Int Tax Public Finance (2016) 23:762–783
DOI 10.1007/s10797-016-9412-5

POLICY WATCH

# Unequal inequalities: Do progressive taxes reduce income inequality?

**Denvil Duncan**[1] · **Klara Sabirianova Peter**[2,3]

Published online: 17 June 2016


**Abstract** This paper analyzes the effect of changes in the structural progressivity of national income tax systems on observed and actual income inequality. Using several unique measures of progressivity over the 1981–2005 period for a large panel of countries, we find that progressivity reduces inequality in observed income, but has a significantly smaller impact on actual inequality, approximated by consumption-based Ginis. An empirical comparative analysis shows that the differential effect on observed versus actual inequality is much larger in countries with weaker legal institutions. We also find that structural progressivity has a greater equalizing effect in environments that support pro-poor redistribution. Substantial differences in inequality response to changes in top versus bottom rates are also uncovered.

**Keywords** Income inequality · Gini · Personal income tax · Structural progressivity · Tax evasion

**JEL Classification** H2 · I3 · J3 · O1 · O2

**Electronic supplementary material** The online version of this article (doi:10.1007/s10797-016-9412-5) contains supplementary material, which is available to authorized users.

✉ Denvil Duncan
duncande@indiana.edu

Klara Sabirianova Peter
kpeter@unc.edu

1 School of Public and Environmental Affairs, Indiana University, SPEA 375F, 1315 East 10th Street, Bloomington, IN 47403, USA

2 Department of Economics, University of North Carolina - Chapel Hill, 107 Gardner Hall, CB 3305, Chapel Hill, NC 27599, USA

3 CEPR, and IZA, Bonn, Germany

## 1 Introduction

A major concern occupying the attention of policymakers across the world is the perceived widening income gap between the rich and the poor. Although there remains some debate, there appears to be a general consensus that within-country income inequality has increased significantly in recent decades (Violante 2010).[1] In an effort to address this problem, some policymakers have recently argued for and/or enacted more progressive income tax schedules. For example, top statutory personal income tax rates increased to 39.6 % in the USA and 45 % in France. Tax policy is commonly viewed as one possible solution to the problem of growing income inequality because higher top tax rates and the resulting increase in progressivity imply that the rich pay a relatively larger share (compared to the poor) of their observed before-tax income to the government in taxes. As a result, observed after-tax income tends to be more equally distributed. This redistributive effect is further strengthened if the tax revenues are redistributed to the poor.

Although this result is intuitive, focusing on the distribution of actual income, which reflects resources available to individuals, rather than observed income, arguably provides a more accurate picture of the distributional effects of tax progressivity. More importantly, responses to taxation such as tax evasion and tax avoidance imply that the effect of tax progressivity on the distribution of actual after-tax income is likely to be different from that on observed after-tax income. Therefore, it is not immediately clear that increased tax progressivity necessarily leads to more equal distribution of resources available to people.

In this paper, we seek to determine the relationship between structural progressivity of personal income taxes[2] and income inequality, with a special emphasis on the differential effect of progressivity on observed versus actual inequality. Although a lot of work has been done to assess the impact of tax reforms on inequality, this is one of the first known attempts to differentiate between these two effects.[3] Verification of this possible differential effect is becoming increasingly important given the number of countries that have raised or are considering raising tax rates after almost three decades of tax reforms aimed at flattening tax schedules (Sabirianova Peter et al. 2010). More recent reform efforts, for example in the USA, include higher tax rates aimed at increasing the tax burden of the rich as well as easing budget deficits. If progressivity and income inequality are negatively related, then there are important implications of such policies for the distribution of income. However, it is not clear that reducing (increasing) the progressivity of income tax schedules will necessarily lead to greater (lower) levels of income inequality in the presence of tax evasion.

[1] A recent paper by Lakner and Milanovic (2015) finds that global inequality declined during the period 1988 to 2008 even as within-country income inequality increased.

[2] The term *structural progressivity* denotes changes in the average (or marginal) tax rate along the income distribution.

[3] Duncan (2014) and Doerrenberg and Duncan (2013) address a similar question to ours using different datasets and empirical techniques. Duncan (2014) uses data from the Russian Longitudinal Household Survey to simulate the redistributive effects of the Russian 2001 personal income tax reform, while Doerrenberg and Duncan (2013) use data from a laboratory experiment to estimate the impact of tax rate changes on income inequality.

Another important contribution of this paper is that we use a unique dataset for a large panel of countries that contains time-varying country-specific measures of structural progressivity of national personal income tax systems over the period 1981–2005. We develop and estimate several measures of structural progressivity for over one hundred countries worldwide by using complete national income tax schedules with statutory rates, thresholds, country-specific tax formulas, and other information. The measures are based on data definitions that are compatible across countries as well as over time. Having such a large dataset is advantageous because it allows us to conduct a comparative analysis that provides useful insights into how the effect of tax policy varies between countries with different institutional characteristics. Unlike our work, most of the existing research in this area tends to be country-specific incidence studies that rely on micro-simulation exercises or computable general equilibrium models (Gravelle 1992; Martinez-Vazquez 2008; Duncan 2014). Our paper also differs from Doerrenberg and Duncan (2013) who use data generated in a laboratory setting.

We acknowledge that macro-analysis has certain limitations as we are not able to examine within-country heterogeneity in individual responses or directly estimate the tax evasion effect on income inequality. We also cannot account for the possible offsetting effects of other taxes.[4] Nevertheless, macro-data provide an excellent opportunity for examining the relationship between structural progressivity and income inequality on a large international scale and for cross-country comparisons in testing several important hypotheses.

Our key prediction is that progressivity affects observed inequality differently than it does true inequality, and that the difference between the two inequality effects is increasing with the extent of tax evasion and its responsiveness to tax changes. To test this hypothesis, we use a country-level dataset of Gini coefficients calculated separately for net income and consumption. We argue that the consumption-based measure of income is closer to true permanent income in comparison with disposable income reported in the household surveys.[5] Our empirical analysis reveals that while progressivity reduces observed inequality in reported net income, it has a significantly smaller impact on inequality in consumption. We show that the differential effect of progressivity on inequality in income versus consumption is increasing with weaker legal institutions that can trigger a large tax evasion response. We find that weaker law and order is associated with significantly smaller negative (and sometimes even positive) effects of progressivity on inequality in consumption. We also find that progressivity changes at the top of the tax schedule are more effective in reducing observed net income inequality than progressivity changes at the bottom, and that the equalizing effect of progressivity on observed income is stronger in the presence of democratic institutions that facilitate pro-poor redistribution.

[4] It is important to emphasize that we focus on the personal income tax only. As such, any equity offsets that may come from other taxes such as the corporate income tax or sales taxes are not taken into account. In principle, policymakers could achieve the same level of income inequality by substituting reduced progressivity of the personal income tax with increased progressivity of the corporate tax.

[5] The empirical micro-literature on developing countries has long pointed out the unreliability of income measures in household budget surveys due to widespread underreporting and called for the use of consumption-based measures of inequality (e.g., see Deaton 1997; Milanovic 1999).

These results suggest that the trade-off between efficiency and equality varies between countries at various stages of development. Because tax evasion is so pervasive in developing countries, our results lead us to speculate that developing countries face much lower equality penalty for increasing efficiency. In other words, compared to developed countries, developing countries with their higher levels of tax evasion, experience a much smaller decline in equality for each increase in efficiency. This difference in the equality efficiency trade-off between developed and developing countries is a likely explanation for the relatively greater popularity of linear personal income tax schedules in developing and transition countries.

The paper proceeds as follows: Sect. 2 provides the theoretical discussion of the effect of progressivity on inequality and describes our implementation strategy. Following that is a description of the data in Sect. 3, the empirical model and results in Sect. 4, and the conclusion in Sect. 5.

## 2 Predictions and implementation strategy

### 2.1 Predictions

Progressive taxes are designed to collect a greater proportion of income from the rich relative to the poor, thus reducing the inequality of disposable income compared to taxable income. However, as the government increases structural progressivity or tax rates facing the rich relative to the poor, individuals may respond by taking steps to reduce their taxable income. Reducing taxable income is achieved by either a productivity response (such as working less) or a tax evasion/avoidance response (i.e., simply reporting a smaller share of true income).[6] It is generally accepted that productivity responses are relatively small compared to tax evasion responses (Slemrod 1995; Gorodnichenko et al. 2009). Additionally, there is evidence that the tax evasion response is much stronger in the upper tail of the income distribution (Slemrod 1994; Feldstein 1995). In other words, the rich tend to be more sensitive to changes in the tax rates because they are better able to hide their income.[7] These two findings suggest that the negative effect of higher and more progressive taxes on observed income inequality will overstate (in absolute terms) their effect on actual income inequality, and that the magnitude of the differential effect increases with evasion. We illustrate these possibilities more formally in Online Appendix B using the Kuznets ratio as a measure of inequality.

To see these predictions more clearly and the implied testable hypotheses, we write observed inequality ($I^o$) and actual inequality ($I^*$) as follows:

[6] We use the term avoidance to refer to all legal *non-real* responses to taxation such as income shifting, timing, and accounting responses. We acknowledge that evasion and avoidance are different, and our primary interest is in the tax evasion response. Because it is often difficult to distinguish evasion responses from avoidance responses empirically, we use the terms interchangeably from here on.

[7] We acknowledge that the level of evasion tends to be more prominent at both tails of the income distribution. However, this should not have any qualitative effect on our results as long as the responsiveness of evasion is relatively higher in the right tail of the income distribution.

$$I^o = \alpha^o + \delta^o P + \beta^o X + \mu^o \tag{1}$$

$$I^* = \alpha^* + \delta^* P + \beta^* X + \mu^* \tag{2}$$

Taking the difference between Eqs. (1) and (2) yields:

$$I^o - I^* = \alpha + (\delta^o - \delta^*)P + \beta X + \mu, \quad E[X\mu] = 0, \quad E[P\mu] \neq 0, \tag{3}$$

where $P$ is the measure of structural tax progressivity, $X$ is the set of exogenous country characteristics uncorrelated with the error terms $\mu^o$ and $\mu^*$, $\alpha = \alpha^o - \alpha^*$, $\beta = \beta^o - \beta^*$, and $\mu = \mu^o - \mu^*$.

Our predictions from the decomposition framework in Online Appendix B.1. imply the following hypotheses: (H1) $\delta^o < 0$; (H2) $\delta^o$ is likely to be smaller (more negative) in countries that facilitate pro-poor government transfers; (H3) $\delta^o < \delta^*$; and (H4) $\triangle \delta (=| \delta^o - \delta^* |)$ increases with the responsiveness of evasion to changes in progressivity.

## 2.2 Implementation strategy

The major issue in testing these hypotheses is approximating reported and actual income in a cross-country setting. Several micro-studies argue that income reported in household surveys is close to the income reported for taxation purposes, but that expenditure levels often exceed that of reported income (Gorodnichenko et al. 2009; Engstrom and Holmlund 2009; Lemieux et al. 1994). (Dis)saving is unlikely to explain these persistent gaps between consumption and income over time. For example, the consumption–income gap for Russia reported in Fig. 1 would require a savings rate of approximately negative 30 % over the 16-year period (Gorodnichenko et al. 2009). Deaton (1997) outlines several reasons why expenditures or consumption data are more reliable than reported income and much closer to actual permanent income.

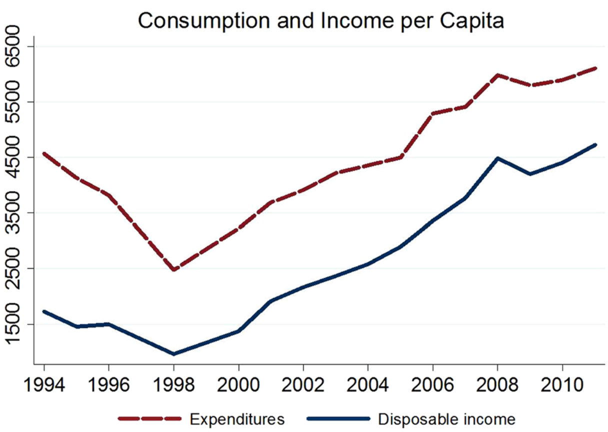


**Fig. 1** Consumption–income gap for Russia, 1994–2010. Amounts are in rubles

Following the arguments in Deaton (1997), we use the available country-level inequality indices calculated from the income reported in typical household surveys, $I^Y$, to approximate $I^o$ and employ the inequality indices based on household expenditures, $I^C$, as noisy measures of $I^*$. The noise in both cases is normalized to zero mean and assumed to be uncorrelated with $X$, but allowed to be correlated with progressivity $P$, such that

$$\begin{aligned} I^o - I^Y &= \xi^Y, \quad E[\xi^Y] = E[X\xi^Y] = 0 \\ I^* - I^C &= \lambda F + \xi^C, \quad E[\xi^C] = E[X\xi^C] = E[F\xi^C] = 0, \end{aligned} \tag{4}$$

where $F$ stands for the observed factors explaining the difference between $I^*$ and $I^C$. For example, more developed financial institutions can better insure household consumption against unanticipated income shocks, so $I^*$ might be larger than $I^C$ in more advanced credit markets.

By substituting Eqs. (3) into (4), we obtain:

$$I^Y - I^C = \alpha + (\delta^o - \delta^*)P + \beta X + \lambda F + \tilde{\mu}, \tag{5}$$

where $\tilde{\mu} = (\mu^o - \mu^*) - (\xi^Y - \xi^C)$, $E[\tilde{\mu}] = E[X\tilde{\mu}] = E[F\tilde{\mu}] = 0$, $E[P\tilde{\mu}] \neq 0$. Tax progressivity could be potentially endogenous if it is correlated either with the "model" component of the error term $(\mu^o - \mu^*)$ or with the measurement error component $(\xi^Y - \xi^C)$. We address endogeneity issues in more detail after the description of data in Sect. 4.

The second implementation issue that we face is measuring tax evasion when testing hypothesis H4. Since country-level measures of evasion are nonexistent, we capture the essence of this hypothesis by relying on proxy factors that are likely to be positively correlated with the evasion elasticity. These include indices of law and order and corruption.[8] We then reformulate our fourth hypothesis in that the differential effect is likely to be larger in countries with higher corruption and weaker law and order.

Finally, we rely on observed country characteristics that tend to be positively associated with pro-poor redistribution policies in order to test the "redistribution" hypothesis (H2). For example, Chong and Gradstein (2007) show that income inequality is lower in countries with higher civil liberties and political rights, suggesting that these factors are associated with pro-poor redistribution.[9] This implies that institutional characteristics such as democratic political systems and larger social liberties are likely to reinforce the negative effect of structural tax progressivity on observed income inequality, i.e., $\delta^o$ is expected to be more negative.

[8] For example, Slemrod and Kopczuk (2002) and Kopczuk (2005) make the point that the elasticity of taxable income is a function of the characteristics of the tax system and is larger in environments with many evasion/avoidance opportunities.

[9] Meltzer and Richard (1981) develop a theoretical model of government spending and show that an expansion of political franchise freedom leads to greater redistribution and larger government. Gradstein and Milanovic (2004) survey a number of other studies with similar findings.

# 3 Measuring inequality and structural progressivity

## 3.1 Income inequality measure

We test the hypotheses developed in the previous section using country-level Gini coefficients obtained from the World Institute for Development Research (WIDER v.3.3), the World Development Index (WDI), the International Labor Office (LABORSTA), and the Review of Economic Dynamics (RED).[10] We rely on the Gini coefficient as our measure of inequality because it has the widest country-year coverage of all the available measures of inequality. The Gini coefficients are based on income and consumption data collected in household surveys by the respective countries. Because surveys vary across countries and time, numerous steps are taken to ensure accuracy and consistency. In particular, careful attention is paid to the types of incomes covered, the use of equivalence scales, and the representativeness of the underlying surveys. Furthermore, each Gini is assigned a quality rating (high, average, and low) based upon area and population coverage, income concepts, survey design, sampling methods, missing information, and overall data quality.

For the purpose of our analysis, we select all Gini coefficients that are based on disposable (net) income and expenditures or consumption. There are some country-years for which only gross income Ginis are available; in these cases, we use gross income Gini as a proxy for net income Gini.[11] Disposable income covers wages and salaries, self-employed income, and property income and is after taxes and transfers. Ginis are classified as expenditure based if durables are included at purchase value and as consumption based if durables are either excluded or included at use value. We exclude Ginis based on ambiguous definitions (e.g., mixed income–consumption) and with limited coverage (e.g., reported for rural areas or country's capital). For a given country, year, and income base, we select one measure using the following set of preferences: National estimates are preferred to urban, rural, and other area coverage estimates, equivalence scales or per capita adjustment is chosen over no or unknown adjustment, higher-quality ranking is preferred to lower-quality ranking.[12]

This selection process leaves us with 2001 Gini estimates for 165 countries from 1982 to 2005. The majority of the estimates meet the best practices as set out by the WIDER including adjustment for household size, national coverage, and quality rat-

[10] The RED data are taken from papers published in the special issue on Cross Sectional Facts for Macroeconomists; Volume 13, Issue 1 (January 2010). It includes data from Canada, Germany, Italy, Mexico, Russia, Spain, Sweden, UK, and the USA.

[11] This represents less than 15 % of the sample. Alternatively, we impute missing net Gini using predicted values from the linear regression of net Gini on gross Gini, time period fixed effects, region fixed effect, standard macroeconomic variables (log of per capita GDP, log of population, government size, and inflation), and characteristics of the country's tax system such as the type of tax allowance (none, standard, complex), an indicator for having local PIT taxes (no local tax, local tax is less than 5 %, and local tax is higher than 5 %), an indicator for having PIT surtaxes, and the number of tax brackets in quadratic form. The explanatory power of the model is fairly high ($R$-squared is 83 %). The results stay the same when the predicted net Gini is used instead of gross Gini, and the results are available upon request.

[12] We also remove a small number of outliers (less than 1 % of the sample) if the reported Gini coefficient deviates from its country-specific linear trend by more than 3 standardized residuals.

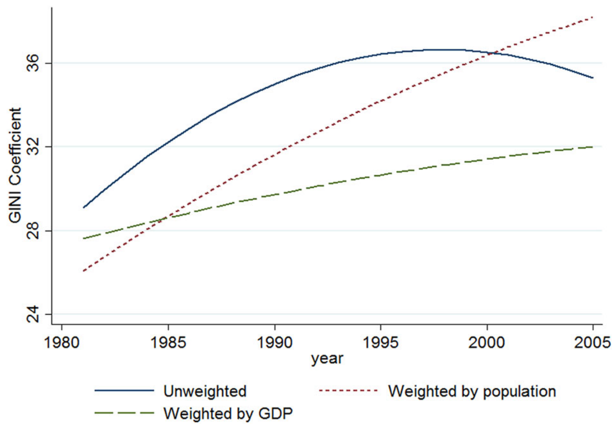


**Fig. 2** Global trend in income inequality, 1981–2005. The trend in income inequality is obtained by regressing the Gini coefficient on a quadratic time trend, controlling for income base, area coverage, income adjustment, and World Bank income classification. A similar, though not identical, procedure is used by Easterly (2007) to address the consistency problem inherent in the Gini data. Country categories are defined using the World Bank country classification based on historical (time-varying) income thresholds

ing. Of the total sample of Gini estimates, 34 % are based on consumption and 66 % on net income. To control for differences in Gini measurement, our estimates include dummy variables for area coverage, household size adjustment, and quality. While we recognize that the use of dummy variables does not eliminate all of the biases resulting from comparability issues (Atkinson and Brandolini 2001), we are constrained by existing inequality estimates. This is especially restricting in cross-country panel studies due to variations in the quality of primary data sources, differences in definition of variables, and other procedures followed by individual countries.

Figure 2 shows that income inequality increased throughout the 1980s and 1990s before declining during the 2000–2005 period. Although the GDP-weighted trend follows the unweighted trend, the population-weighted trend shows income inequality increasing throughout the sample period, which is consistent with rising within-country inequality in China, India, and other developing countries with large populations. We also report summary statistics on the Gini coefficient by income definition across time in Panel A of Table 1 and find patterns that are consistent with the literature (e.g., Easterly 2007).

### 3.2 Tax progressivity measures

In contrast to income inequality, the measures of tax progressivity are not readily available for cross-country comparison. The existing measures implemented in the literature fall into one of three groups: (1) the top statutory PIT rate, (2) effective inequality-based measures of progressivity, and (3) structural progressivity measures. Neither of these measures is perfectly suitable for our analysis in their original form.

Although the top statutory PIT rate has occasionally been used in empirical cross-country research as a proxy variable for tax progressivity, it might be a misleading

**Table 1** Summary statistics: Gini and progressivity

| Income base | 1981–1985 | 1986–1990 | 1991–1995 | 1996–2000 | 2001–2005 |
|---|---|---|---|---|---|
| *Panel A: Gini by income base* | | | | | |
| Consumption | 33.561 | 35.880 | 39.030 | 36.839 | 36.629 |
| | (8.979) | (8.083) | (10.756) | (8.231) | (7.453) |
| | [50] | [77] | [157] | [181] | [219] |
| Net income | 33.203 | 33.424 | 38.206 | 37.299 | 35.539 |
| | (10.771) | (11.038) | (11.788) | (11.011) | (9.015) |
| | [173] | [262] | [304] | [307] | [271] |
| *Panel B: progressivity measures* | | | | | |
| Top PIT rate | 46.768 | 40.397 | 34.263 | 32.493 | 30.109 |
| | (23.300) | (21.199) | (17.478) | (15.795) | (14.735) |
| | [572] | [614] | [736] | [824] | [867] |
| MRP | 5.669 | 5.427 | 4.730 | 4.750 | 4.585 |
| | (4.934) | (4.563) | (3.710) | (3.600) | (3.582) |
| ARP | 4.124 | 4.013 | 3.471 | 3.452 | 3.335 |
| | (4.156) | (3.757) | (3.130) | (2.960) | (2.868) |
| ARP-bottom | 3.754 | 3.791 | 3.216 | 3.238 | 3.176 |
| | (3.711) | (3.488) | (2.924) | (2.812) | (2.842) |
| ARP-middle | 6.032 | 5.758 | 5.051 | 4.934 | 4.695 |
| | (7.511) | (6.602) | (5.772) | (5.470) | (5.176) |
| ARP-top | 2.759 | 2.616 | 2.284 | 2.302 | 2.218 |
| | (2.454) | (2.270) | (1.858) | (1.815) | (1.815) |
| | [459] | [517] | [624] | [734] | [748] |

Standard deviation in parentheses and number of observations in brackets. ARP and MRP are average and marginal tax rate progression for income between 0.4 and 4 times a country's per capita GDP, respectively. ARP-bottom, ARP-middle, and ARP-top are average tax rate progression for income between 0.4–2, 1–3, and 2–4 times a country's per capita GDP, respectively. All non-missing values are used in calculating averages

indicator of progressivity since both proportional and highly progressive tax systems may, in principle, have the same top statutory rate. The effective progressivity measures require information on pre- and post-tax inequality and the distribution of the tax burden (Musgrave and Thin 1948; Kakwani 1977; Suits 1977). Information on these variables is either not available or not comparable across countries. The more serious problem, however, is the issue of simultaneity in determination of income inequality and inequality-based progressivity, which inhibits the identification of the direct effect of tax progressivity on inequality.

From this perspective, the measures of structural progressivity are more suitable for the purpose of our analysis. The term structural progressivity denotes changes in average and marginal rates along the income distribution (Musgrave and Thin 1948). These changes can be identified without knowing after-tax inequality, making the endogeneity problem less severe. However, the calculations require information on gross income distribution, which is difficult to gather in a comparable way at the cross-country level. Because structural progressivity changes along the income distribution, it produces a range of possible measures instead of one. However, we need a single, comprehensive measure of PIT progressivity, which is comparable across countries, is available for a large representative sample of countries, and varies over time. We propose the following procedure to derive such a measure.

The first step in calculating structural progressivity is to obtain tax rates at different points of the income distribution. Instead of actual income distribution, we use a country's per capita GDP and its multiples as a measure of the income base that is comparable across space and time. The GDP figures are rescaled to get 100 units of pre-tax income for each country and year, ranging from 4 to 400 % of a country's per capita GDP. We then apply the tax schedule information to these units of income to obtain tax liability and average and marginal tax rates. The data on national tax schedules are collected for 189 countries from 1981 to 2005 and described in detail in Sabirianova Peter et al. (2010). Here we only note that our average and marginal tax rates take into account standard deductions, basic personal allowances, tax credits, major national surtaxes, multiple schedules, nonstandard tax formulas, and other provisions in addition to statutory rates and thresholds.[13]

The progressivity measures are obtained by regressing average tax rates on gross income using 100 data points that are formed relative to a country's per capita GDP in a given year. The slope coefficient on the income variable measures the percentage point change in the tax rate resulting from a 1 % point change in gross income and is our measure of structural progressivity. The PIT structure is interpreted as progressive, proportional, or regressive if the slope coefficient is positive, zero, or negative, respectively. This procedure gives us average rate progression (ARP) for each country and year in our dataset. In a similar vein, we use marginal rates to compute marginal rate progression (MRP). However, since the concept of tax progressivity generally applies to average taxes, we focus on ARP and use MRP for comparative purposes only. Figure 3 in Online Appendix A illustrates how the MRP is obtained for a hypothetical case with no allowances or other provisions.

Both the average and marginal rate progression measures, as defined by Musgrave and Thin (1948), decline as one moves up the income distribution. In other words, the tax schedule is less progressive at the top of the income scale. In an effort to capture this nonlinearity, we also calculate ARP for the bottom, middle, and top portion of the income scale, respectively.[14]

Panel B of Table 1 reports summary statistics on the six progressivity measures across time for the estimation sample. The pattern that stands out is that all of the measures declined throughout the 1980s and early 1990s and then remained stable during the latter period. In concordance with the nonlinear properties of progressivity (Musgrave and Thin 1948), our measures calculated for the bottom portion of the income scale tend to be larger than those for the full income scale. Panel B of Table 1 also reports summary statistics on the top statutory PIT rate. The top marginal tax rate has declined steadily from a high of 47 % in the 1981–1985 period to a low of 30 %

[13] We do not account for deductions, allowances, and credits that vary by individual characteristics; for example, child credits are not included in our calculations.

[14] This corresponds to 4–200, 100–300, and 200–400 % of per capita GDP, respectively.

in the 2001–2005 period. Since these global trends follow closely those reported in Sabirianova Peter et al. (2010), we refer the reader to that paper for a more detailed description of the changes that have taken place over the 25-year period 1981–2005.

## 4 Empirical methodology and results

This section describes the empirical methodology used to test the hypotheses discussed in Sect. 2 and presents our findings. We begin with hypothesis H1 by modeling inequality as:

$$I_{it} = \beta_o + \delta P_{it-1} + \beta_1 Z_{it-1} + \beta_2 W_{it} + \zeta_t + \epsilon_{it}, \tag{6}$$

where $I_{it}$ is the Gini coefficient for reported income in country $i$ and year $t$, $P_{it-1}$ is the relevant measure of PIT progressivity, $Z_{it-1}$ is a vector of control variables (discussed below), $W_{it}$ is a vector of auxiliary variables, $\zeta_t$ captures time effects, and $\epsilon_{it-1}$ is the error term. Vector $W$ controls for consistency of the Gini measures and includes a set of dummy variables for national area coverage, quality of Gini measure, and equivalent scale adjustment. The key parameter of interest, $\delta$, captures the effect of progressivity on reported income inequality.

### 4.1 Identification issues

There are several econometric issues that have to be addressed before we can test our hypotheses. To begin with, our measure of structural progressivity is likely to be endogenous for multiple reasons. First, the error term can be correlated with structural progressivity due to reverse causality. The political economy literature has long established a reverse relationship between income inequality and taxes (Meltzer and Richard 1981; Persson and Tabellini 2002; Borge and Rattso 2004; Adam et al. 2014). This reverse causality implies that OLS estimation of $\delta$ is likely to be biased toward zero.

Second, the use of fixed effects is problematic given the limited within variation in the dependent variable for some countries; the Gini data are mostly sparse for a number of countries in our sample.[15] To the extent that constant country characteristics are correlated with the error term, omitted fixed effects create an endogeneity bias. Finally, structural progressivity is an estimated parameter with associated standard errors, which can lead to attenuation bias in the estimated coefficients, assuming standard errors follow the properties of the classical error-in-variables problem.

These endogeneity issues are addressed by (1) using the lagged value of progressivity measures and other macro-variables and (2) exploiting the spatial correlation in tax rates among neighboring countries to create instrumental variables (IVs) (Lee and Gordon 2005; Ferede and Dahlby 2012). There are several channels through which this spatial correlation might arise. First, PIT rates will be correlated if countries actively

[15] While the between-countries standard deviation of the net income Gini is 11.6, the within-country standard deviation is only three points. Comparable numbers for the consumption Gini are 9.3 and 3, for between and within standard deviations, respectively.

compete for the PIT base (Kleven et al. 2013; Egger and Radulescu 2009). Spatial correlation may also arise through policy diffusion (Whalley 1990) and yardstick competition (Besley and Case 1995).

We first construct IVs by assigning each contiguous neighbor an equal weight and calculating the mean of various tax progressivity measures for the reference country's neighbors. Theoretically, we require tax variables in country $i$ to be correlated with tax variables in neighboring country $j$ and orthogonal to the error term in Eq. (6). While the strength of the correlation can be established empirically (see discussion Sect. 4.2), we must rely on intuitive arguments to support the orthogonality condition. For example, one may argue that an inflow of the PIT base in response to higher tax rate in neighboring countries will cause neighbors' tax rate to have a direct effect on inequality in country $i$, thus violating the orthogonality condition. However, the level of migration required for this to affect the orthogonality condition in any meaningful way likely outweighs observed flows of labor between countries. For example, according to the OECD's International Migration Outlook (2011) (OECD 2011), net migration in selected OECD countries peaked at approximately 0.3 % of total resident population in the 2001–2006 period. Therefore, we do not believe this is a major concern for identification.

We also use a number of alternative IVs, including the 1-year lag of neighbors' progressivity, and average progressivity measures for "neighbors of the neighbors," which represent the next layer of countries in the spatial dimension. Allowing the IVs to be based on neighbors from different points in space and time should help to break correlation between the IVs and the error term. Finally, we include a 1-year lagged indicator of complexity of neighboring countries' tax systems as an additional IV in some specifications.

In an effort to minimize bias due to omitted fixed effects, we control for a variety of country characteristics in vector $Z$.[16] We include country size measured as log of population, cultural background proxied by the major religion of a country, average inflation rate, and 1-year lagged log of per capita GDP in quadratic form (see Table 8 in Online Appendix A for variable definitions). Inflation is included to account for the possible equalizing effect of price stability on the distribution of income (Minarik 1979; Bulr 2001). The quadratic GDP term is added to account for the existence of the Kuznets curve which postulates there is a nonlinear (inverted U) relationship between income inequality and per capita GDP. We also control for the ratio of financial deposits to GDP and for the interest spread as proxies for financial development since a well-developed financial sector is likely correlated with both progressivity (Gordon and Li 2007, 2009) and inequality (Clarke et al. 2006; Law and Tan 2009; Mookerjee and Kalipioni 2010; Kim and Lin 2011).[17]

[16] This allows us to control for time constant variables that are likely correlated with our measures of progressivity. We acknowledge that this approach does not fully solve the problem and is a limiting feature of the empirical analysis. It is worth noting that a similar approach is used by Dobson and Ramlogan-Dobson (2012) who use the same inequality measure as we do. Table 7 in Online Appendix A reports summary statistics for the covariates.

[17] We use financial sector deposit as a share of GDP and interest rate spread as our primary measure of financial development (Beck et al. 2000), and the ratio of deposit money bank assets to all bank assets (central bank plus money bank) as alternative measures in robustness checks. We also replace inflation rate

**Table 2** Effect of progressivity on income inequality: IV estimates

| | Top rate | MRP | ARP | ARP | | |
|---|---|---|---|---|---|---|
| | | | | Bottom | Middle | Top |
| *Panel A: unweighted* | | | | | | |
| First-stage diagnostics | | | | | | |
| IV1-neighbor ARP-mid | 1.137** | 0.265** | 0.383*** | 0.844*** | 0.243* | 0.116* |
| | (0.525) | (0.130) | (0.113) | (0.200) | (0.127) | (0.064) |
| Under-ID $p$ value | 0.035 | 0.049 | 0.002 | 0.000 | 0.060 | 0.076 |
| $F$-stat (first stage) | 4.691** | 4.127** | 11.569*** | 17.885*** | 3.685* | 3.285* |
| Second-stage results | | | | | | |
| Progressivity | −1.370** | −5.883** | −4.069*** | −1.846*** | −6.406* | −13.465* |
| | (0.635) | (2.979) | (1.306) | (0.490) | (3.460) | (7.522) |
| $F$-stat | 1.763** | 4.396*** | 12.181*** | 10.614*** | 6.258*** | 2.771*** |
| *Panel B: weighted* | | | | | | |
| First-stage diagnostics | | | | | | |
| IV1-neighbor ARP-mid | 0.955 | 0.406*** | 0.456*** | 0.784*** | 0.431*** | 0.176** |
| | (0.623) | (0.144) | (0.132) | (0.240) | (0.130) | (0.068) |
| Under-ID $p$ value | 0.158 | 0.025 | 0.011 | 0.012 | 0.017 | 0.032 |
| $F$-stat (first stage) | 2.349 | 7.980*** | 11.951*** | 10.669*** | 10.951*** | 6.725** |
| Second-stage results | | | | | | |
| Progressivity | −1.247 | −2.935** | −2.610** | −1.520** | −2.766** | −6.756** |
| | (0.851) | (1.388) | (1.095) | (0.614) | (1.293) | (3.266) |
| $F$-stat | 4.446*** | 9.620*** | 20.011*** | 15.212*** | 13.889*** | 7.656*** |

Robust standard errors clustered at the country level are in parentheses
* Significant at 10 %; ** significant at 5 %; *** significant at 1 %. Estimates are based on Eq. (6) with the dependent variable defined as Gini index for net income. The measure of structural progressivity is indicated by column titles. The following covariates are included, but not shown: log of population, the quadratic function of the log of GDP per capita, country's major religion, average inflation rate, the share of financial deposits in GDP, the interest rate spread, time period dummies, and a vector of auxiliary variables for the quality of Gini measures. See Table 9 in Online Appendix A for a complete specification. $N = 1012$

## 4.2 Results for observed income inequality (H1)

We test hypothesis H1 by estimating Eq. (6) with six measures of progressivity: the top statutory PIT rate, MRP, ARP, ARP-bottom, ARP-middle, and ARP-top. Panel A of Table 2 starts with a summary of the first-stage results obtained from estimating our baseline model with ARP-middle of neighboring countries as the IV. The estimated coefficient on the instrument is statistically different from zero in five specifications, indicating that the IV is strongly correlated with our endogenous variable. This conclusion is supported by the $F$ values for the excluded instrument, which allows us to reject

Footnote 17 continued
with currency depreciation rate in robustness checks. These changes do not affect the results, which are available upon request.

the null that the instrument has no explanatory power in the first-stage models. Finally, we reject the null that first-stage equations are under-identified based on results from the Kleibergen and Paap Lagrange multiplier under-identification test (Kleibergen and Paap 2006). Together, these first-stage results provide evidence that ARP-middle is well suited as an IV for our purposes.

The second-stage results presented in Panel A indicate that a 1-unit increase in either MRP or ARP reduces the Gini coefficient by 4–6 points, which is about a half standard deviation of the Gini coefficient in our estimation sample.[18]

We also find that increasing progressivity at the top of the income scale is a more effective method of reducing inequality in observed income. The estimated effect ranges from −1.8 for ARP-bottom to −13.5 for ARP-top. This result is consistent with our theoretical framework under the plausible assumption that the rich have a higher combined productivity and evasion response than the poor.

The above findings are supported by several modifications to the model. First, we use the 1-year lag of neighbors' ARP-middle, $\overline{P}^{n}_{i,t-1}$, as our IV to account for the possibility that countries do not adjust their tax rates in the same year as their neighbors (Panel A of Table 3). Second, we define our instrumental variable, ARP-middle, using the neighbors of a country's neighbors, $\overline{P}^{nn}_{i,t}$, where *nn* represents the next tier of non-contiguous countries in the spatial dimension (Panel B of Table 3). This approach should strengthen our assumption that the IV is exogenous. We note that in some models with $\overline{P}^{nn}_{i,t}$, the correlation between $P_{i,t}$ and $\overline{P}^{nn}_{i,t}$ is smaller than the correlation between $P_{i,t}$ and $\overline{P}^{n}_{i,t}$, which may explain higher standard errors. The estimated inequality responses are statistically significant at the 10 % level (with one exception), but estimates are a bit larger than in the baseline specification reported in Panel A of Table 2.

Third, we use contiguous neighbors' ARP and MRP for the full income scale as well as neighbors' ARP-bottom, ARP-middle, and ARP-top as IVs and find that the estimates are similar regardless of the instrument used (Panel C of Table 3). We acknowledge that our IV estimates reflect the local average treatment effect (LATE) of the progressivity on observed income inequality, and that the inequality response may not be same for every country, i.e., there is an underlying distribution of $\delta_i's$ that may vary with certain country characteristics. For example, as we show in subsequent sections, the inequality response is estimated to be higher in absolute value in countries with more developed democratic and legal institutions. If countries' reaction to the instrument depends on their $\delta_i$, the LATE estimate of inequality response might be sensitive to the choice of IVs (see Imbens and Angrist 1994).

The reported results reflect the composition of our estimation sample that is biased toward more developed countries. While we have Ginis for many low and lower middle-income countries (32 % of our sample), their fraction still does not match the 60 % of world country composition. Since developed countries are more likely to have a larger effect of progressivity on inequality, the overall effect in the representative

[18] Table 9 in Online Appendix A presents results for the complete OLS and IV specification of the baseline model with MRP and ARP as measures of progressivity. The OLS estimates of progressivity are negative and statistically significant in ARP specification, but they are much smaller than the IV estimates, as implied in Sect. 4.1. The other coefficients are mostly consistent with expectations.

**Table 3** Effect of progressivity on observed income inequality: alternative IVs

| IV | Top rate | MRP | ARP | ARP | | |
|---|---|---|---|---|---|---|
| | | | | Bottom | Middle | Top |
| *Panel A: lagged IV (contiguous neighbor), $N = 997$* | | | | | | |
| Neighbor ARP-middle | −1.439** | −7.864*** | −5.050*** | −2.030*** | −12.401* | −15.918*** |
| | (0.583) | (2.604) | (1.145) | (0.391) | (7.531) | (5.706) |
| *Panel B: IV (neighbor of neighbor), $N = 1008$* | | | | | | |
| Neighbor ARP-middle | −1.167** | −8.459** | −5.039*** | −1.942*** | −13.678 | −17.215* |
| | (0.471) | (4.076) | (1.208) | (0.388) | (9.846) | (8.906) |
| *Panel C: IV (contiguous neighbor), $N = 1012$* | | | | | | |
| Neighbor MRP | −1.558* | −5.967* | −4.077*** | −1.904*** | −5.630** | −14.349 |
| | (0.917) | (3.352) | (1.275) | (0.484) | (2.517) | (9.558) |
| Neighbor ARP | −1.134*** | −6.107*** | −3.808*** | −1.650*** | −6.519*** | −14.097** |
| | (0.369) | (2.214) | (0.904) | (0.346) | (2.444) | (5.783) |
| Neighbor ARP-bottom | −0.998*** | −6.141*** | −3.640*** | −1.529*** | −6.863*** | −13.942*** |
| | (0.285) | (1.913) | (0.778) | (0.311) | (2.234) | (5.040) |
| Neighbor ARP-middle | −1.370** | −5.883** | −4.069*** | −1.846*** | −6.406* | −13.465* |
| | (0.635) | (2.979) | (1.306) | (0.490) | (3.460) | (7.522) |
| Neighbor ARP-top | −1.779 | −6.109 | −4.035*** | −1.903*** | −5.240** | −15.343 |
| | (1.395) | (3.941) | (1.363) | (0.528) | (2.373) | (12.311) |

Robust standard errors clustered at the country level are in parentheses
* Significant at 10 %; ** significant at 5 %; *** significant at 1 %. Estimates are based on Eq. (6) with the dependent variable defined as Gini in net income. Each specification is similar to that in panel A of Table 2, but only progressivity coefficient in the second stage is reported. The measure of structural progressivity is indicated by column titles. IV in panels A and B is ARP-middle of neighbors. IVs in panel C are indicated by the name of each row. All models pass the Kleibergen and Paap Lagrange multiplier under-identification test. The $F$ test for significance of IVs in the first stage is statistically different from zero in all models

world sample should be smaller. We create sample probability weights to check if the composition of our estimation sample influences the results (see Online Appendix E.2 for details). The results are reported in Panel B of Table 2. Indeed, the effect of progressivity on inequality is smaller in weighted estimates, but the overall conclusions stay the same.

### 4.3 Role of democratic institutions (H2)

In order to test the second hypothesis, we extend the baseline equation model in Eq. (6) to include an interaction term between the progressivity measures and democratic institutional variables. In other words, we estimate the following model:

$$I_{it} = \beta_0 + \delta_0 G_{it} + \delta_1 P_{it} + \delta_2 P_{it} * G_{it} + \beta_1 Z_{it} + \beta_2 W_{it} + \zeta_t + \epsilon_{it}, \quad (7)$$

where $G_{it}$ is an indicator of pro-poor redistribution; all other variables are as defined above. Following our discussion in Sect. 2, we use two democratic institutional vari-

**Table 4** Role of redistribution

| | Top rate | MRP | ARP | ARP | | |
|---|---|---|---|---|---|---|
| | | | | Bottom | Middle | Top |
| *Panel A: political rights* | | | | | | |
| Progressivity | −2.742* | −9.853*** | −7.833*** | −4.216*** | −13.680* | −21.507** |
| | (1.465) | (3.721) | (2.460) | (1.447) | (7.132) | (8.841) |
| Interaction | 0.944* | 3.012*** | 2.281*** | 1.349** | 4.568* | 6.787** |
| | (0.550) | (1.111) | (0.750) | (0.541) | (2.409) | (2.727) |
| Under-ID $p$ value | 0.036 | 0.005 | 0.000 | 0.001 | 0.048 | 0.010 |
| $F$-stat | 1.736** | 4.918*** | 8.842*** | 6.183*** | 2.995*** | 3.888*** |
| *Panel B: civil liberties* | | | | | | |
| Progressivity | −1.940** | −10.161** | −7.975*** | −4.199*** | −13.809* | −22.849* |
| | (0.832) | (5.029) | (2.475) | (1.342) | (7.934) | (12.898) |
| Interaction | 0.568** | 3.222** | 2.783*** | 1.650*** | 5.123* | 7.398* |
| | (0.258) | (1.529) | (0.878) | (0.627) | (2.917) | (4.018) |
| Under-ID $p$ value | 0.013 | 0.042 | 0.001 | 0.001 | 0.084 | 0.075 |
| $F$-stat | 4.237*** | 3.294*** | 10.404*** | 7.227*** | 2.662*** | 2.231*** |

Robust standard errors clustered at the country level are in parentheses

* Significant at 10 %; ** significant at 5 %; *** significant at 1 %. The dependent variable is Gini index for net income. Each model uses the same variables as in Panel A of Table 2 along with the specified institutional variable, but only progressivity and its interaction with the institutional variable are reported. The measure of structural progressivity is indicated by column titles. IV is ARP-middle of contiguous neighbors and its interaction with the specified institutional variable in both panels. All models pass the Kleibergen and Paap Lagrange multiplier under-identification test (under-ID $p$ value). Political rights and civil liberties are measured on a reverse scale from 1 to 7, where 7 indicates no political rights or no civil liberties. $N = 1012$

ables to proxy for the extent to which pro-poor redistribution exists in each country: political rights and civil liberties. Both variables are from Freedom House and are rated on a scale from 1 to 7, with the highest value indicating *no* liberty or rights. Given hypothesis H2, we expect the coefficient on the interaction term $\delta_2$ to be positive. The coefficient estimates using political rights and civil liberty are reported in Table 4. We show results with ARP-middle of contiguous neighbors and its interaction with either political rights or civil liberty as IVs. First-stage diagnostics indicate the models are not under-identified.

The estimates indicate that progressivity has a greater equalizing effect in countries with greater access to political rights and civil liberties. These results show that equalizing the distribution of income via personal income taxes requires not only progressive tax structures, but also institutions that facilitate pro-poor redistribution policies on the expenditure side of the budget. The findings are consistent with Goni et al. (2011) who argue that the difference in inequality between Latin American and Western European countries is partly explained by the relatively more generous redistribution present in Europe's tax and transfer fiscal system.

### 4.4 The effect of progressivity on inequality in consumption (H3 and H4)

One of the main predictions in Sect. 2 is that changes in progressivity may affect true and observed income inequality differently and that the magnitude of the differential effect is increasing with the share of unreported income in the economy (H3 and H4).

This section of the paper tests hypotheses H3 and H4 by extending Eq. (6) to include interaction terms. First, to test H3, we write down the following specification:

$$I_{it} = \beta_0 + \delta_0 N_{it} + \delta_c P_{it} + \pi P_{it} * N_{it} + \beta_1 Z_{it} + \beta_2 W_{it} + \zeta_t + \epsilon_{it}, \tag{8}$$

where $N$ is a dummy variable (=1 if Gini is based on reported net income) and all other variables are as defined previously. The coefficient $\pi = (\delta_0 - \delta_*)$ tells us if the effect of progressivity on reported net income differs from that on consumption and is expected to be negative.[19] In addition to being negative, hypothesis H4 tells us that the magnitude of the difference increases with tax evasion. We therefore extend Eq. (8) to account for this possibility by estimating the following three-way interaction model:

$$\begin{aligned} I_{it} = \gamma_0 + \beta_1 Z_{it} + \beta_2 W_{it} + \gamma_1 E + \gamma_2 N_{it} + \gamma_3 N_{it} * E + \gamma_4 P_{it} \\ + \gamma_5 P_{it} * E + \gamma_6 P_{it} * N_{it} + \gamma_7 P_{it} * N_{it} * E + \zeta_t + \epsilon_{it}, \end{aligned} \tag{9}$$

where $E$ is an index that captures the prevalence of tax evasion (larger E indicates evasion is less prevalent) and all other variables are as defined previously. Following our discussion in Sect. 2, we use the corruption and law and order indices as proxies for the prevalence of tax evasion. Both variables are measured on a scale from 0 to 6, with 0 representing the worst case, and are available from the International Country Risk Guide (ICRG) for the period 1984–2005. We note that in Eqs (8) and (9) we also add gross savings rate to control for the potential effects of savings on the differences in inequality between actual net income and consumption.

The key parameters of interest are $\gamma_6$, which shows the differential effect of progressivity on net income versus consumption inequality in high-evading countries, and $\gamma_7$, which shows how this differential effect varies between high- and low-evading countries. Our a priori expectation is that $\gamma_6 < 0$ (H3) and $\gamma_7 > 0$ (H4).

Table 5 shows the IV estimates for three models: Eq. (6) for consumption-based Ginis; Eq. (8) for combined income and consumption Ginis; and Eq. (9) with the three-way interaction between progressivity, income base, and tax evasion proxied by the corruption index. The results are shown for MRP and ARP using average ARP-middle in bordering countries as IVs. The direct effect of progressivity on consumption inequality reported in column "Model 1" is negative, but not statistically different from zero. Adding income-based Ginis to consumption-based Ginis in "Model 2" makes the coefficient on ARP statistically significant. Additionally, we find evidence of an

[19] We note that if consumption expenditures used in Gini estimates are disproportionately underreported among the rich, then the difference between consumption inequality and income inequality based on survey data will be smaller than the actual difference in inequality. This implies that the differential effect posited in hypothesis 3 based on true expenditures will be even bigger in absolute value than the effect we estimate. Our estimate in this case can be interpreted as a lower-bound estimate.

**Table 5** Differential effect of progressivity on inequality in consumption and observed income

| | MRP | | | ARP | | |
|---|---|---|---|---|---|---|
| | Model 1 | Model 2 | Model 3 | Model 1 | Model 2 | Model 3 |
| $P$ | −0.987 | −2.450* | 2.411 | −2.243 | −2.927** | 0.717 |
| | (1.034) | (1.478) | (2.105) | 2.336 | 1.590 | 1.823 |
| $P$ × Corruption | | | −1.214** | | | −0.930* |
| | | | (0.559) | | | (0.360) |
| $P$ × Net | | −1.507** | −5.720* | | −0.545** | −4.646** |
| | | (0.735) | (3.422) | | (0.732) | (2.348) |
| $P$ × Net × Corruption | | | 0.889 | | | 0.991** |
| | | | (0.621) | | | (0.464) |
| Corruption | | | 5.012* | | | 2.944 |
| | | | (2.835) | | | (1.734) |
| Net income | | 11.875** | 24.433** | | 5.584 | 17.708** |
| | | (4.640) | (12.253) | | (3.848) | (6.898) |
| Net × Corruption | | | −2.218 | | | −3.271** |
| | | | (2.653) | | | (1.743) |
| Under-ID $p$ value | 0.005 | 0.005 | 0.012 | 0.072 | 0.002 | 0.008 |
| $F$-stat (first stage) | 10.250*** | 4.830*** | 3.333*** | 3.152*** | 8.340*** | 3.776*** |
| $F$-stat (second stage) | 4.241*** | 4.596*** | 4.081*** | 4.646*** | 10.496*** | 11.117*** |
| $N$ | 579 | 1531 | 1424 | 519 | 1531 | 1424 |

Robust standard errors clustered at the country level are in parentheses
* Significant at 10 %; ** significant at 5 %; *** significant at 1 %. $P$ is progressivity, Corruption is corruption index, and Net is a dummy for net income versus consumption. The dependent variable is Gini in net income or expenditures/consumption. Specifications in models 1, 2, and 3 are defined in Eqs. (6), (8), and (9), respectively, but only coefficients of interest are reported. IVs are ARP-middle in bordering countries and its respective interactions. Model 1 uses consumption-based Ginis only. The corruption index is measured on a scale from 0 to 6, with 0 representing the worst case

economically and statistically significant differential effect between consumption and net income inequality.

We show in Sect. 2 that the differential effect is likely dependent on evasion opportunities. If this is the case, then Eq. (9) is the more appropriate model. Results from this specification are presented in "Model 3" of Table 5 for the ICRG corruption index. For a typical country in the sample, we find that progressivity has a much greater influence on net income-based Ginis than on consumption-based Ginis ($\gamma_6 < 0$) (H3). The coefficient on $\gamma_7$ is positive and statistically significant for the ARP measure of progressivity, which suggests that this differential effect is more pronounced in economic environments conducive to tax evasion (H4). For the MRP measure, the coefficient on a three-way interaction term is positive but not statistically significant. In Table 6, we replicate the analysis for the ICRG corruption index and for the ICRG law and order index using two instruments: neighbors ARP-middle and the complexity of tax

**Table 6** Differential effect of progressivity on inequality in consumption and observed income

| Top rate | MRP | ARP | ARP | | |
|---|---|---|---|---|---|
| | | | Bottom | Middle | Top |
| *Corruption—two IVs* | | | | | |
| $P \times$ Net | −3.974** | −4.401** | −2.210 | −5.145** | −7.669* |
| | (1.908) | (2.028) | (1.393) | (2.500) | (4.011) |
| $P \times$ Net $\times$ Corruption | 0.618* | 0.855** | 0.458* | 0.936* | 1.037 |
| | (0.369) | (0.382) | (0.234) | (0.494) | (0.776) |
| Under-ID $p$ value | 0.064 | 0.010 | 0.240 | 0.004 | 0.150 |
| $F$-stat (first stage) | 2.914*** | 2.420** | 2.408** | 4.371*** | 2.982*** |
| Hansen J $p$ value | 0.532 | 0.871 | 0.980 | 0.424 | 0.301 |
| $F$-stat (second stage) | 5.882*** | 13.916*** | 15.278*** | 8.234*** | 4.895*** |
| *Law and order—two IVs* | | | | | |
| $P \times$ Net | −3.483* | −4.288* | −2.355 | −4.235 | −6.585 |
| | (1.918) | (2.336) | (1.753) | (2.598) | (4.092) |
| $P \times$ Net $\times$ Law | 0.698* | 0.924** | 0.519* | 0.903* | 1.210 |
| | (0.399) | (0.414) | (0.267) | (0.491) | (0.859) |
| Under-ID $p$ value | 0.020 | 0.004 | 0.131 | 0.007 | 0.040 |
| $F$-stat (first stage) | 3.504*** | 3.651*** | 2.770*** | 6.850*** | 3.801*** |
| Hansen J $p$ value | 0.633 | 0.573 | 0.835 | 0.513 | 0.621 |
| $F$-stat (second stage) | 6.288*** | 11.866*** | 14.419*** | 7.721*** | 5.432*** |

Robust standard errors clustered at the country level are in parentheses
* Significant at 10 %; ** significant at 5 %; *** significant at 1 %. $P$ is progressivity, Law is the law and order index, and Net is a dummy for net income-based versus consumption-based Ginis. The dependent variable is Gini in net income or expenditures/consumption. Each specification is similar to Model 3 in Table 5 and defined in Eq. (9), but only coefficients of interest are reported. IVs are ARP-middle in bordering countries and tax complexity and their respective interactions. The law and order and corruption indices are measured on a scale from 0 to 6, with 0 representing the worst case. $N = 1424$

systems.[20] Overall, the results draw a similar picture confirming hypotheses H3 and H4.[21]

## 5 Conclusions

We test four hypotheses about the relationship between tax progressivity and income inequality. First, increased structural progressivity of the PIT structure reduces observed income inequality (H1), especially in highly democratic environments (H2).

[20] Neighbors' ARP-middle appear to be a weaker predictor for tax progressivity in the first stage compared to our other models ($F$-stat $< 4$). We had to use additional instruments to make sure that IVs are valid before interpreting the results. The result in Table 5 is not affected by the use of two IVs; results are available upon request.

[21] The results reported in Table 6 are estimated without sample weights. The use of sample weights does not affect the results for corruption. However, the law and order results are somewhat sensitive to the inclusion of weights; sign and magnitude are similar, but estimates are imprecise in some specifications.

The third hypothesis is that structural progressivity has a differential effect on observed versus actual income inequality (H3), and that the difference between the two effects is positively related to the spread of tax evasion in the economy (H4).

These predictions are tested using time-varying measures of structural progressivity of national income tax systems. As predicted, we find that PIT progressivity reduces observed inequality in reported net income and show that this negative effect on observed income inequality is particularly strong in countries with more developed democratic institutions. At the same time, we find a significantly smaller negative effect of PIT progressivity on true inequality, approximated by consumption-based measures of the Gini coefficient. We also establish that the effect of tax progressivity on consumption inequality can be positive, especially in countries with weak law and order that increases the likelihood of tax evasion. Finally, we find evidence that changing progressivity at the top of the tax schedule is more effective in reducing observed net income inequality than similar changes at the bottom of the income scale.

Assuming that lower progressivity is associated with greater efficiency, our empirical analysis implies that the trade-off between equity and efficiency does in fact exist. This follows from the negative relationship that we identify between progressivity and reported income inequality. The result suggests that as taxes become more efficient, via lower progressivity, observed income inequality tends to increase. This result by itself points to the importance of taking into account the equity effects of shifts in tax policy toward greater efficiency.

What we find particularly interesting, however, is that the cost of efficiency differs across country groups. Because tax evasion is so pervasive in developing countries, our results lead us to speculate that developing countries face much lower equality penalty of efficiency. In other words, to the extent that efficiency is achieved by lowering the progressivity of taxes, developing countries with their higher levels of tax evasion experience a much smaller decline in equality for each increase in efficiency compared to developed countries. If flatter taxes can reduce the size of the underground economy, then developing countries may actually see an improvement in the distribution of income via the direct compliance response and via pro-poor redistribution of increased tax revenues from higher levels of compliance. Developed countries, on the other hand, may not benefit much from this evasion effect due to higher tax compliance rates to begin with. This may possibly explain why flat taxes are relatively more popular in developing countries than in developed countries.

**Acknowledgments** We would like to thank participants at the 71st Annual Congress of the International Institute of Public Finance, and two anonymous referees for their feedback and suggestions.

**Compliance with ethical standards**

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

Adam, A., Kammas, P., & Lapatinas, A. (2014). Income inequality and the tax structure: Evidence from developed and developing countries. *Journal of Comparative Economics, 43*(1), 138–154.

Allingham, M., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, *1*(3–4), 323–338.

Atkinson, A. B., & Brandolini, A. (2001). Promise and pitfalls in the use of 'secondary' data-sets: Income inequality in oecd countries as a case study. *Journal of Economic Literature*, *39*(3), 771–799.

Beck, T., Demirg-Kunt, A., & Levine, R. (2000). A new database on financial development and structure. *World Bank Economic Review*, *14*(3), 597–605.

Besley, T., & Case, A. (1995). Incumbent behavior: Vote-seeking, tax-setting, and yardstick competition. *The American Economic Review*, *85*(1), 25–45.

Borge, L.-E., & Rattso, J. (2004). Income distribution and tax structure: Empirical test of the Meltzer–Richard hypothesis. *European Economic Review*, *48*(4), 805–826.

Bulr, A. (2001). Income inequality: Does inflation matter? *IMF Staff Papers*, *48*(1), 139–159.

Chong, A., & Gradstein, M. (2007). Inequality and institutions. *The Review of Economics and Statistics*, *89*(3), 454–465.

Clarke, G. R. G., Xu, L. C., & Zou, H.-F. (2006). Finance and income inequality: What do the data tell us? *Southern Economic Journal*, *72*(3), 578–596.

Clotfelter, C. T. (1983). Tax evasion and tax rates: An analysis of individual returns. *The Review of Economics and Statistics*, *65*(3), 363–373.

Deaton, A. (1997). *The analysis of household surveys: A microeconometric approach to development policy*. Baltimore, MD: Johns Hopkins University Press.

Dhami, S., & Al-Nowaihi, A. (2007). Why do people pay taxes? Prospect theory versus expected utility theory. *Journal of Economic Behavior and Organization*, *64*(1), 171–192.

Dobson, S., & Ramlogan-Dobson, C. (2012). Why is corruption less harmful to income inequality in Latin America? *World Development*, *40*(8), 1534–1545.

Doerrenberg, P., & Duncan, D. (2013). Distributional implications of tax evasion: Evidence from the lab. *Public Finance Review, 42*(6), 720–744.

Duncan, D. (2014). Behavioral responses and the distributional effects of the russian 'flat' tax. *Journal of Policy Modeling*, *36*(2), 226–240.

Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics*, *84*(2), 755–776.

Egger, P., & Radulescu, D. M. (2009). The influence of labour taxes on the migration of skilled workers. *The World Economy*, *32*(9), 1365–1379.

Engstrom, P., & Holmlund, B. (2009). Tax evasion and self-employment in a high-tax country: Evidence from Sweden. *Applied Economics*, *41*(19), 2419–2430.

Feldstein, M. (1995). The effect of marginal tax rates on taxable income: A panel study of the 1986 tax reform act. *Journal of Political Economy*, *103*(3), 551–572.

Ferede, E., & Dahlby, B. (2012). The impact of tax cuts on economic growth: evidence from the Canadian provinces. *National Tax Journal*, *65*(3), 563–594.

Goni, E., Lopez, J. H., & Serven, L. (2011). Fiscal redistribution and income inequality in Latin America. *World Development*, *39*(9), 1558–1569.

Gordon, R., & Li, W. (2007). *Puzzling tax structures in devloping countries: A comparison of two alternative explanations, chapter 1* (Vol. 16). Chicago, IL: The University of Chicago Press.

Gordon, R., & Li, W. (2009). Tax structures in developing countries: Many puzzles and a possible explanation. *Journal of Public Economics*, *93*(7–8), 855–866.

Gorodnichenko, Y., Martinez-Vazquez, J., & Sabirianova Peter, K. (2009). Myth and reality of flat tax reform: Micro estimates of tax evasion response and welfare effects in russia. *Journal of Political Economy*, *117*(3), 504–554.

Gorodnichenko, Y., Sabirianova-Peter, K., & Stolyarov, D. (2009). Inequality and volatility moderation in russia: Evidence from micro-level panel data on consumption and income. *Review of Economic Dynamics*, *13*(1), 209–237.

Gradstein, M., & Milanovic, B. (2004). Does libert = egalit? A survey of the empirical links between democracy and inequality with some evidence on the transition economies. *Journal of Economic Surveys*, *18*(4), 515–537.

Gravelle, J. G. (1992). Equity effects of the tax reform act of 1986. *Journal of Economic Perspectives*, *6*(1), 27–44.

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*(2), 467–475.

Kakwani, N. C. (1977). Measurement of tax progressivity: An international comparison. *Economic Journal*, *87*(345), 71–80.
Kim, D.-H., & Lin, S.-C. (2011). Nonlinearity in the financial development-income inequality nexus. *Journal of Comparative Economics*, *39*(3), 310–325.
Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, *133*(1), 97–126.
Kleven, H., Landais, C., & Saez, E. (2013). Taxation and international migration of superstars: Evidence from the European football market. *American Economic Review, 103*(5), 1892–1924.
Kopczuk, W. (2005). Tax bases, tax rates and the elasticity of reported income. *Journal of Public Economics*, *89*(11–12), 2093–2119.
Lakner, C., & Milanovic, B. (2015). Global income distribution: From the fall of the berlin wall to the great recession. *The World Bank Economic Review* **(forthcoming)**.
Law, S. H., & Tan, H. B. (2009). The role of financial development on income inequality in malaysia. *Journal of Economic Development*, *34*(2), 153–168.
Lee, Y., & Gordon, R. H. (2005). Tax structure and economic growth. *Journal of Public Economics*, *89*(5–6), 1027–1043.
Lemieux, T., Fortin, B., & Frechette, P. (1994). The effect of taxes on labor supply in the underground economy. *American Economic Review*, *84*(1), 231–254.
Martinez-Vazquez, J. (2008). *The impact of budgets on the poor: Tax and expenditure benefit incidence analysis, chapter 5* (pp. 113–162). Washington, DC: World Bank.
Meltzer, A. H., & Richard, S. F. (1981). A rational theory of the size of government. *Journal of Political Economy*, *89*(5), 914–27.
Milanovic, B. (1999). Explaining the increase in inequality during transition. *Economics of Transition*, *7*(2), 299–341.
Minarik, J. J. (1979). The size distribution of income during inflation. *Review of Income and Wealth*, *25*(4), 377–392.
Mookerjee, R., & Kalipioni, P. (2010). Availability of financial services and income inequality: The evidence from many countries. *Emerging Markets Review*, *11*(4), 404–408.
Musgrave, R., & Thin, T. (1948). Income tax progression, 1929–48. *Journal of Political Economy*, *56*(6), 498–514.
OECD. (2011). *International migration outlook: SOPEMI 2011*. Paris: OECD Publishing.
Persson, T., & Tabellini, G. (2002). *Political Economics and Public Finance* (Vol. 3). Amsterdam: Elsevier.
Sabirianova Peter, K., Buttrick, S., & Duncan, D. (2010). Global reform of personal income taxation, 1981–2005: Evidence from 189 countries. *National Tax Journal*, *63*(3), 447–478.
Slemrod, J. (1985). An empirical test for tax evasion. *The Review of Economics and Statistics*, *67*(2), 232–238.
Slemrod, J. (1994). *On the high-income Laffer curve, chapter 6* (pp. 177–210). Cambridge: Cambridge University Press.
Slemrod, J. (1995). Income creation or income shifting? behavioral responses to the tax reform act of 1986. *The American Economic Review*, *85*(2), 175–180.
Slemrod, J., & Kopczuk, W. (2002). The optimal elasticity of taxable income. *Journal of Public Economics*, *84*(1), 91–112.
Suits, D. B. (1977). Measurement of tax progressivity. *American Economic Review*, *67*(4), 747–752.
Violante, G. (2010). Special issue: Cross-sectional facts for macroeconomists. *Review of Economic Dynamics*, *13*(1), 1–264.
Whalley, J. (1990). *Foreign responses to U.S. tax reform*. Cambridge: MIT Press.
Yitzhaki, S. (1974). Incometax evasion: A theoretical analysis. *Journal of Public Economics*, *3*(2), 201–202.