced a 7% reduction in their net-of-tax shares. While exploiting rich full-population administrative data up to 2017, we evaluate earnings responses to the reform. When graphically and econometrically comparing earnings growth at different segments of the distribution, we estimate a significant relative earnings reduction in the treatment group immediately appearing in 2016, and growing in 2017. The implied earnings elasticity is fairly low and around 0.1. We interpret the essential features of the response using a simulation model, in which people have noisy perceptions of the piece-wise linear tax code. To simulate the empirically observed response, we need to add more noise to perceptions than what is motivated by earnings uncertainty alone.
Table of Contents

Acknowledgments iii
Abstract v
Table of Contents vii

Essay I: Should Natives and Immigrants Face Different Income Tax Schedules?
Essay II: Earnings Responses to Even Higher Taxes
Should Natives and Immigrants Face Different Income Tax Schedules?

Dingquan Miao⇤†

Abstract

I compute how optimal income tax schedules optimally should be differentiated between immigrants and natives. I use a calibrated optimal tax model with heterogeneous labor supply elasticities across and within groups and employ Utilitarian and Rawlsian social welfare functions. As compared to an optimal tax system that treats both groups the same, the optimal differentiated tax system increases marginal tax rates for the majority of natives, with a decrease in the marginal tax rates of immigrants. However, there is not much redistribution between the groups.

Keywords: Optimal income taxation, Tagging, Immigrants.

⇤Department of Economics and Statistics, Linnaeus University. Email: dingquan.miao@lnu.se
†I am most grateful to my supervisors Spencer Bastani and Thomas Giebe. I thank Tobias König, Gazi Salah Uddin, Malte Zoubek, and seminar participants at Linnaeus University, UCFS PhD workshop in Public Economics in Uppsala, and SWEGPEC workshop in Blekinge for comments. I am very grateful to Håkan Selin for providing me with the tax calculator.
1 Introduction

The seminal work of Mirrlees [1971] describes how the government re-distributes income and raises revenue through nonlinear (progressive) income taxation while minimizing distortions to work incentives. The shape of the optimal income tax schedule depends on three determinants, the distribution of skill levels, the elasticity of labor supply, and the governments’ attitudes towards inequality.

Beginning with Akerlof [1978], a literature has developed that explores how redistribution can be achieved at a lower efficiency cost by conditioning taxes on different non-income characteristics, known as ‘tagging’. Tagging implies that different groups of the population face different tax schedules. The benefit of this is that the group-specific tax schedules can be tailored to characteristics of the taxpayers in each group, more specifically, to the group-specific skill distributions and the group-specific labor supply behavior. Examples of tagging schemes that have been studied in the previous literature are tagging by age, gender and height (Kremer, 2002; Blomquist and Micheletto, 2008; Weinzierl, 2011; Cremer et al., 2010; Alesina et al., 2011; Bastani et al., 2013; Mankiw and Weinzierl, 2010). The most useful tags are strongly related to the three determinants of optimal tax rates. Age-dependent income taxation is based on the fact that different age groups have different income distributions, e.g., the hazard rate is much larger for young workers than prime-age workers (Kremer, 2002). Gender-based income taxation relies on the assumption that women have lower elasticity of labor supply than men and their wage distributions differ (Alesina et al., 2011; Bastani, 2013).

In this paper, I study immigration-status-based income taxation by computing optimal income tax schedules for native and foreign-born individuals, taking into account the different shape of the wage distribution for these two groups. This is important to study for the following reasons. First, immigrants are likely to differ from natives in both income distribution and elasticity of labor supply. Second, over the past years, we observe increasing migration to Sweden and many other European countries. The percentage of immigrants in the total population of Sweden is shown in Figure A1, which shows an increasing pattern of the share of immigrants during the last decade. Third, the presence of income tax competition for high skill im-
migrants and compensation for labor market discrimination can be reasons to differentiate taxes between immigrants and natives. For example, several countries have adopted preferential tax schemes to attract high-skill immigrants (see e.g., Kleven et al. 2014). This paper aims to study the following questions. First, what is the optimal tax schedule given the empirical wage data? Second, what are the optimal tax schemes if we account for immigration status?

I calculate optimal income tax rates based on immigration status. The calibrated models employ utilitarian and Rawlsian social welfare functions and heterogeneous labor supply elasticities for different groups. The key primitives for optimal income tax models are individuals’ skills levels, and wage rates are used as a proxy for skill levels in the literature. I access individual-level survey data in 2016, and the sample contains wage information of the population aged 18-66. By using this empirical wage data, I compare calibrated marginal tax rates with the empirical tax scheme. In the standard tagging model, the target group is (on average) a needy group. However, the recipient group is not necessarily a needy group in the immigration-status-based model. I (endogenously) determine the recipient group and the optimal amount of inter-group transfer, i.e., either immigrants or natives can be a recipient group.

I follow the Stiglitz discrete income taxation model (Stiglitz 1982). The discrete model provides more economic insights with a low level of complexity (Guesnerie and Seade 1982). Hellwig (2007) has shown that the finite discrete and continuous models are more or less equivalent. Moreover, simulation evidence shows almost identical patterns between continuous and discrete models as long as the number of types is large (Bastani 2015).

Based on the static nonlinear income taxation model with a utilitarian social welfare function, there are several findings. First, the optimal tax rates from the calibrated model are systematically lower than the empirical tax rates (except for high-income earners, e.g., income levels above the second central government tax kink). The optimal tax rates converge to the real-world tax rates for high and very high income earners. Second, the

\[1\] The wage data is collected by Statistics Sweden and contains information about salary levels, the extent of work, etc. All the public sector employees and half of the private sector employees are surveyed.

\[2\] Bastani calibrates a discrete income model using a set of types from 125 to 2000.
tax system using immigration status will increase marginal tax rates of the majority of natives to a level that is very close to the current empirical tax scheme, while immigrants’ marginal tax rates are systematically lower for all income levels. The magnitude of change is more significant for immigrants. Third, the tax system using immigration status lowers the tax burden for high skilled natives and immigrants. Forth, the optimal inter-group transfer is approximately 0.04% of total consumption (or 0.03% of the total income) of natives. It is optimal for natives to subsidize immigrants with the transfer, but overall, the extent of inter-group redistribution is small.

Tagging models expect the target group to locate at a clear place of the income distribution, as the tags should relate to the earning potentials. Immigrants, however, distribute over the entire income distribution, and immigration status itself reveals no information about the skill level. This further aggravates the violation of horizontal equity in standard tagging models, in which people with the same level of income face different tax rates. Previous researchers simulate optimal tax rates by using a combination of tagging and means-tested systems, and they find that the marginal tax rates are decreasing in income for the richer groups and increasing in income for the poorer groups [Immonen et al. 1998]. The intra-group income redistribution mitigates the violation of the horizontal equity problem.

This paper only studies labor income taxation, and does not consider capital income taxation. The Swedish dual tax system has a flat capital tax rate of 30%. This is likely to affect the calibration of labor supply in my paper, especially for high income earners, because of potential income shifting and tax avoidance. Additionally, I assume quasi-linear preferences, which ignores income effects on labor supply.

The conventional tagging literature assumes that tags are exogenous. This paper uses “place of birth” to determine whether an individual is an immigrant. However, in practice immigration status can be based on citizenship, which is endogenous in the sense that people choose to migrate to Sweden. This “extensive margin” has not been addressed in standard tagging models, and people focus on the implied skill property of the target group. An immigration-status-based tax policy could lead to a self-selection problem: People could choose to become immigrants because of the tax system. This holds especially for workers with low mobility cost (e.g., Nordic or EU cit-
There are many studies related to the endogenous shifting in the design of tax systems. For example, Balestrino et al. (2002) has endogenized fertility in designing optimal household taxation. Scheuer (2014) also investigates the endogenous selection of occupational choice in entrepreneurial taxation, e.g., to be a worker or an entrepreneur. Moreover, based on Piketty and Saez (2013) two tax bases income shifting model, Selin and Simula (2020) have shown that policymakers can increase social welfare contingent on the setting of endogenous tagging. For a more complete study of immigration-status-based income taxation, one can rely on a more realistic model that takes into account the extensive margin of shifting between natives and immigrants.

The rest of the paper proceeds as follows. Section 2 introduces the discrete nonlinear income taxation model and the extension to an immigration-status-based income taxation model. Section 3 provides an empirical calibration of discrete income taxation models. Section 4 presents optimal tax structures. Section 5 concludes.

2 Model

2.1 Discrete income taxation model

In this paper, I use an extension of the two-type Stiglitz (1982) discrete optimal income tax model. The discrete optimization problem faced by the government is the maximization of social welfare of a set of discrete types of agents, subject to an aggregate budget constraint, and a set of incentive compatibility constraints. Agents differ only in skill levels.

Suppose there are two types of agents with distinct productivity that is measured by output per hour in the labor market, e.g., \( w_1 \) and \( w_2 \) are the wages for the low-skill and the high-skill agents respectively. The type \( i \) agent with productivity \( w_i \) produces \( y_i \) units of output (before-tax income) and consumes \( c_i \) units of good (after-tax income) gains the payoff \( u(c_i, y_i, w_i) \).

---

3 According to Swedish Migration Agency, Nordic citizens do not need to apply for a residence permit, and they simply need to register with the Swedish Tax Agency to live in Sweden. A similar rule applies to EU citizens.

4 By fully accounting for extensive margin, Selin and Simula highlight that people who shift easily are also more elastic in labor supply, giving them a lower tax rate can increase both efficiency and equity.
The utility function is specified as

\[ u(c_i, y_i, w_i) = u^i(c_i, y_i). \]

In this specification, \( \frac{u_i}{w_i} \) is the number of working hours for agent \( i \), denoted as \( l_i \). The utility function can be written as a separable function \( u(c_i, l_i) = x(c_i) - v(l_i) \), and \( v(l_i) \) is an increasing convex function of \( l_i \).

An allocation is a pair of before-tax and after-tax income bundles \( \{c, y\} \) contingent on the productivity \( w \). The budget constraint is a function of the before-tax income, \( c = y - T(y) \). The marginal rate of substitution is \( MRS = -\frac{dc}{dy} = 1 - T'(y) \), where \( T'(y) \) is the marginal tax rate. Applying the implicit function theorem we get \( MRS = \frac{\partial u}{\partial y} \frac{\partial c}{\partial c} = \frac{\partial u}{\partial l} \frac{\partial l}{\partial c} \), which implies that indifference curve is steeper for low-skill agents than high-skill agents. The indifference curves thus satisfy the single-crossing condition (or Spence-Mirrlees condition). Moreover, the (implicit) marginal tax rate is defined as \( T'(y) = 1 - \frac{\partial u/\partial y}{\partial u/\partial c} \).

Governments redistribute income and raise revenue such that aggregate production exceeds aggregate consumption. Given that the number of agents of type \( i \) is \( n_i \), the government reaches a feasible allocation if

\[ n_1 c_1 + n_2 c_2 \leq n_1 y_1 + n_2 y_2 - E, \]

where \( E \) is the government expenditures. The skill level is not observable by the government. High-skill individuals can earn the same income as low-skill individuals and enjoy more leisure. One prerequisite of the income redistribution is that no one deviates from his true productivity. Hence, governments need to impose incentive compatibility constraints to prevent mimicking behavior:

\[ u^1(c_1, y_1) \geq u^1(c_2, y_2), \text{ and } u^2(c_2, y_2) \geq u^2(c_1, y_1). \]

Each agent is assumed to behave rationally and choose an allocation that maximizes his utility. Given the single-crossing condition and the incentive compatibility constraints, individuals with different skill levels have different optimal allocations.
The social welfare function can be expressed as
\[ \sum_i G(u^i(c_i, y_i))n_i, \]
where \( G(u^i) \) is a function of utility. The expression is a utilitarian social welfare function when \( G(u^i) = u^i \). Arnott et al. (1987) has shown that given a utilitarian social welfare function and a separable utility function, the only possibility is that the downward incentive compatibility is binding whereas the upward incentive compatibility is not binding.

2.2 Immigration-status-based income taxation model

Tax schedules can be customized for different groups to increase the efficiency of the tax system. More specifically, individuals can be identified as either natives or immigrants. In this setting, the social welfare maximization problem is a two-part problem. The first part is the intra-group income redistribution, i.e., the maximization of a social welfare function subjects to a budget constraint and incentive compatibility constraints of each group. The second part is the inter-group redistribution, i.e., the budget for each group is optimally determined (assuming there is no government expenditure). The immigration-status-based income taxation model can be described as follows. Suppose an economy in which agents differ in both skill levels and place of birth. I denote \( w_{ij} \) as the productivity for the type \( i \) agent in group \( j \), where \( j = A \) \((j = B)\) if the agent is a native \( (\text{an immigrant}) \). The utility function can be expressed as
\[ u(c_{ij}, y_{ij}, w_{ij}) = u^{ij}(c_{ij}, y_{ij}). \]

We need a separate budget for each group to study this two dimensional problem, since the redistribution is not only between different skill types but also between natives and immigrants. The budget constraints for natives

\[ \frac{du}{dc_2} - \frac{du}{dy_2} = 0, \]
\[ \frac{du^2}{dy_2} = -1, \]

the marginal tax rate for the high-skill agent is 0. We can further show that the low-skill agent has a marginal tax rate between 0 and 1. This result is inline with Mirrlees (1971) continuous nonlinear income taxation model.
and immigrants are,

$$\sum_i n_{iA}(y_{iA} - c_{iA}) \geq E_A \quad \text{and} \quad \sum_i n_{iB}(y_{iB} - c_{iB}) \geq E_B.$$ 

The government reaches a feasible allocation if $E_A + E_B \geq E$, where $E$ is the (exogenous) amount of spending. I assume that government expenditure is 0, such that budgets from two groups are pure inter-group redistribution. The inter-group transfer can be expressed as $E_A = -E_B = E$. The optimal inter-group transfer is endogenously determined. This does not imply that immigrants (or natives) are paying lower tax rates than another group at all income levels, there is also a redistribution within each group, see [Immonen et al.] (1998).

Since the immigration status (place of birth) is exogenous and observable, only the intra-group redistribution requires incentive compatibility constraints. The tax system needs to be incentive compatible within each group for the same reason as in the nonlinear income taxation model, i.e., governments cannot observe individuals’ skill levels. The incentive compatibility constraints can be expressed as

$$\forall i \geq 2 : u^{ij}_i(c_{ij}, y_{ij}) \geq u^{ij}_{i-1}(c_{i-1,j}, y_{i-1,j}), \quad \text{and} \quad u^{ij}_i(c_{ij}, y_{ij}) \geq u^{ij}_{i+1}(c_{i+1,j}, y_{i+1,j}).$$

The group-wise taxation reduces the number of incentive compatibility constraints, as there are no inter-group constraints. [Hellwig] (2007) has shown that the upward incentive compatibility constraints can be substituted with the consumption monotonicity constraints in a ”weakly relaxed income tax problem”. Hence, I write the incentive constraints as

$$\forall i \geq 2 : u^{ij}(c_{ij}, y_{ij}) \geq u^{ij}_{i-1}(c_{i-1,j}, y_{i-1,j}), \quad \text{and} \quad c_{ij} \geq c_{i-1,j}.$$ 

The general form of social welfare function is

$$\sum_i \sum_j G(u^{ij}_i(c_{ij}, y_{ij}))n_{ij},$$

where $G(u^{ij})$ assigns welfare weights to individuals and reflects the governments’ aversion towards inequality.
3 Empirical calibration

In this section, I present the calibration of the nonlinear income taxation model and the immigration-status-based income taxation model. The key elements of the models are introduced, e.g., wage rates, preferences, and the social welfare function. The calibrated models employ utilitarian and Rawlsian welfare functions, and homogeneous and heterogeneous labor supply elasticities across and within groups. The baseline model uses a utilitarian social welfare function and a homogeneous labor supply elasticity for all individuals.

3.1 Data

I access register and survey data from 2016. The data is provided by Statistics Sweden. The two variables of interest are immigration status (place of birth) and wage rates. The register data provides demographic information for the entire population, e.g., age, gender, income, and place of birth. The survey data covers the salary structure of the entire labor market, which includes data from the central government, county councils, primary municipalities and the private sector. All the data from the public sector is collected. By collecting wage information from 4% of all companies/organizations/foundations, about 50% of the employees in the private sector are surveyed. The survey contains information on salary levels, the extent of work, and industry code.

The wage variable is constructed by dividing the monthly salary by the extent of work. A full-time worker’s labor supply is 174 hours per month. Figure 1 shows kernel density estimates of wage distributions for immigrants and natives. Many suggest that income distributions should follow a lognormal distribution with a Pareto tail (Saez 2001; Mankiw et al. 2009). The

---

6 The sample is stratified based on 83 industry groups and 8 company size classes. Companies with 500 and more employees are fully surveyed, and nearly 2% of the smallest companies (1-9 employees) are surveyed.

7 In general, the maximum working hour during a year is 5840 hours, i.e., 16 hours per day times 365 days. Based on The Working Hours Act (Arbetstidslagen) in Sweden, a regular working week should not exceed 40 hours that is equivalent to a maximum 2088 hours per year. The maximum labor supply of each month should not exceed $40 \times \left(\frac{365}{7}/12\right) \approx 174$ hours.

8 This is a “zoom in” figure that excludes the top 5% of observations. The right tails of two distributions are almost identical, and large extreme outliers are natives.
lognormal parametric estimation of wage distribution for the entire population, natives and immigrants are \((\mu, \sigma) = (5.33, 0.51)\), \((\mu_{native}, \sigma_{native}) = (5.34, 0.51)\) and \((\mu_{immigrant}, \sigma_{immigrant}) = (5.28, 0.53)\) respectively.\(^9\) The immigrants’ wage distribution is more dispersed, that is, most of the immigrants locate at the bottom 40% and the top 10% of the wage distribution, see Table 1. The survey contains a sample of individuals aged 18-66. I exclude data with zero working hours from the sample (approximately 1% of the total sample) due to the missing hourly wage information.\(^10\)

### 3.2 Construction of the discrete skill distributions

There are two ways to construct discrete wage rates from a wage distribution. Bastani (2015) constructs wage rates by dividing the entire distribution into a set of equidistant wage rates, and the distance between any two consecutive wage rates is fixed. I instead divide the entire distribution into a

---

\(^9\)The four best estimations that fit the wage distributions are generalized extreme value, t location-scale, log-logistic and lognormal distributions (in descending order).

\(^10\)The non-participating observations have 0 working hours, and the possible reasons are sick leave, parental leave, holiday, severance and so on.
set of equal proportional bins, such that the number of individuals at different representative wage rates are the same. Each representative wage is the median value of each bin. The representative wage rates in each group are not necessarily equivalent by using the “equal proportion” approach, i.e., $w_{iA} \neq w_{iB}$. This does not affect the general shapes of the calibrated tax schemes, since the difference between $w_{iA}$ and $w_{iB}$ is expected to be small for the majority of wage earners (e.g., those who are not outliers). I denote the ratio between the share of natives and the share of immigrants as $r \equiv \gamma_A / \gamma_B$, where $\gamma_A$ and $\gamma_B$ are the total population of natives and immigrants. The proportion of each type of natives and immigrants are $n_{iA} = \frac{\gamma_A}{k} = \frac{r}{(1+r)k}$ and $n_{iB} = \frac{\gamma_B}{k} = \frac{1}{(1+r)k}$, where $k$ is the number of skill types. The number of representative wage rates I use is $k = 2000$. The ratio of population weights is $r = 4.55$ in the simulation. 

### 3.3 Preferences

I adopt the following utility function:

$$u(c, l) = c - \frac{\alpha}{1 + 1/e} l^{1+1/e},$$

where $c$ is the after-tax income, and $l$ is the number of hours of labor supply. The before-tax income is $y = wl$. Each individual faces a nonlinear budget constraint $c = y - T(y)$. The quasi-linear form of utility function excludes the income effect so that the choice of labor supply depends purely on the tax rate. The optimal choice of labor supply is

$$-\alpha l^{\frac{1}{e}} + \lambda w(1 - T'(wl)) = 0,$$

where $\lambda$ is the Lagrange multiplier which measures how much the indirect utility would change given a one Swedish krona increase in the consumption. The labor supply elasticity is

$$\frac{d \log l}{d \log w} = e.$$
There are many discussions about the labor supply elasticity and elasticity of taxable income (also elasticity of total income, elasticity of earned income, etc.). In particular, the labor supply elasticity measures how the change in the wage rate affects the labor supply, while elasticity of taxable income captures all behavioral responses to changes in marginal tax rates (e.g., career paths, income shifting, etc.), see Saez et al. (2012). In my simple model, taxable income is only a function of labor supply. The labor supply elasticity and the elasticity of taxable income are equivalent in this setting.\textsuperscript{13} According to the previous literature, a reasonable assumption regarding the labor supply elasticity is around 0.1 to 0.3. For example, Chetty et al. (2011) suggest the Hicksian elasticity for empirical calibration is 0.25. I choose a elasticity $e = 0.2$ for all individuals as the baseline case. The survey data contains mainly public sector and large private sectors, and individuals are less able to change their contract working hours in response to the change of income.

<table>
<thead>
<tr>
<th>Case I: Increasing profile of elasticities</th>
<th>Bottom 40%</th>
<th>Middle 50%</th>
<th>Top 10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natives</td>
<td>0.12</td>
<td>0.2</td>
<td>0.5</td>
</tr>
<tr>
<td>Immigrants</td>
<td>0.1</td>
<td>0.2</td>
<td>0.6</td>
</tr>
<tr>
<td>All</td>
<td>0.116</td>
<td>0.2</td>
<td>0.518</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case II: Decreasing profile of elasticities</th>
<th>Bottom 40%</th>
<th>Middle 50%</th>
<th>Top 10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natives</td>
<td>0.3</td>
<td>0.15</td>
<td>0.1</td>
</tr>
<tr>
<td>Immigrants</td>
<td>0.28</td>
<td>0.15</td>
<td>0.12</td>
</tr>
<tr>
<td>All</td>
<td>0.296</td>
<td>0.15</td>
<td>0.104</td>
</tr>
</tbody>
</table>

| Share of Immigrants                        | 20%        | 0.13%      | 0.18%  |

Notes: Share of Immigrants measures the proportion of immigrants in the total population in different parts of the wage distribution. The baseline elasticity is 0.2 for all individuals. The average elasticity is 0.2 in all cases.

Table 1: Labor Supply Elasticities for Different Wage Group

The baseline case assumes a homogeneous elasticity for all individuals, regardless of their skill levels or immigration status. Compared to the sim-

\[ u(c, y) = c - \frac{a}{1 + \frac{y}{w}} (\frac{y}{w})^{1+1/e}. \]

First-order optimality condition of $u(c, y)$, subject to a nonlinear budget constraint $c = y - T(y)$, gives us $\frac{w^{e+1}}{a} (1 - T'(y)) = y$. The elasticity of taxable income is $\frac{\partial \log y}{\partial \log (1 - T'(y))} = e$.\textsuperscript{14}

\[ \frac{\partial \log y}{\partial \log (1 - T'(y))} = e. \]
ple (but unrealistic) baseline case, I assumed different elasticities for different wage levels, see Table [1]. Elasticities either increase or decrease in wage levels.

Labor supply elasticity might vary across and within the groups of natives and migrants, as immigrants might have a different composition of the labor force, e.g., less female workers or more young workers. Blundell et al. (2013) have shown that the labor supply responses differ systematically by age, gender and family composition. Natives and immigrants are assumed to have different elasticities at the bottom and top wage levels. The average elasticity for natives and immigrants are the same, and equal to the baseline case, i.e., 0.2. The labor supply elasticity for the single tax system is based on the share of immigrants in each wage segment.  

3.4 Calibration of the labor supply

In this section, I describe how I calibrate labor supply. The first-order optimality condition of the individual maximization problem gives

\[
\frac{\partial u(c,l)}{\partial l} = w - T'(wl)w - \alpha \frac{l^{1/\epsilon}}{\epsilon} = 0 ,
\]

\[
\left( \frac{w(1 - T'(wl))}{\alpha} \right)^\epsilon = l.
\]

I calibrate labor supply to match empirical working hours. The calibration uses the median wage earner’s labor supply as the benchmark. On average, the median wage earner works 164 hours per month and earns nearly 191 SEK per hour. The labor supply is normalized such that \( l = 1 \) corresponds to 164 hours. The calibration of labor supply depends on the elasticity of labor supply and the marginal tax rates as seen from equation (1). With the aid of the scale parameter \( \alpha \), the labor supply of median wage earners can be normalized to 1. Hence, we find that \( \alpha \) should satisfy,

\[
191(1 - T'(wl)) = \alpha.
\]

There is a one-to-one mapping between \( \alpha \) and the marginal tax rate. One reasonable set of marginal tax tax rates for the calibration is the set of empir-
ical tax brackets (excluding the non-convex kink) in 2016, see Table 2. I loop over empirical marginal tax rates to find an $\alpha$ that normalizes median wage earners’ labor supply to 1. The calibration of labor supply for different tax brackets is shown in Table 2.

<table>
<thead>
<tr>
<th>Marginal tax rate</th>
<th>Utilitarian $\alpha$</th>
<th>$l$ (median)</th>
<th>Rawlsian $l$ (median)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>190.80</td>
<td>0.939</td>
<td>0.788</td>
</tr>
<tr>
<td>0.21</td>
<td>150.732</td>
<td>0.985</td>
<td>0.826</td>
</tr>
<tr>
<td>0.29</td>
<td>135.468</td>
<td>1.006</td>
<td>0.844</td>
</tr>
<tr>
<td>0.32</td>
<td>129.744</td>
<td>1.014</td>
<td>0.851</td>
</tr>
<tr>
<td>0.52</td>
<td>91.584</td>
<td>1.088</td>
<td>0.913</td>
</tr>
<tr>
<td>0.55</td>
<td>85.86</td>
<td>1.102</td>
<td>0.925</td>
</tr>
<tr>
<td>0.6</td>
<td>76.32</td>
<td>1.128</td>
<td>0.947</td>
</tr>
<tr>
<td>0.7</td>
<td>57.24</td>
<td>—</td>
<td>1.003</td>
</tr>
</tbody>
</table>

Notes: Since median income earners face higher marginal tax rates in a Rawlsian social welfare function than in a utilitarian setting, I borrow a pseudo-marginal tax rate of 70%.

Table 2: Calibration of Labor Supply

3.5 Choice of the social welfare functions

I employ Bergson-Samuelson and Rawlsian social welfare functions. The Bergson-Samuelson social welfare function is a concave transformation of the utilitarian welfare function, and it can be expressed as

$$W = \sum_{i}^{k} \log(u_{iA}(c_{iA}, y_{iA})) + \sum_{i}^{k} \log(u_{iB}(c_{iB}, y_{iB})).$$

The above can be viewed as a weighted utilitarian social welfare function, and the welfare weight for the $i$th agent in group $j$ is $\frac{1}{u^{j}(c_{ij}, y_{ij})}$. Compared to the (weighted) utilitarian objective function, the Rawlsian social welfare function captures an extreme case of inequality aversion, where governments only desire to maximize the welfare of the worst-off individual. The least well-off natives and immigrants have productivities $w_{1A}$ and $w_{1B}$, see Section 2. One additional constraint needs to be satisfied to employ a Rawls-
sian welfare function,

\[ u^{1A}(c_{1A}, y_{1A}) = u^{1B}(c_{1B}, y_{1B}). \]

Provided that the above constraint is satisfied, the Rawlsian social welfare function can be written as:

\[ W = \max u^{1A}(c_{1A}, y_{1A}). \]

## 4 Optimal tax structures

Table 3 shows optimal marginal tax rates for different income levels (percentiles) and optimal inter-group transfers. It is optimal for natives to subsidize immigrants, while the optimal amount of inter-group transfer is relatively small in percentage terms. For example, the optimal inter-group transfer is approximately 0.04% of the total consumption (or 0.03% of the total income) of natives in the baseline case. The optimal tax schemes are illustrated graphically for all income levels in Figure 2 and Figure A2-A4 in the Appendix. The marginal tax rate for the highest income earner reaches zero in all cases, which is not shown in the figures. The empirical piece-wise linear tax rates are those that apply to earned income, excluding social transfers.\(^{16}\)

Figure 2 depicts the general shape of the optimal tax schemes for a utilitarian social welfare function. Compared to the uniform optimal tax scheme, the tax system using immigration status lowers the tax burden for top income earners, i.e., both natives and immigrants, see Figure 2a. The calibrated tax rates converge to the real-world tax rates for top income earners.

The comparison of the optimal and empirical marginal tax rates are shown in Figure 2b.\(^{17}\) The optimal marginal tax rate is increasing in income, and its shape is close to the empirical (progressive) tax scheme. The second central government tax kink in 2016 is at the income level, e.g., SEK 638,800, and there are 6% of the population who face a higher tax rate above the kink. This implies that the empirical piece-wise linear tax rates for the bot-

---

\(^{16}\)The tax calculator for empirical marginal tax rates is provided by Håkan Selin. The calibrated income is also a pure function of wages without social transfers.

\(^{17}\)The “zoom in” figure covers the bottom 99.8% of the income earners.
(a) Optimal marginal tax rates

(b) Comparison between optimal and empirical marginal tax rates

Figure 2: Optimal Tax Rates. Utilitarian Social Welfare Function. 2016
<table>
<thead>
<tr>
<th>Percentile</th>
<th>All</th>
<th>Natives</th>
<th>Immigrants</th>
</tr>
</thead>
<tbody>
<tr>
<td>10th</td>
<td>0.207</td>
<td>0.270</td>
<td>0.049</td>
</tr>
<tr>
<td>30th</td>
<td>0.222</td>
<td>0.275</td>
<td>0.076</td>
</tr>
<tr>
<td>50th</td>
<td>0.262</td>
<td>0.347</td>
<td>0.133</td>
</tr>
<tr>
<td>70th</td>
<td>0.308</td>
<td>0.411</td>
<td>0.264</td>
</tr>
<tr>
<td>90th</td>
<td>0.543</td>
<td>0.600</td>
<td>0.532</td>
</tr>
<tr>
<td>99th</td>
<td>0.708</td>
<td>0.693</td>
<td>0.631</td>
</tr>
<tr>
<td>Inter-group transfer</td>
<td>0.04% (0.03%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Rawlsian; $\varepsilon = 0.2$**

<table>
<thead>
<tr>
<th>Percentile</th>
<th>All</th>
<th>Natives</th>
<th>Immigrants</th>
</tr>
</thead>
<tbody>
<tr>
<td>10th</td>
<td>0.879</td>
<td>0.872</td>
<td>0.839</td>
</tr>
<tr>
<td>30th</td>
<td>0.748</td>
<td>0.740</td>
<td>0.704</td>
</tr>
<tr>
<td>50th</td>
<td>0.688</td>
<td>0.705</td>
<td>0.690</td>
</tr>
<tr>
<td>70th</td>
<td>0.650</td>
<td>0.672</td>
<td>0.711</td>
</tr>
<tr>
<td>90th</td>
<td>0.731</td>
<td>0.737</td>
<td>0.796</td>
</tr>
<tr>
<td>99th</td>
<td>0.758</td>
<td>0.728</td>
<td>0.734</td>
</tr>
<tr>
<td>Inter-group transfer</td>
<td>0.0005% (0.0005%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Utilitarian; Increasing elasticity profile**

<table>
<thead>
<tr>
<th>Percentile</th>
<th>All</th>
<th>Natives</th>
<th>Immigrants</th>
</tr>
</thead>
<tbody>
<tr>
<td>10th</td>
<td>0.240</td>
<td>0.301</td>
<td>0.071</td>
</tr>
<tr>
<td>30th</td>
<td>0.259</td>
<td>0.309</td>
<td>0.108</td>
</tr>
<tr>
<td>50th</td>
<td>0.224</td>
<td>0.303</td>
<td>0.117</td>
</tr>
<tr>
<td>70th</td>
<td>0.281</td>
<td>0.388</td>
<td>0.262</td>
</tr>
<tr>
<td>90th</td>
<td>0.591</td>
<td>0.598</td>
<td>0.596</td>
</tr>
<tr>
<td>99th</td>
<td>0.578</td>
<td>0.547</td>
<td>0.463</td>
</tr>
<tr>
<td>Inter-group transfer</td>
<td>0.07% (0.05%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Utilitarian; Decreasing elasticity profile**

<table>
<thead>
<tr>
<th>Percentile</th>
<th>All</th>
<th>Natives</th>
<th>Immigrants</th>
</tr>
</thead>
<tbody>
<tr>
<td>10th</td>
<td>0.170</td>
<td>0.231</td>
<td>0.044</td>
</tr>
<tr>
<td>30th</td>
<td>0.192</td>
<td>0.299</td>
<td>0.066</td>
</tr>
<tr>
<td>50th</td>
<td>0.337</td>
<td>0.420</td>
<td>0.168</td>
</tr>
<tr>
<td>70th</td>
<td>0.389</td>
<td>0.475</td>
<td>0.318</td>
</tr>
<tr>
<td>90th</td>
<td>0.695</td>
<td>0.746</td>
<td>0.502</td>
</tr>
<tr>
<td>99th</td>
<td>0.749</td>
<td>0.721</td>
<td>0.610</td>
</tr>
<tr>
<td>Inter-group transfer</td>
<td>0.07% (0.05%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The optimal inter-group transfer goes from natives to immigrants in all cases, and it is calculated as the proportion of total consumption (total consumption in parentheses) of natives.

Table 3: Optimal Marginal Tax Rates

Tom 94% of the population are higher than the optimal marginal tax rates. Moreover, optimal tax schemes suggest that top income earners should face even higher tax rates than the current tax rates (except the highest income earners). Interestingly, the tax system using immigration status increases the
marginal tax scheme of the majority of natives to a level that is very close to the current empirical tax scheme. In contrast, immigrants’ marginal tax rates are now systematically lower for all income levels. The magnitude of change is more significant for immigrants.

With the Rawlsian social welfare function, governments have the strongest aversion to inequality. Figure A2 shows a U-shaped marginal tax scheme covers most income earners, and a decreasing trend of marginal tax rates follows it. The high marginal tax rates at low-income levels allow low-income earners to receive a transfer while preventing high-income earners from working less and claiming those transfers. In general, U-shaped marginal tax rates start to increase at the income level a little below the second central government kink. An immigration-status-based income tax system lowers the marginal tax rates for top income earners, see Table 3. There is no big change in marginal tax rates for the majority of natives (except the top income earners) when allowing tax rates to depend on immigration status. In contrast, the marginal tax rates increase for high-income immigrants.

It is well-known that the shape of the optimal marginal tax rates is sensitive to the labor supply elasticity. To see the effect of the choice of labor supply elasticity, I, ceteris paribus, employ heterogeneous elasticities in calibrated models that are displayed in Table 1. Figure A3 and Figure A4 illustrate two different elasticity scenarios across income levels for natives and immigrants. One clear expected pattern is lower marginal tax rates for the more elastic group. Compared to the baseline case, the optimal tax model with increasing elasticities across income levels does not preserve the general shape of the marginal tax rates, for example, the non-convex budget in Figure A3.

5 Conclusion

I use a numerical simulation approach to derive optimal marginal tax rates for natives and immigrants in Sweden. I exploit survey and register data to construct empirical variables, which are immigration status and wage rates. The optimal tax models employ utilitarian and Rawlsian social welfare functions, and homogeneous and heterogeneous labor supply elasticities across and within groups. The baseline model employs a utilitarian
social welfare function and a utility function that is quasi-linear in consumption, and a homogeneous labor supply elasticity for all individuals.

My most important results are the following. First, optimal tax rates are systematically lower than the empirical tax rates except for the very high income earners. The tax system that takes into account the immigration status lowers the tax burden for top income earners, i.e., both natives and immigrants. The tax system using immigration status will increase the marginal tax rates of the majority of natives to a level that is very close to the current empirical tax scheme. In contrast, immigrants’ marginal tax rates are systematically lower for all income levels. The optimal inter-group transfer is approximately 0.04% of total consumption (or 0.03% of the total income) of natives. A model with a Rawlsian social welfare function provides a U-shaped marginal tax scheme. The calibrated marginal tax rates are sensitive to the labor supply elasticities that are in line with the “inverse-elasticity” intuition, i.e., high elasticities tend to imply lower optimal marginal tax rates.
References


Appendix

Figure A1: Share of Immigrants, Sweden 2007-2016
Figure A2: Optimal Tax Rates. Rawlsian Social Welfare Function. 2016
(a) Optimal marginal tax rates

(b) Comparison between optimal and empirical marginal tax rates

(a) Optimal marginal tax rates

(b) Comparison between optimal and empirical marginal tax rates

**Figure A4:** Optimal Tax Rates. Utilitarian Social Welfare Function. Elasticity decreases in income. 2016
Earnings Responses to Even Higher Taxes*

Dingquan Miao†, Håkan Selin‡, and Martin Söderström §

Abstract

Sweden introduced a phase-out of the earned income tax credit in 2016. As a consequence, taxpayers belonging to the top 5 percent of the earnings distribution, already facing high taxes, experienced a 7% reduction in their net-of-tax shares. While exploiting rich full-population administrative data up to 2017, we evaluate earnings responses to the reform. When graphically and econometrically comparing earnings growth at different segments of the distribution, we estimate a significant relative earnings reduction in the treatment group immediately appearing in 2016, and growing in 2017. The implied earnings elasticity is fairly low and around 0.1. We interpret the essential features of the response using a simulation model, in which people have noisy perceptions of the piece-wise linear tax code. To simulate the empirically observed response, we need to add more noise to perceptions than what is motivated by earnings uncertainty alone.

Keywords: Earnings supply, Income taxation.

JEL Classification: H24; J22.

---

*We thank Spencer Bastani, Tobias König, and seminar participants at the Linnaeus University in Växjö, IFAU, and Umeå School of Business, Economics and Statistics for comments. This is a preliminary version of an ongoing and unpublished IFAU project.

†Department of Economics and Statistics, Linnaeus University. Email: Dingquan.Miao@lnu.se

‡Institute for Evaluation of Labour Market and Education Policy (IFAU). Email: hakan.selin@ifau.uu.se

§Institute for Evaluation of Labour Market and Education Policy (IFAU). Email: Martin.Soderstrom@ifau.uu.se
1 Introduction

In progressive tax systems a disproportionately large share of tax revenues are collected from high-income earners. Needless to say, the earnings elasticity at the top of the distribution has far-reaching implications for the optimal taxation of labor incomes. In particular, top income earners’ sensitivity to taxes is a key determinant of the revenue maximizing tax rate in the top bracket (“the peak of the Laffer curve”) – a low elasticity \textit{ceteris paribus} implies a high revenue-maximizing tax rate and vice versa. It is often pointed out that a top tax rate exceeding this level is socially inefficient, because the government could then make everyone better off by lowering the tax rate.

In this article we exploit a unique opportunity to learn about top earners’ responses to higher taxes. In 2015, Swedish high income earners faced an effective marginal tax rate of around 73% when accounting for both direct and indirect taxes on labor and consumption. This figure is high in international comparison, and it happens to coincide with the revenue maximizing tax rate computed by Diamond and Saez (2011, p.171) for the U.S.\footnote{This follows from an elasticity of 0.25 and a Pareto parameter of 1.5. The right tail of the Swedish income distribution is thinner and, \textit{ceteris paribus}, the Swedish revenue maximizing tax rate is lower. But the value of the Pareto parameter in Sweden tends to be highly sensitive to the inclusion of capital incomes. Several economists who analyzed the Swedish income tax system and income distribution in the pre-2016 period concluded that the effective tax rate applying to high incomes was either very close to or above the peak of the Laffer curve, see Holmlund and Söderström (2011), Pirttilä and Selin (2011), and Sørensen (2010).} In 2016, Sweden introduced a phase-out of the earned income tax credit (EITC) at high income levels. Consequently, taxpayers in the top 5 percentile groups of the earnings distribution experienced a 7% reduction in their net-of-tax shares. We evaluate this reform by comparing treated and non-treated percentile groups of the earnings distribution.

In several respects, the Swedish EITC phase-out reform provides a promising source of quasi-experimental variation for the purpose of estimating the earnings elasticity, i.e. the percentage change in before tax earnings in response to a percentage change in the net-of-share (1-marginal tax rate). First, the Swedish earnings distribution has been surprisingly stable since the turn of the millennium. More specifically, we will demonstrate that earnings growth evolved similarly in the treatment and control groups in the years preceding the reform. Second, the reform brought about an iso-
lated policy change – an introduction of a new tax bracket, while essentially leaving other relevant aspects of the tax system intact. Historically, many tax reforms combined tax rate changes with changes to the tax base, and it has often been challenging for researchers to separate between the two (Kopczuk, 2005). Third, we access population-wide administrative data up to 2017, and the richness of our data allows us to control for a large number of factors, and we are also able to examine potential shifting between the labor and capital income tax bases.

Since gross earnings (before deductions) is the base for the EITC, we are estimating an earnings rather than a taxable income elasticity. One could claim that the earnings elasticity is more interesting from the perspective of optimal taxation since it has a stronger connection to real behavior (labor supply and effort responses). Tax planning responses, which are captured by the taxable income elasticity, are arguably local to the tax function in place (Slemrod and Kopczuk, 2002).

Still, from a methodological perspective our study is part of the taxable income literature. There is no consensus on how to estimate the elasticity of taxable income (ETI). While the pioneering study by Lindsey (1987) was conducted on cross-sectional tax return data, the main approach in the ETI literature since Feldstein (1995) has been to estimate difference-in-difference models using individual level panel data. Identification comes from tax reforms treating different income groups differentially. However, unlike most modern applications of the difference-in-difference estimator, ETI papers often lack graphical evidence that corresponds to the parameters being estimated. Certainly, the absence of such graphs can be explained by the fact that taxpayers typically are assigned to treatment and control groups based on pre-reform income, and individual incomes vary stochastically from one year to another for non-tax reasons, especially at the very top of the distribution.

According to the terminology of the ETI literature, panel data methods must account for the “mean reversion problem”. This challenge has led to interesting econometric proposals, see e.g. Blomquist and Selin (2010), Matikka (2018), who used municipal variation in Finland, and Burns and Ziliak (2017), who exploited state level variation in the U.S.

A similar point was made by Kleven and Schultz (2014). In Section 4.3 below, we argue that our setting is even more suitable for an informative graphical analysis.
Holmlund and Söderström (2011), Kumar and Liang (2017), and Weber (2014). But the methodological advances have come with the cost of lost transparency. Actually, when Saez et al. (2012) surveyed the ETI literature a couple of years ago, the superiority of the panel data approach was questioned with reference to the mean reversion problem. And a recent simulation study by Aronsson et al. (2017) confirmed that panel data estimators of the ETI are highly sensitive to the modeling of the stochastic income process. In this article we instead rely on a simple and transparent cross sectional approach. We compare earnings growth in different percentile groups over time, and we discuss potential disadvantages and alternatives using our rich data. Our full population data of course also contain a panel element, which we exploit in different ways.

When graphically and econometrically comparing earnings growth at different segments of the distribution, we estimate a significant relative earnings reduction in the treatment group immediately appearing in 2016, and growing in 2017. The implied earnings elasticity is fairly low and around 0.1. The effect is not driven by income shifting of corporate owners between the labor and capital income tax bases.

A justified question when examining this quasi-experiment is whether individuals’ perceptions of the tax system are sufficiently precise for a clear response to appear when applying our methodology. Actually, in the empirical analysis we find that earnings growth slows down also in the control group, and we find that the effect increases when omitting percentile groups around the treatment cut-off. A tentative interpretation would be that perceived incentives change gradually rather than sharply around the new kink. We have therefore built a simulation model that allows us to interpret these aspects numerically.

Up to now, earnings dynamics has been discussed in terms of a statistical problem in the ETI literature. But the stochastic element of earnings may also impact the underlying behavioral economic model. When people facing non-linear tax schedules cannot perfectly control their incomes, perceptions of marginal incentives will be noisy as well, because they do not know the

---

Saez et al. (2012, p.29) remarked that “the advantage of longitudinal analysis relative to repeated cross-section analysis has been somewhat exaggerated in the empirical literature following Feldstein (1995), especially when one wants to analyze tax changes happening primarily at the top of the income distribution. In some contexts, repeated cross-section analysis or sharebased time-series analysis may be a more robust and transparent approach.”
tax they will pay on their realized income, even if they are fully aware of the tax code. Following Saez (1999) we hypothesize that taxpayers maximize expected utility under \textit{ex ante} uncertainty about \textit{ex post} realized earnings. This model framework has earlier been used to explain why people do not bunch at kinks, but it has not been used to simulate broader responses to tax changes.

We use our individual level panel data on earnings to determine the standard deviation of shocks to earnings. When simulating tax responses using this value, we are actually quite far from recovering the empirically observed response. Maybe this was to be expected, because the model assumes that taxpayers are fully aware of, and act according to, the true piece-wise linear tax schedule, an hypothesis that has been questioned in recent years.\footnote{Liebman and Zeckhauser (2004) and Rees-Jones and Taubinsky (2019) claim that people rely on simple mental heuristics rather than perfect information when responding to complex non-linear tax schedules. In particular, people tend to confuse marginal and average tax rates. These studies, however, abstract from earnings uncertainty as a source of misperception.} Interestingly, when doubling the standard deviation of the noise term, while assuming an underlying elasticity of 0.2, we obtain simulated elasticities that are very close to the empirically estimated elasticities.

If our results reflect long-run responses to taxes the earnings elasticity is fairly low in Sweden. It is important to keep in mind, however, that we only considered two post-reform years, and we cannot exclude that the response is larger (or smaller) in the third post-reform year.

The rest of the paper is organized as follows. In Section 2 we describe the Swedish tax system, and we discuss various aspects of the EITC phase-out reform. Section 3 provides a brief account of the data source. We report the empirical analysis in Section 4. In Section 5 we move on to the simulation model. Section 6 finally, concludes the paper.

\section{Institutional setting}

\subsection{The Swedish system}

The Swedish income tax system is a dual income tax system in which labor earnings and capital incomes are taxed separately. Additionally, the income tax is individual based rather than family based, i.e. spouses are
taxed separately. All individuals aged up to 65 essentially face the same tax schedule, with some variations in the local tax rate. The basic structure of Swedish labor income taxation is fairly simple. A proportional local tax rate applies to the sum of all earned income and taxable transfers (net of some deductions). The average local income tax rate (unweighted) in 2016 was 32.1 percent. For total labor incomes exceeding a certain threshold (SEK 443,200 in 2016) a central government income tax is due. 17 percent of the population aged 20-65 paid the central government income tax in 2016. The central government income tax schedule consists of two brackets; the marginal tax rates in each bracket are 20 percent (for incomes between 443,200 and 638,800 in 2016) and 25 percent (for incomes above 638,800 in 2016) respectively.

The Swedish Earned Income Tax Credit (EITC) was first introduced in 2007 by a center-right wing government coalition, see Edmark et al. (2016) for more details. The base for the EITC is not identical to the base for the local and central government tax, because the EITC is solely a function of earned income. The tax reduction is not granted for social transfers (like unemployment insurance and sickness insurance). In a stepwise fashion, the EITC has become more generous since 2007, and in 2016 the maximum tax credit was around SEK 26,500. The EITC slightly varies with the local tax rate. The Swedish EITC is very general: all individuals aged below 66 face the same tax credit scheme, regardless of marital status or number of children in the household.

2.2 The 2016 reform

It is a standard feature of in-work tax credit policies that the tax credit tapers off when earnings rise. This was not the case, however, in Sweden until 2016, when a left wing-green government reformed the EITC schedule.

---

6The legal term in Swedish is “fastställd förvärvsinkomst”. The income tax is assessed on the basis of yearly income, and the tax year coincides with the calendar year. A basic allowance affects marginal tax rates at lower and medium incomes. The basic allowance does not affect marginal tax rates for the income groups we study in this paper.

71 USD ⇡ 10 SEK.

8On January 1, 2020, the top bracket will be abolished, and the central government tax will be 20 percent on all total labor incomes exceeding the first central government kink point.

9The legal Swedish term is “Skattereduktion för förvärvsinkomster”, but the tax credit is typically referred to as “jobbskatteavdraget”, also in the information from the Swedish tax agency.
While the tax credit in other countries, e.g. the US and the U.K., is phased out at relatively low earnings levels, the Swedish phase-out impacts taxpayers at the upper end of the earnings distribution. In the 2016 reform, the reduction rate was set to 3 percent. Accordingly, if the individual increases her income by SEK 100 she forgoes SEK 3 in tax credit.

Figure 2a visualizes the EITC schedule in 2016 with and without the phase-out. We have plotted the compressed Swedish earnings distribution in the background. Evidently, the phase-out impacts work incentives for a significant number of high-income earners. In Figure 2b we visualize the effect of the EITC phase-out on the marginal tax schedule, again with the earnings distribution in the background. The new EITC kink was placed just below the second central government kink. In a sense, the 3 percentage point increase in the marginal tax rate appears small. But then one should keep in mind that the marginal tax rate in this region was at a very high level already before the reform, 0.57. Hence, the percentage change in the net-of-tax rate is \[ \frac{0.03}{1-0.57} \approx 7\% \]. The reform also implied that a kink point was created at the income level at which the entire tax credit had been phased out. At this (non-convex) kink, located at around SEK 1.5 million in annual earnings, the marginal tax rate decreased by 3 percentage points. Figure 2 illustrates that very few taxpayers – not more than 0.2 percent of the population – earned incomes above this point. It is well-known since earlier tax studies that the right tail of the Swedish earnings distribution is thin.

We wish to emphasize that Figure 2b plots the marginal tax schedule under the assumption that the individual does not receive any taxable transfers, e.g. sickness insurance benefits or parental insurance benefits. Remember that earned income is the base for the EITC, whereas the sum of earned income and taxable transfers is the base for local and central government taxes. Consequently, EITC varies more with sickness absence and parental leave spells. This should be kept in mind when we in Section 5 model uncertainty in earnings realizations.

\[ \text{10If one also takes payroll and consumption taxes into account, the level of marginal tax rates was even higher after the reform, around 75\% (Lundberg, 2017). However, these indirect taxes do not affect the percentage change in the net-of-tax rate. Accordingly, we do not include them in the further analysis.} \]

\[ \text{11See e.g. Bastani and Lundberg (2017, Figure 5a), who report that the Pareto coefficient is around 3. However, the Pareto coefficient of the distribution of factor incomes is substantially smaller and around 2 (the right tail is thicker) (Pirttilä and Selin, 2011).} \]
The EITC phase-out reform came into effect on January 1, 2016. Did other reforms that are relevant to our study occur at the same time? The EITC phase-out was part of a government bill, in which several taxes were adjusted upwards. This was the first budget proposal from the new social democratic-green party coalition that gained parliamentary support. The new budget also contained stricter rules for the household tax reduction, higher payroll taxes for people aged over 65 (who are excluded from our analysis), higher taxes on tax-preferred savings accounts, and higher energy taxes. Finally, the kink points of the central government tax schedule were not fully adjusted for wage growth in 2016 and 2017.

3 Data

We utilize full-population individual register data from Statistics Sweden. The most important register we use (“LOUISE”) contains data from the labor market, educational and social sector and is updated each year. The key variable of interest is earned income, i.e. the base for the EITC. We construct this variable by taking the sum of wage and self-employment income. We also have detailed information on other demographic characteristics. These include age, gender, level and field of education, country of origin, county of residence, marital status, number of children and industry codes. The only sample restriction we make is that we limit the sample to individuals aged 20-65. No other restrictions are imposed, which implies that our empirical analysis is easily reproducible.

We also want to study income shifting between the labor and capital income tax bases. Therefore, we use separate data from Statistics Sweden, which include full population tax registers linked with tax return data from owners of closely held corporations (“FRIDA”). These data are analyzed in Section 4.5 below.

12 A similar budget proposal from the same government coalition was sensationaly rejected by the parliament in December 2014. Center-right wing budgets shaped tax policy in Sweden 2007-2015.

13 We are grateful to SNS for support with these data. Unfortunately, we are not able to link FRIDA to our main analysis data.
Figure 1: EITC in SEK as function of earned income.

(a) Marginal tax rate, with and without EITC, as function of earned income. Social transfers are assumed to be zero.

Figure 2: Taxes as function of earned income in 2016
4 Empirical analysis

4.1 Empirical model

The basic idea behind the empirical model is to compare earnings growth at treated and untreated parts of the earnings distribution. For each year \( t \) we rank all individuals aged 20 to 65 by their earnings, \( z \), and we partition the population into 100 equally sized percentile groups, which we follow over time. One may think of a percentile group \( j \) as a synthetic unit who a given year face marginal tax rate \( \tau_{jt} \). In the main analysis we focus on the upper part of the earnings distribution, where the central government tax applies, during the time period 2012-2017. We start the analysis in 2012, when the Swedish economy had recovered from the 2008 financial crisis. The upper percentile groups were essentially unexposed to marginal tax changes 2012-2015. In 2016 synthetic units belonging to percentile groups 96-100 faced a marginal tax increase due to the EITC phase-out. We will refer to percentile groups 96-100 as the treatment group, while percentile groups 88-95 constitute the control group. We want the individuals in the control group to be well above the first central government kink point, and therefore percentile group 88 is the lower limit of the control group. In some specifications, we exclude four percentile groups in a symmetric window around the treatment cut-off, i.e. percentile groups 94-97. Given earnings dynamics we hypothesize that the marginal tax increase is more salient to people who earn incomes at some distance from the kink, see Section 5 for a more systematic discussion.

We first estimate reduced form regressions of the following type on percentile groups 88-100:

\[
\log z_{ijt} = \sum_{t \geq 2016} \gamma_{jt}^{post} D_{jt} + \sum_{t < 2015} \gamma_{jt}^{pre} D_{jt} + \kappa_t + \mu_j + \delta X_{ijt} + \epsilon_{ijt}
\]

(1)

, where \( D_{jt} \) is an indicator that takes the value of 1 if percentile group \( j \) falls in the interval 96-100 and the year is \( t \), and it is 0 otherwise. \( \kappa_t \) is a

\footnote{We include people with zero earnings to reduce the influence of unemployment and non-participation.}

\footnote{The extreme high-income earners in the top 0.2 percent group did not experience increasing marginal tax rates and were unaffected by the policy (their entire tax credit was phased-out on infra-marginal earnings). For simplicity, we include these taxpayers in the treatment group in the main analysis.}
shorthand for the vector of time dummies, and $\mu_j$ represents a fixed effect at the percentile group level. In some specifications, we will also control for a vector of individual characteristics, which we denote by $X$. These include age, gender, education level, field of education, immigrant status, marital status, number of children, county, and 3-digit industry-codes. The year immediately preceding the reform, 2015, is the reference year. We cluster the standard errors at the percentile group level.

The analysis requires two central identifying assumptions. The parallel trends assumption implies that the treatment and control groups would evolve in the same way in the absence of the 2016 reform. If the pre-reform trends are parallel in the treatment and control groups we expect $\gamma_t^{\text{pre}}$ to be zero for 2012-14. By contrast, $\gamma_t^{\text{post}}$ for 2016 and 2017 should be negative if the EITC phase-out has a negative impact on earnings in the post-reform period. The constant group composition assumption fails if the reform brings about non-random compositional changes to the treatment and control groups. A concern could e.g. be that responsive people in the treatment group respond to the reform by transitioning into the control group (or migrating from the analysis sample). We will discuss the validity of this assumption below in Section 4.2.

If the earnings distribution is transformed by the EITC reform, it is reasonable to think that the effect materializes gradually rather than immediately. Therefore, the treatment estimates for 2016 and 2017 are likely to imply different earnings elasticities. A simple Wald estimator for the elasticity, $e_t$, is

$$
e_t = \frac{\Delta_t E(\log z|X, D = 1) - \Delta_t E(\log z|X, D = 0)}{\Delta_t E(\log (1 - \tau)|X, D = 1) - \Delta_t E(\log (1 - \tau)|X, D = 0)}$$

(2)

, with $t = \{2016,2017\}$, and the base year is always 2015. $D$ is an indicator that is 1 if the percentile group is 96-100, and zero otherwise. Hence, $\Delta_t$ denotes the change in mean quantities in the treatment and control groups between $t$ and 2015.

\[^{16}\text{We report analytical standard errors, but we have also computed standard errors using wild bootstrap. The results are similar.}\]
4.2 Group composition over time

Before turning to the central graphical analysis we briefly comment on the distribution of observable characteristics. Table A1 in the Appendix shows descriptive statistics, 2012-2017, for the treatment group and the control group. The control group is larger than the treatment group, since it contains more percentile groups. On average, the treatment group contains almost 300,000 individuals per year, and the control group more than 450,000 individuals. Individuals in the treatment group are slightly older (48 years of age compared to 46), fewer are females, and they have higher education. This was to be expected since earnings is higher in the treatment group. It is maybe a little bit more surprising that the immigrant shares in the two groups are quite similar.

Remember that a disadvantage of the cross sectional approach – as opposed to the longitudinal approach – is that the composition of treatment- and control groups may change endogenously due to the reform. From this perspective it is interesting to look at how observable characteristics develop over time in the two groups. If there is a discontinuous change in the distribution of covariates in the reform year, compositional changes is likely to be a major issue. Therefore, we graphed the evolution of four central covariates in the treatment and control groups over time in Figure 3. Individuals have higher education over time, the immigrant share increases, and the average age increases in both groups. This is in line with the evolution of the composition of the labor force in general. The share of females among high-income earners (both in the treatment and control group) increases over time, which is consistent with a shrinking gender wage-gap in Sweden in recent years.\footnote{See https://www.mi.se/app/uploads/LS_18.pdf, Diagram 1.2}

Note that the trends in the two groups are not always parallel. E.g., the share with university degree grows substantially faster in the control group. But it is of central importance that there are no discontinuous changes in 2016.\footnote{To test for this, we have run regressions with the covariate in question on the left hand side. The right hand side featured time dummies, a treatment group dummy, a treatment group dummy interacted with a time dummy and a linear treatment group specific trend. The interaction between the post-reform period and the treatment group dummy was always insignificant.}

Another possibility is to exploit the panel element of the data, and to study transitions between the groups over time. The probability that an individual, who in year $t$ belonged to the treatment group, is part of the

\[^{17}\text{See } https://www.mi.se/app/uploads/LS_18.pdf, \text{ Diagram 1.2}\]

\[^{18}\text{To test for this, we have run regressions with the covariate in question on the left hand side. The right hand side featured time dummies, a treatment group dummy, a treatment group dummy interacted with a time dummy and a linear treatment group specific trend. The interaction between the post-reform period and the treatment group dummy was always insignificant.}\]
treatment group also in $t + 1$ is surprisingly stable over time. The fraction of stayers is around 83 % in the whole time period. Taken together, these specification tests did not detect significant compositional changes of the treatment- and control groups.

### 4.3 Graphical evidence and main results

In Figure 4a we graph average log earnings in the treatment and control groups 2012-17. The graph entirely reflects raw data, where we have normalized the levels to be zero in 2015. There is a small tendency that log earnings grow faster in the control group between 2012 and 2013. However, the trends are extremely parallel 2013-15. In the reform year of 2016 earnings growth begins to divert, and in 2017 there is a salient gap between the two lines. Interestingly, earnings growth goes down also in the control group, and we will return to this phenomenon in Section 5 below. When we in Figure 4b exclude two percentile groups at each side of the cut-off, the pre-reform trends are still very parallel. However, the post-reform gap is
larger now. When evaluating elasticities using the Wald estimator in (2) we obtain an elasticity of 0.1 when we include all percentile groups 88-100, and we obtain an elasticity of 0.14 when excluding 2 percentile groups at each side of the cut-off. This aspect of the observed response will also be further discussed in the simulation section below.

Figure 4 provides a standard graphical difference-in-difference comparison. Still, clear graphical evidence of responses to income taxes is rare in the ETI literature. The closest example we know of is Kleven and Schultz (2014, Figure 4), who illustrate earnings responses of individuals who experienced decreasing and increasing marginal tax rates in the Danish 1987 reform. Our graphical analysis is in fact even more basic, because we simply partition the raw data into percentile groups and compare them, and our policy experiment is a clean introduction of a new bracket.

The regression results of Table 1 complement the visual analysis. In the absence of controls (columns 1 and 3), the coefficients for the interaction between the treatment group and the dummy for year \( t \), \( \gamma^\text{pre} \) and \( \gamma^\text{post} \), correspond to the vertical distance between the solid line and the dashed line in year \( t \) in Figure 4. Interestingly, pre-reform interactions are insignificant across specifications, while the opposite holds for post-reform interactions. This indicates a causal effect of the EITC reform. What about the magnitudes? When no percentile groups are excluded, the 2017 effect amounts to \(-0.64\) log points without controls. Since the log net-of-tax decreased by around 7 log points, the implied elasticity is around 0.1. When excluding percentile groups 94-97, we estimate a larger effect: \(-0.90\) without controls in column 3. The 2017 effect is always estimated to be larger than the 2016 effect.

The results of columns 2 and 4, where we control for a rich set of potentially confounding factors, deserve special attention. If the interactions of interest would be correlated with the error term in (1), we expect control variables to have a large impact on the estimated reform effects. Since we are working on a large administrative data set, we are able to include a rich set of controls in a very flexible way. We use dummies for age, gender, education level, education field, immigrant status, marital status, number

---

19 The Danish 1987 was more complex than the reform we consider here, and it included both tax rate and tax base changes. Kleven and Schultz (2014) restrict their sample to a balanced panel of individuals who are observed throughout the period.
Table 1: DiD-regressions

<table>
<thead>
<tr>
<th></th>
<th>(1) All percentiles</th>
<th>(2) Percentiles 94-97 excluded</th>
<th>(3) No controls</th>
<th>(4) Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Treated * 2012</strong></td>
<td>0.264 (0.268)</td>
<td>0.256 (0.243)</td>
<td>0.244 (0.339)</td>
<td>0.261 (0.304)</td>
</tr>
<tr>
<td><strong>Treated * 2013</strong></td>
<td>0.050 (0.293)</td>
<td>0.043 (0.267)</td>
<td>-0.004 (0.370)</td>
<td>0.017 (0.330)</td>
</tr>
<tr>
<td><strong>Treated * 2014</strong></td>
<td>-0.097 (0.209)</td>
<td>-0.094 (0.187)</td>
<td>-0.173 (0.260)</td>
<td>-0.154 (0.231)</td>
</tr>
<tr>
<td><strong>Treated * 2016</strong></td>
<td>-0.314*** (0.092)</td>
<td>-0.306*** (0.104)</td>
<td>-0.473*** (0.100)</td>
<td>-0.497*** (0.139)</td>
</tr>
<tr>
<td><strong>Treated * 2017</strong></td>
<td>-0.639*** (0.128)</td>
<td>-0.633*** (0.139)</td>
<td>-0.901*** (0.088)</td>
<td>-0.943*** (0.118)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>4,500,619</td>
<td>4,500,619</td>
<td>3,115,812</td>
<td>3,115,812</td>
</tr>
</tbody>
</table>

Notes: All regressions (columns 1–4) include controls for year and percentile group. Regressions in columns (2) and (4) include controls for age, gender, education level, education field, immigrant status, marital status, number of children, county, and 3-digit industry-codes. All control variables are also interacted with year dummies. Standard errors are clustered at percentile groups. Significance levels: *** p < 0.01, ** p < 0.05, * p < 0.1

of children, county, and a 3-digit industry-codes. Moreover, we interact all these dummies with the full set of time dummies. Still, it turns out that the controls only have a negligible effect in the specification including all percentile groups (column 2). Similarly, it only has a slight impact on the results when excluding percentile groups 94-97. Given that endogenous compositional changes is one of the main concerns with the cross sectional approach, we find these results reassuring. They are also consistent with the analysis of Section 4.2, which demonstrated that there are no discontinuous changes in central observable characteristics in the reform year.

4.4 Alternative specifications

In this section we summarize what we get from some alternative approaches.
Figure 4: Average log earnings in treatment- and control groups. Average log earnings are normalized to be zero in both groups in 2015. Raw data.
Predicted income

The main analysis is simple and transparent, and it has a close connection to basic descriptive analysis. However, a potential criticism is that we assign individuals into treatment- and control groups based on the outcome variable (log earnings). As already discussed, this is controversial from an econometric perspective, because the composition of these groups may change if people respond to the reform in such a way the ordering of taxpayers is altered. Therefore, we also estimated a model, in which we use percentile groups based on predicted rather than actual income. In a first step, we predict income as functions of pre-determined characteristics, which are plausibly orthogonal to the reform. Actual earnings now differ less in different percentile groups, because someone who is in the top percentile group may report very low actual earnings, and vice versa. Unfortunately, when predicting income based on genuinely pre-determined characteristics we do not obtain a sufficiently strong first stage, see Appendix B.1 for more details. Intuitively, observable characteristics simply explains too little of the variation in earnings at high income levels.

Panel data

We have run Gruber and Saez (2002) like regressions, which means that we regress individual level changes in log earnings on individual level changes in log net-of-tax rates 2015-17. We construct a tax instrument as a function of base year earnings, and we control for log base year earnings. The results are not that informative: estimates varies a lot depending on specification. As expected, transitory incomes play an important role. More information is provided in Appendix B.2.

4.5 Is the response driven by income shifting?

An important feature of the Swedish income tax system is that labor incomes are taxed progressively, whereas capital incomes are taxed at a low proportional rate. In the Swedish dual income tax system high-income earners therefore face substantial incentives to shift income between the tax bases. Usually, regular wage earners cannot transform earnings into capital income, because their wage income is third-party reported. The situation is
different, however, for active owners of closely held corporations (CHCs), who are working in their own firms, and are able to distribute themselves dividends instead of wages. Alstadsæter and Jacob (2016) have documented that such activities are important in Sweden. Already before the 2016 reform there was a large gap of almost 30 percentage points between the top effective labor marginal tax rate and the effective tax rate on dividends from CHCs after accounting for payroll taxes and corporate taxes. When the EITC phase-out was introduced in 2016, the gap widened to 32 percentage points.

Against this background it is natural to ask whether the response we observe in the main analysis is driven by income shifting of active CHC owners. The most simple way to examine this is to exclude the potential group of "shifters", namely the active owners of the CHCs. In Table 2, we do this in two steps. In column 2 we first exclude CHC owners who receive dividends from their own corporation. They correspond to 7% of the baseline sample. However, all active CHC owners do not take out dividends a specific year. In column 3 we exclude all CHC owners, who make up 13% of our original study population in percentile groups 88-100. When re-estimating the model we use the same percentile limits as in the main analysis.

We infer from Table 2 that there are no dramatic changes to the results when excluding CHC owners. If anything, there is a tendency that the reform effects amplify in columns 2 and 3. Therefore, we do not think that the estimated response is driven by income shifting between the labor and

---

20 Swedish CHC owners cannot distribute lightly taxed dividends freely. Each year there is a cap on dividends that can be taxed at the low rate, commonly referred to as the dividend allowance. The dividend allowance has become more generous over time.

21 The effective marginal tax rate on labor income can be written $1 - \frac{1}{1 + t_p}$, where $t$ is the personal marginal tax rate and $t_p$ is the payroll tax rate. The effective dividend tax rate can be written $1 - (1 - \tau_d)(1 - \tau_{11})$, where $\tau_d$ is the dividend tax rate and $\tau_{11}$ is the corporate tax rate. (Here we do not account for consumption taxes, because they do not change the relation between the two tax rates.) In 2015 we had $\tau = 0.57$, $t_p = 0.3142$, $\tau_d = 0.2$, and $\tau_{11} = 0.22$. Hence, the effective labor tax rate was around 0.67, whereas the effective dividend tax rate amounted to around 0.38. In the 2016 phase-out reform $\tau$ increased from 0.57 to 0.6, and the effective labor tax rate rose to around 0.7.

22 As discussed in Section 3 in these estimations we used a separate data source that contains information on the corporate owners' income tax returns. To identify the groups of "shifters" we used information from the K10 form. Column 2 excludes everyone who reports dividends on a K10 form, and column 3 excludes all those filing a K10 form. The K10 form must be filled in by all active CHC owners who want to accumulate or use dividend allowances. Note also that there are slight differences in parameter estimates and in the number of observations in column 1 of Table 2 and column 1 of Table 1, reflecting that we use a different data source.
Table 2: DiD-regressions excluding "shifters"

<table>
<thead>
<tr>
<th></th>
<th>(1) All</th>
<th>(2) No CHC dividends</th>
<th>(3) No CHC owners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treated * 2012</td>
<td>0.267 (0.261)</td>
<td>0.315 (0.260)</td>
<td>0.422* (0.227)</td>
</tr>
<tr>
<td>Treated * 2013</td>
<td>0.068 (0.283)</td>
<td>0.200 (0.272)</td>
<td>0.334 (0.273)</td>
</tr>
<tr>
<td>Treated * 2014</td>
<td>-0.078 (0.203)</td>
<td>0.025 (0.205)</td>
<td>0.136 (0.186)</td>
</tr>
<tr>
<td>Treated * 2016</td>
<td>-0.281*** (0.075)</td>
<td>-0.333*** (0.107)</td>
<td>-0.310*** (0.098)</td>
</tr>
<tr>
<td>Treated * 2017</td>
<td>-0.598*** (0.118)</td>
<td>-0.701*** (0.143)</td>
<td>-0.704*** (0.146)</td>
</tr>
<tr>
<td>Observations</td>
<td>4,499,813</td>
<td>4,183,412</td>
<td>3,913,033</td>
</tr>
</tbody>
</table>

Notes: All regressions (columns 1–3) include controls for year and percentile group. Standard errors, clustered at percentile groups, in parenthesis. Significance levels: *** p < 0.01, ** p < 0.05, * p < 0.1

5 Simulation model

In this section we outline a model framework. The purpose of the simulation exercise is to interpret Figure 4. In particular, we want to interpret why the estimated response grows when we exclude percentile groups around the treatment cut-off, and we want to understand why earnings growth slows down also in the control group after the reform. Our underlying hypothesis is that people have noisy perceptions of tax rates.

5.1 Model

Consider a model economy in which agents differ with respect to potential incomes (skills) \( z_0 \). Individuals derive utility from consumption, \( c \), and disutility from earnings supply, \( z \). The budget constraint is \( c = z - T(z) \), where \( T(z) \) is a piece-wise linear tax function, perfectly perceived by the
individuals. Under certainty, an individual maximizes the utility function

\[ U = z - T(z) - \frac{z_0}{1 + \frac{1}{e}} \left( \frac{z}{z_0} \right)^{1 + \frac{1}{e}} \]  

with respect to \( z \). If the individual has an optimum at a linear segment of the budget constraint, the optimal earnings supply function is \( z = (1 - \tau)cz_0 \), where \( \tau \) is the marginal tax rate. Under these assumptions, the earnings supply only depends on the marginal tax rate, \( \tau \), and the potential income \( z_0 \). If \( \tau = 0 \) we have \( z = z_0 \). The (compensated) earnings elasticity is \( e \). Note that there are no income effects on earnings supply. This is quite an innocuous assumption in this context. Most treated taxpayers only experienced small changes in disposable incomes. We choose to refer to the “earnings elasticity” rather than the “taxable income elasticity”, since labor earnings is the base for the EITC (the tax rate variation we are using).

Following Saez (1999) we now extend the standard earnings supply model to a choice environment featuring uncertainty. Historically, this framework has been used to rationalize that taxpayers do not bunch at kinks, see e.g. Blomquist and Simula (2019). We instead use it to interpret broader transformations of the earnings distributions when taxes change. Suppose that individuals are able to control expected, but not realized, earnings. More specifically, suppose that realized earnings is given by

\[ \tilde{z} = z + \varepsilon \]  

, where \( z \) is chosen by the individual, and \( \varepsilon \) is a stochastic term, which the individual cannot control. An example of a positive shock could be an end-of-year bonus. A negative shock could be an unexpected sickness absence spell. The agent does, however, know the distribution of \( \varepsilon \). For simplicity, we assume that the stochastic element of realized earnings is normally distributed with mean zero, i.e. \( \varepsilon \sim N(0,\sigma^2) \). The agent maximizes expected utility

\[ EU = \int \{ z + \varepsilon - T(z + \varepsilon) - \frac{z_0}{1 + \frac{1}{e}} \left( \frac{z}{z_0} \right)^{1 + \frac{1}{e}} \} f(\varepsilon) d\varepsilon \]  

, where \( T(\tilde{z}) \) is the piecewise linear tax function and \( f(\varepsilon) \) is the pdf of the normal distribution. Note that the disutility of earnings supply is known
with certainty, while consumption differs depending on the realization of $\varepsilon$. When utility is quasi-linear in consumption, i.e. under risk neutrality, (5) can be rewritten as

$$EU = z - \hat{T}(z) - \frac{z_0}{1 + \frac{1}{\varepsilon}} \left( \frac{z}{z_0} \right)^{1 + \frac{1}{\varepsilon}}.$$  \hspace{1cm} (6)

\(\hat{T}(z) = \int T(z + \varepsilon)f(\varepsilon)d\varepsilon\) can be thought of as the effective tax schedule facing the individual. In a way, the choice problem under uncertainty has been transformed into a choice problem under certainty. Maximizing (6) with respect to $z$ is akin to the certainty problem given by (3). The only difference is that the actual piece-wise linear tax function $T(z)$ is replaced by a smooth tax function $\hat{T}(z)$.

Intuitively, when people do not know their incomes by the end of the year with certainty, the marginal incentives to work will be given by a weighted average of actual marginal tax rates around their expected earnings levels. As most other tax systems, the actual Swedish income tax system features sharp kink points, where marginal tax rates change. These kinks are absent in the effective tax schedule, $\hat{T}(z)$, where marginal tax rates instead changes gradually around the actual kink. When a new tax bracket is introduced, like the Swedish EITC phase-out in 2016, effective tax rates of those with realized incomes below the new statutory kink will also be affected. The standard deviation of $\varepsilon$, $\sigma$, is a key parameter determining the shape of the effective marginal tax schedule. If $\sigma$ is small, effective and actual marginal tax rates will differ only locally around kinks. In the special case in which $\sigma = 0$ the two schedules will be identical. If $\sigma$ is large, the introduction of a new kink to the actual tax system will impact effective marginal tax rates over wide ranges of income that would be unaffected in the certainty model. In Figure 5 we illustrate the EITC reform in the effective marginal tax schedule for 2016, with and without an EITC phase-out for different values of $\sigma$. When $\sigma = 70,000$ many taxpayers both to the left and to the right of the EITC kink are affected by the reform.

\footnote{In Appendix C we describe how we smooth the actual tax schedule.}
Figure 5: Reform in effective marginal tax schedules for different values of $\sigma$. 
5.2 Calibration of the simulation model

What is a reasonable value of \( \sigma \) from the perspective of earnings dynamics? Suppose that taxes are constant over time and that individuals have skill level, \( z_0 \), both in year \( t - 1 \) and \( t \). Since \( \tilde{z}_t \sim N(z_t, \sigma^2) \) we have that the difference is also normally distributed with twice the variance, i.e. \( \tilde{z}_t - \tilde{z}_{t-1} \sim N(0,2\sigma^2) \). Accordingly, if \( \sigma^{\text{diff}} \) is the standard deviation of the distribution of individual realized earnings differences, it holds true that \( \sigma = \frac{\sigma^{\text{diff}}}{\sqrt{2}} \). Following Saez (1999) we use this relationship to quantify \( \sigma \) on panel data data from 2012-15, i.e. the pre-reform years with no major tax changes and a stable earnings distribution. In the 95th percentile group the standard deviation amounts to \( \sigma = 70,000 \), and we think that this is an upper bound of a relevant value \( \sigma \). We elaborate more on this in Appendix C.1. We pick \( \sigma \) for the 95th percentile group, because the kink that we want to smooth is located there.

Another important input to the simulation model is the skill distribution. In the spirit of Saez (2001), we recover the distribution of \( z_0 \) from the empirically observed earnings distribution, see Appendix C for details. Consequently, our simulated pre-reform realized earnings distribution will have properties that are similar to the empirical earnings distribution, and we will examine how earnings react to the 2016 reform on the simulated data. In our baseline simulations we used \( \epsilon = 0.2 \), which is an elasticity that matches the observed responses quite well. Potential income growth is assumed to be constant over time and the same for all individuals. Hence, if nothing happened to the tax system in 2016, log earnings growth would be the same in all percentile groups the entire period 2012-17. To preserve the linear and parallel trends in log earnings, we will plot log \( z \) rather than log \( (z + \epsilon) \). Note also that we do not model adjustment costs. Accordingly, the tax reform response will materialize immediately in 2016 in the simulations, and it will not grow 2016-17.

\[24\] When quantifying \( \sigma \) it is important to distinguish between \( \epsilon \) and the error term in the empirical equation, which e.g. is denoted by \( \epsilon \) in the reduced form equation above. While the latter represents factors unobserved by the econometrician, the former refers to factors that are random from the individual’s perspective. Of course, if \( \sigma \) is large, the variance of the transitory empirical error term will probably be large as well. However, the individual may non-randomly choose different earnings levels from year to year for reasons that are known by the individual but unobserved by the econometrician.
5.3 Simulated responses

We first study simulated responses when the noise in tax perceptions is motivated by earnings dynamics, i.e. $\sigma =$SEK 70,000. Figure 6 is the simulated counterpart to Figure 4a of Section 4.3. In Figure 6, the response under uncertainty is contrasted to the response under certainty, when people respond to changes in the piece-wise linear tax schedule with extreme precision. In the certainty model, people in the treatment group reduce their earnings according to the assumed elasticity, which is 0.2. By contrast, there is no reaction in the control group. Earnings growth continues at the pre-reform pace also in 2016 and 2017 in the control group.

When turning to the output from the uncertainty model in Figure 6, we see that the smoothening of the tax schedule has led to a visually detectable deviation from the certainty model. The estimated elasticity is now 0.15, i.e. 5 percentage points lower than the underlying behavioral elasticity of 0.2. In the uncertainty model, individuals in the control group, especially those located near the new bracket cut-off, experience increasing effective marginal tax rates. Therefore, we simulate a lower earnings growth also in the control group 2015-16. On the same note, individuals in the treatment group who are close to the cut-off will receive a lower tax increase compared to the certainty model, and they will react less 2015-16. Still, when $\sigma =$SEK 70,000 we do not obtain the empirically estimated elasticity, which is 0.1 when including all percentile groups and 0.14 when excluding 4 percentile groups symmetrically around the treatment cut-off. Actually, when excluding 4 percentile groups from the simulated data, we estimate an elasticity of 0.19, which is close to the underlying elasticity of 0.2, but relatively far from the empirically estimated elasticity of 0.14.

Maybe it should not come as a surprise that earnings dynamics alone cannot rationalize the fuzziness of the response. When formulating the model in Section 5.1 we assumed that individuals have full knowledge about the piece-wise linear tax function – the smoothness of effective tax rates entirely comes from earnings shocks. However, taxpayers’ knowledge of the tax code has been questioned in recent years. Liebman and Zeckhauser [2004] and Rees-Jones and Taubinsky [2019] claim that people rely on simple mental heuristics rather than perfect information when responding to complex non-linear tax schedules. In particular, people tend to confuse marginal and
average tax rates. These studies, however, abstract from earnings uncertainty as a source of misperception. Therefore, we also performed simulations where we impose more noise than what is motivated by earnings dynamics. In Figure 6, we show data generated by $\sigma =$SEK 140,000. In all other respects, we use exactly the same parameter values as in Figure 6. Interestingly, when doubling the standard deviation of the noise term, while still assuming an underlying elasticity of 0.2, we obtain simulated elasticities that are very close to the empirically estimated elasticities. In Figure 7, we obtain an elasticity of 0.11, which is very close to the corresponding empirical elasticity estimate of 0.10, reported in Figure 4a. And when four percentile groups are omitted from the simulated data we estimate 0.14, which is identical to the empirical estimate reported in Figure 4b.

The empirical estimates are summarized in Table 3. We wish to emphasize that the purpose of the simulation exercise not is to provide causal statements, but to provide a systematic interpretation of earnings growth in the treatment- and control groups. It is difficult to tell whether $\sigma =$SEK 140,000 is a surprisingly small or large value of $\sigma$. On the one hand, it is sufficiently small for a response to appear when the elasticity is 0.2. On the other
hand, $\sigma = \text{SEK} 140,000$ suggests that the effective tax schedule is very different from the piece-wise linear schedule, a finding that is consistent with the no-bunching result from wage earners in Bastani and Selin (2014).

6 Concluding section

We evaluate earnings responses to a 7% reduction in the net-of-tax share, which in 2016 affected a large number of Swedish high-income earners already facing high taxes. We exploit full-population administrative data, and we graphically compare earnings growth at different parts of the income distribution in a simple and transparent way. With two years of post-reform data, we estimate an earnings elasticity of 0.1. We discuss potential disadvantages of the cross-sectional approach, and we find that the response is not driven by active owners of closely held corporations.

When people facing non-linear tax schedules cannot perfectly control their incomes, perceptions of marginal incentives will be noisy as well, especially for high-skilled workers. We therefore interpret essential features of the response using a simulation model, in which people have noisy perceptions
Table 3: Simulated and empirical elasticities

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All percentile groups</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>+2 Percentile groups excluded</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>σ = SEK 0</td>
<td>0.15</td>
<td>0.19</td>
</tr>
<tr>
<td>σ = SEK 70,000</td>
<td>0.11</td>
<td>0.14</td>
</tr>
<tr>
<td>σ = SEK 140,000</td>
<td>0.10</td>
<td>0.14</td>
</tr>
<tr>
<td>Empirical point estimates</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The details of the simulations are described in the main text and Appendix C. The empirical estimates have been obtained from the Wald estimator in (2), evaluated in 2017 and are also reported in Figure 4.

of the piece-wise linear tax code. To simulate the empirically observed response, we need to add more noise to perceptions than what is motivated by earnings uncertainty alone.

If our results reflect long-run responses to taxes, the earnings elasticity is fairly low in Sweden, and the revenue maximizing tax rate is high. It is important to keep in mind, however, that we only considered two post-reform years, and we cannot exclude that the response is larger (or smaller) in the third and fourth post-reform year.

References


Appendix

A  Summary statistics

Summary statistics are reported in Table A1.

B  IV estimates using predicted earnings and panel data models

B.1  Predicted earnings

To avoid potential problems with endogeneity, we have estimated models where we group individuals into percentile groups based on predicted earnings rather than actual earnings. We regress earnings on a set of predetermined characteristics that are arguably exogenous to reform, and we use the predicted values from these regressions to group individuals to percentiles. The treatment group is defined as individuals belonging to percentile groups 96 and above, as when we grouped on actual earnings.

After classifying individuals to percentile groups based on predicted earnings, we estimate the following model on data from 2015 and 2017:

\[
\log(z)_{ijt} = \alpha + \beta \times \log(ntr)_{ijt} + \mu_t + \mu_j + \eta_{ijt}.
\]

Log earnings (log(z)) for individual \(i\) in percentile \(j\) in year \(t\) is regressed on log net-of-tax rates (log(ntr)), time fixed effects (\(\mu_t\)), and percentile fixed effects (\(\mu_j\)). Log(ntr) is instrumented by belonging to the treatment group in 2017 (i.e. the interaction of treatment status and a dummy for 2017). The model is estimated by 2sls, the control group are individuals in percentile groups 88-95, and the standard errors are clustered at the percentile group level. As a sensitivity check, we include controls for age, gender, education (level and field), marital status, immigrant status, municipality, industry, and occupation.

It is difficult to find pre-determined characteristics that are exogenous to reform, and are able to predict individuals correctly to treatment and control groups, i.e. are able to provide a significant first stage. Since the reform
### Table A1: Descriptives, 2012–2017

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2012</td>
<td>2013</td>
<td>2014</td>
<td>2015</td>
<td>2016</td>
<td>2017</td>
</tr>
</tbody>
</table>

#### Panel A: Treated

<table>
<thead>
<tr>
<th>Variable</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>47.86</td>
<td>48.00</td>
<td>48.11</td>
<td>48.16</td>
<td>48.24</td>
<td>48.35</td>
</tr>
<tr>
<td>Female</td>
<td>0.250</td>
<td>0.259</td>
<td>0.267</td>
<td>0.275</td>
<td>0.283</td>
<td>0.289</td>
</tr>
<tr>
<td></td>
<td>(0.433)</td>
<td>(0.438)</td>
<td>(0.442)</td>
<td>(0.446)</td>
<td>(0.451)</td>
<td>(0.453)</td>
</tr>
<tr>
<td>Immigrant</td>
<td>0.080</td>
<td>0.085</td>
<td>0.089</td>
<td>0.096</td>
<td>0.103</td>
<td>0.107</td>
</tr>
<tr>
<td></td>
<td>(0.272)</td>
<td>(0.279)</td>
<td>(0.285)</td>
<td>(0.294)</td>
<td>(0.304)</td>
<td>(0.309)</td>
</tr>
<tr>
<td>Secondary educ</td>
<td>0.211</td>
<td>0.209</td>
<td>0.208</td>
<td>0.207</td>
<td>0.206</td>
<td>0.203</td>
</tr>
<tr>
<td></td>
<td>(0.408)</td>
<td>(0.407)</td>
<td>(0.406)</td>
<td>(0.405)</td>
<td>(0.404)</td>
<td>(0.402)</td>
</tr>
<tr>
<td>University</td>
<td>0.757</td>
<td>0.759</td>
<td>0.762</td>
<td>0.762</td>
<td>0.763</td>
<td>0.767</td>
</tr>
<tr>
<td></td>
<td>(0.429)</td>
<td>(0.428)</td>
<td>(0.426)</td>
<td>(0.426)</td>
<td>(0.425)</td>
<td>(0.423)</td>
</tr>
<tr>
<td>Married</td>
<td>0.651</td>
<td>0.650</td>
<td>0.649</td>
<td>0.648</td>
<td>0.646</td>
<td>0.645</td>
</tr>
<tr>
<td></td>
<td>(0.477)</td>
<td>(0.477)</td>
<td>(0.478)</td>
<td>(0.478)</td>
<td>(0.478)</td>
<td>(0.478)</td>
</tr>
<tr>
<td>Observations</td>
<td>283,717</td>
<td>285,234</td>
<td>287,147</td>
<td>288,889</td>
<td>291,815</td>
<td>294,203</td>
</tr>
</tbody>
</table>

#### Panel B: Control

<table>
<thead>
<tr>
<th>Variable</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>45.82</td>
<td>46.00</td>
<td>46.16</td>
<td>46.30</td>
<td>46.37</td>
<td>46.44</td>
</tr>
<tr>
<td>Female</td>
<td>0.305</td>
<td>0.314</td>
<td>0.321</td>
<td>0.329</td>
<td>0.335</td>
<td>0.343</td>
</tr>
<tr>
<td></td>
<td>(0.461)</td>
<td>(0.464)</td>
<td>(0.467)</td>
<td>(0.470)</td>
<td>(0.472)</td>
<td>(0.475)</td>
</tr>
<tr>
<td>Immigrant</td>
<td>0.076</td>
<td>0.080</td>
<td>0.086</td>
<td>0.092</td>
<td>0.100</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td>(0.265)</td>
<td>(0.272)</td>
<td>(0.280)</td>
<td>(0.289)</td>
<td>(0.300)</td>
<td>(0.307)</td>
</tr>
<tr>
<td>Secondary educ</td>
<td>0.360</td>
<td>0.352</td>
<td>0.347</td>
<td>0.342</td>
<td>0.336</td>
<td>0.327</td>
</tr>
<tr>
<td></td>
<td>(0.480)</td>
<td>(0.478)</td>
<td>(0.476)</td>
<td>(0.474)</td>
<td>(0.472)</td>
<td>(0.469)</td>
</tr>
<tr>
<td>University</td>
<td>0.579</td>
<td>0.590</td>
<td>0.597</td>
<td>0.604</td>
<td>0.612</td>
<td>0.624</td>
</tr>
<tr>
<td></td>
<td>(0.494)</td>
<td>(0.492)</td>
<td>(0.491)</td>
<td>(0.489)</td>
<td>(0.487)</td>
<td>(0.484)</td>
</tr>
<tr>
<td>Married</td>
<td>0.538</td>
<td>0.538</td>
<td>0.539</td>
<td>0.539</td>
<td>0.540</td>
<td>0.539</td>
</tr>
<tr>
<td></td>
<td>(0.499)</td>
<td>(0.499)</td>
<td>(0.498)</td>
<td>(0.498)</td>
<td>(0.498)</td>
<td>(0.498)</td>
</tr>
<tr>
<td>Observations</td>
<td>453,948</td>
<td>456,376</td>
<td>459,436</td>
<td>462,223</td>
<td>466,906</td>
<td>470,725</td>
</tr>
</tbody>
</table>

Notes: Means, with standard deviations in parenthesis.
affects individuals with very high earnings, we need to find observable pre-determined characteristics that can differentiate between individuals at the very top of the earnings distribution. Typically, when good earnings predictors are observable, they are related to labor market characteristics (e.g. occupation), which in turn could be argued to be endogenous to reform.

In Table [B.1] we show results for two types of predictions. In the first case (columns 3 and 4), we aim to use strict pre-determined characteristics. First, we use information on age, gender, immigrant status, and education. Second, we use information that potentially could be endogenous (occupation, industry, municipality of residence and marital status), but we lag those characteristics two years in order to circumvent endogeneity issues. In the second case of predictions (column 5 and 6), we use all the above mentioned information contemporary. The idea is to get best possible predictions. In these specifications, we should get a stronger first stage, but potentially have a bigger issue with endogeneity. As a comparison, we replicate the results where we have grouped individuals based on actual earnings (columns 1 and 2).

In the case with strictly pre-determined characteristics, we get a large, marginally significant estimate, when we do not include control variables (Table [B.1] column 3). However, when we do include control variables, we no longer have a first stage (Table [B.1] column 4). Our conclusion is that we are not able to establish a significant first stage (correct predictions of treatment status) with strictly pre-determined characteristics.

If we allow ourselves to view all characteristics as being pre-determined, our first stage gets somewhat stronger (column 6). However, the estimates of the log net-of-tax rate are imprecisely estimated. We conclude that even with this stronger assumption, we are not able to predict earnings with adequate precision. Maybe this should not be that surprising, it is genuinely hard to predict earnings in the very top of the earnings distribution where individuals’ earnings vary substantially between years.

### B.2 Panel data models

We have estimated panel data models, using data from 2015 and 2017, to study the effects of changes in log net-of-tax-rates on changes in log earnings. More specific, we have estimated the following model using 2sls:
Table A2: IV-estimates using predicted earnings

<table>
<thead>
<tr>
<th>Outcome variable: ln(earnings)</th>
<th>Actual earnings</th>
<th>Predicted earnings (Lagged charac.)</th>
<th>Predicted earnings (Contemporary charac.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No controls (1) With controls (2)</td>
<td>No controls (3) With controls (4)</td>
<td>No controls (5) With controls (6)</td>
</tr>
<tr>
<td>Ln(ntr)</td>
<td>0.099***        0.103***              0.475*        4.13*                      0.273         0.178</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.024)         (0.028)                (0.263)       (2.31)                    (0.734)       (0.187)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>No             Yes                    No             Yes                       No             Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-values</td>
<td>37             37                     26             3                        14             24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clusters</td>
<td>13             13                     13             13                       13             13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,527,703      1,527,703               1,527,703     1,527,703                      1,527,703     1,527,703</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The outcome variable is ln(earnings). All specifications include controls for percentiles, treatment status, and year effects. The other control variables included in columns 2, 4, and 6 are: age, gender, education (level and field), marital status, immigrant status, municipality, industry (3-digit), and occupation (3-digit). Standard errors are clustered at percentile level and shown in parentheses. Asterisks indicate that the estimates are significantly different from zero at the * p < 0.1, ** p < 0.05, *** p < 0.01 level.

\[
\Delta \log(z)_{ijt} = \alpha + \beta \times \Delta \log(ntr)_{ijt} + \eta_{ijt}
\] (B2)

where the we instrument the changes in log net-of-tax rates. We create tax instruments by using previous earnings, i.e. we calculate net-of-tax rates for 2017 using previous earnings (Gruber and Saez, 2002). To correct for mean reversion, we include measures of previous earnings in the model. We estimate the model on individuals aged 20–50 in 2015, earning more than 500 000 SEK, and cluster standard errors at the individual level.

We have tried three different measures of previous earnings, creating three different tax instruments: earnings from 2015 (base-year earnings, as standard in the literature), earnings from 2016, as well as average earnings 2015–2017. For all three instruments, we present results with and without including controls for the previous earnings measure (Table B.2).

The overall picture from Table B.2 is that the estimates varies considerably, both between instruments, and with/without controlling for previous earnings. Our conclusion is that this type of model appear not to be suited for analyzing this reform.

\[^{25}\text{We have controlled for previous earnings in different ways, and the results vary substantially between specification.}\]
Table A3: IV estimates (Gruber/Saez)

<table>
<thead>
<tr>
<th>Instrument:</th>
<th></th>
<th>Instrument:</th>
<th></th>
<th>Instrument:</th>
</tr>
</thead>
<tbody>
<tr>
<td>D \ln(ntr)</td>
<td>0.492***</td>
<td>-0.893***</td>
<td>-2.133***</td>
<td>0.309***</td>
</tr>
<tr>
<td>(0.099)</td>
<td>(0.122)</td>
<td>(0.041)</td>
<td>(0.079)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>Controlling for \ln(e)</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>365,648</td>
<td>365,648</td>
<td>365,648</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The sample consists of individuals aged 20–50 in 2015, earning more than 500,000 SEK. Differences corresponding to changes between 2015 and 2017. Standard errors clustered at individual level. Asterisks indicate that the estimates are significantly different from zero at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ level.

C Details of the simulation model

To simulate tax policy effects on earnings we essentially need, (i), a model for individual behavior, (ii), data on the relevant tax function, and, (iii), data on the skill distribution. In Section 5.1 we elaborated on (i). In this Appendix we describe in greater detail how we deal with (ii) and (iii), and how we solve the models. First, we show how to calibrate a value of $s$ using earnings dynamics. Second, we outline how we go about to smoothen the income tax schedule given assumptions about $s$. Third, we show how to calibrate a potential income (skill) distribution with desirable properties. Fourth, we solve the model for the piece-wise linear case, and finally, fifth, we solve the model for the uncertainty (smooth) case.

C.1 Calibrating $s$ using earnings dynamics

A key parameter when smoothening the tax schedule is the standard deviation of the earnings noise term, $s$. Ideally, the noise term should capture variation in earnings that the individual cannot control when choosing optimal expected earnings. Following Saez (1999), we back out a value of $s$ by using empirical year-to-year differences in earnings. We assume that empirical earnings realizations are generated by $\tilde{z} = z + \varepsilon$, where $z$ is determined by the optimization problem in (3). Crucially, we only use pre-reform data 2012-15, which was a period without any substantial changes to tax rates. Hence, we do not expect $z$ to change because of tax changes. When
\( \varepsilon \sim N(0, \sigma^2) \), it follows that \( \tilde{z}_t - \tilde{z}_{t-1} \sim N(0, \sigma_{\text{diff}}^2) \) where \( \sigma_{\text{diff}}^2 \equiv 2\sigma^2 \). We pool \( \tilde{z}_t - \tilde{z}_{t-1} \) for \( t = 2013, 2014, 2015 \), and we study the distribution of first differences in realized earnings. Since the variability in earnings differs a lot across the earnings distribution, we compute different \( \sigma \) for different percentile groups.

Our ambition is that \( \sigma \) should capture earnings dynamics at the new EITC kink, which we wish to smoothen. Figure A1 illustrates the distribution of \( \tilde{z}_{t-1} - \tilde{z}_t \) in the 95th percentile group, which is located just at the earnings level of the kink. When imposing the normality assumption we obtain \( \sigma \approx \text{SEK 70,000} \) in this group. Having said that, the validity of the normality assumption can be questioned. The histogram displays the empirical distribution of earnings differences. Note that the mode value is larger than zero due to real wage growth. Obviously, when fitting the normal density to the raw distribution (solid line) it turns out that the fit is poor. In particular, the empirical distribution contains a substantially larger mass at small earnings differences. Moreover, the distribution is skewed to the left – it is more common with large earnings reductions than with large increases in earnings.

It is important to keep in mind that year-to-year differences in earnings actually may reflect active choices that the individual is able to control, but the econometrician cannot observe. When considering all these aspects, we believe that our approach overestimates the earnings shocks facing the individual.

As already mentioned, year-to-year variation in earnings differs depending on the earnings level. Figure A2 illustrates that the calibrated value of \( \sigma \) tends to increase in earnings. In the 95th percentile group we have \( \sigma \approx \text{SEK70,000} \), which will be the baseline choice of \( \sigma \) in the simulation exercise.

### C.2 Piece-wise linear marginal tax rates

In Section 2 we outlined the main features of the Swedish tax system. A salient feature of the Swedish system is that the marginal tax rate jumps dramatically at the income level where the central government tax kicks in ("the first central government kink point"), see Bastani and Selin (2014). In this simulation exercise we will consider individuals with earnings exceeding the 87th percentile 2012-17, i.e. well above the first central government
Figure A1: Distribution of $\tilde{z}_t - \tilde{z}_{t-1}$ in the 95th percentile group

Figure A2: $\sigma$ by percentile group
kink, which is located at the 84th percentile. These individuals report earnings well above the first central government kink point. As can be seen from Figure 2b the pre-reform system contains only one kink above the first central government kink point. This is the “second central government kink”, where the central government tax increases by 5 percentage points. The EITC phase-out reform of 2016 introduced a new convex kink point just below the second central government kink. In the baseline simulation model we will consider the following three-bracket structure of the piece-wise linear marginal tax schedule for 2016-17:

\[
T'(\bar{z}) = \begin{cases} 
\tau_1 & \text{if } \bar{z} \leq z_1^* \\
\tau_2 & \text{if } z_1^* < \bar{z} \leq z_2^* \\
\tau_3 & \text{if } \bar{z} > z_2^*
\end{cases}
\] (C3)

, while the pre-reform system has a similar two-bracket structure. In the baseline specification we omit the non-convex kink that emerged in 2016 at the earnings level at which the entire EITC was phased-out. The lower tax rate in the “top bracket” affected very few taxpayers (around 0.2 percent of the population). We have in fact performed simulations while including the non-convex kink, and the results are fairly similar. Still, we think the simulation results are easier to interpret if the chosen elasticity directly corresponds to the simulated earnings response in the treatment group, which is the case when everyone in the treatment group experiences the same marginal tax change.

C.3 Smoothened marginal tax rates

Following Saez (1999) we showed in Section 5.1 that the optimization problem under uncertainty is similar to an optimization problem under certainty, with the piecewise linear tax function \( T(z) \) replaced by the new “effective” tax function

\[
\hat{T}(z) = \int T(z + \epsilon)f(\epsilon)d\epsilon = \int T(\bar{z})f(\bar{z} - z)d\bar{z}
\]

The effective marginal tax rate can be expressed as

\[
\hat{T}'(z) = \int T(\bar{z}) \frac{\partial f(\bar{z} - z)}{\partial \bar{z}} d\bar{z}.
\]

Since \( \frac{\partial f(\bar{z} - z)}{\partial z} = -\frac{\partial f(\bar{z} - z)}{\partial \bar{z}} \) we can write \( \hat{T}'(z) = -\int T(\bar{z}) \frac{\partial f(\bar{z} - z)}{\partial \bar{z}} d\bar{z} \). Using in-
Integration by parts, we obtain
\[ \hat{T}'(z) = \int T'(\tilde{z}) f(\tilde{z} - z) d\tilde{z} \]  
(C4)

Intuitively, at a given level of earnings, \( z \), the effective marginal tax rate is a weighted sum of the true marginal tax rates of the piece-wise linear tax function, \( T'(\tilde{z}) \). Combining (C3) and (C4) we obtain
\[ \hat{T}'(z) = \tau_1 \int_{0}^{z_1} f(\tilde{z} - z) d\tilde{z} + \tau_2 \int_{z_1}^{z_2} f(\tilde{z} - z) d\tilde{z} + \tau_3 \int_{z_2}^{\infty} f(\tilde{z} - z) d\tilde{z}, \]
(C5)
where \( F(\epsilon) \) is the cumulative density function. To compute the effective marginal tax rate at \( z \), \( \hat{T}'(z) \), it is hence sufficient to consider the kink points of the piece-wise linear schedule and the cdf of the normal distribution with standard deviation \( \sigma \). We use (C5) to compute the effective marginal tax schedule 2016-17. A similar expression, but with one kink only, i.e. \( \hat{T}'(z) = \tau_1 F(z_1^* - z) + \tau_2 (1 - F(z_1^* - z)) \), is used for 2012-15.

C.4 The skill distribution

We want the simulated pre-reform earnings distribution to have similar properties as the empirically observed one. Therefore, we calibrate a skill (potential income) distribution for 2015 that generates the pre-reform earnings distribution. Our approach roughly follows Saez (2001), who simulated optimal tax schedules and recalibrated the skill distribution for different values of the elasticity. As we show below in Section C.6, when agents optimize subject to a smooth tax schedule the following endogenous relationship between optimal \( z \) and skill \( z_0 \) holds in the individual’s optimum
\[ z_0(z) = \frac{z}{[1 - \hat{T}'(z)]^e}. \]  
(C6)
When calibrating skills, we treat the empirically observed realized earnings distribution as a proxy for the distribution of deterministic earnings, \( z \). We recover the frequency distribution of \( z_0 \) by plugging in the smoothened 2015 marginal tax rate \( \hat{T}'(z) \) (computed for \( \sigma = \text{SEK 70,000} \), \( e = 0.2 \), and \( z \) into (C6). We have verified that there is a one-to-one mapping between \( z_0 \) and \( z \).
We use the calibrated skill distribution for 2015 to obtain similar distributions for 2012-14 and 2016-17. We assume homogenous skill growth across all skill types. Moreover, we assume that potential incomes grow at a constant rate all years 2012-17, the growth rate is $g = 2.4\%$. This value corresponds to the average earnings growth for pre-reform years 2012-2015. As long as we impose $e = 0.2$ we will use the same calibrated skill distribution when solving the model, regardless of the value of $\sigma$.

C.5 Solving the model – piece-wise linear tax function

When $\sigma = 0$ agents solve

$$U = z - T(z) - \frac{z_0}{1 + \frac{1}{e}} \left( \frac{z}{z_0} \right)^{1 + \frac{1}{e}}$$

, where $T(z)$ is a piece-wise linear tax function. In this environment agents may have optima at interior points of segments, or they may have optima at convex kink points, where marginal tax rates increase. We solve for the individuals’ optima numerically, and our own code builds on MATLAB scripts originally constructed by Spencer Bastani for the simulations in Bastani and Selin (2014). The optimization routine finds the value of $z$ that maximizes indirect utility. The set of optimal solutions contains two parts: the interior solution of realized income is $z = (1 - \tau)^{e} z_0$, and the bunching solution for individuals with skill level $z_0 \in [z^*/(1 - \tau_1)^{e}, z^*/(1 - \tau_2)^{e}]$ is $z = z^*$ where $z^*$ is the kink point, and $\tau_1$ and $\tau_2$ are the marginal tax rates before and after the kink respectively. With a three-bracket structure, as shown in (C3), the optimization problem is identical to the two-bracket case. The highest income earner who bunches at the kink point $z_1^*$ has potential income $z_1^*/(1 - \tau_2)^{e}$. The highest income earner who bunches at the next kink point $z_2^*$ has potential income $z_2^*/(1 - \tau_3)^{e}$. Individuals with potential income $z_0 \in (z_1^*/(1 - \tau_2)^{e}, z_2^*/(1 - \tau_3)^{e}]$ have optima at either the second segment or kink $z_2^*$.

C.6 Solving the model – smooth tax function

In Section 5.1 we demonstrated that the uncertainty model is equivalent to a setting in which agents choose expected earnings, $z$, subject to a smooth
tax schedule $\hat{T}(z)$:

$$EU = z - \hat{T}(z) - \frac{z_0}{1 + \frac{1}{\hat{\tau}}} \left( \frac{z}{z_0} \right)^{1 + \hat{\tau}}.$$  

Individuals now choose optimal deterministic income $z$, and realized income $\tilde{z}$ is unknown. The first order condition can be written

$$1 - \hat{T}'(z) = \left( \frac{z}{z_0} \right)^{\frac{1}{\hat{\tau}}}.$$  \hspace{1cm} (C7)

As the left hand side is endogenous, it is not possible to obtain an analytical solution for optimal $z$. We instead use a simulation approach to solve this problem. For each observation with skill $z_0$, we loop over all values $z$ to find the unique value that satisfies (C7). As mentioned in Appendix Section C.3 above, we have verified that there is a one-to-one mapping between $z_0$ and $z$ for chosen values of $\hat{e}$ and $\sigma$. 