

Giertz, Seth H.; Ramezani, Rasoul

Working Paper

Taxes, Wage Capitalization and the Ability of States to Redistribute Income

GLO Discussion Paper, No. 291

Provided in Cooperation with:
Global Labor Organization (GLO)

Suggested Citation: Giertz, Seth H.; Ramezani, Rasoul (2018) : Taxes, Wage Capitalization and the Ability of States to Redistribute Income, GLO Discussion Paper, No. 291, Global Labor Organization (GLO), Maastricht

This Version is available at:
<https://hdl.handle.net/10419/190970>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

Taxes, Wage Capitalization and the Ability of States to Redistribute Income

By Seth H. Giertz and Rasoul Ramezani*

Abstract

Local and state governments attempt to lessen after-tax income inequality via progressive taxation. Migration responses of capital and labor undermine such attempts. Location theory predicts that cross-state migration will continue until the redistributive effects from taxation are fully capitalized into gross wages leaving after-tax wages unchanged. Empirical evidence has not reached a consensus on this issue. At one extreme, Feldstein and Wrobel (1998) report evidence of full tax capitalization for US states. At the other extreme, Leigh (2008) reports very little to no wage capitalization. We revisit this question by creating a pseudo panel from CPS data spanning years 1997 to 2015. Our “best” estimate is that pre-tax wages adjust in response to redistributive state and local taxes, negating roughly 50 percent of effect compared to counterfactual with no behavioral responses.

Keywords: Fiscal federalism; Redistribution; State taxation; Tax capitalization; Progressivity; Migration

JEL classification: H20, H71, H77

* Giertz: University of Texas at Dallas, EPPS, 800 W Campbell Rd, GR31, Richardson, TX 75080-3021 (e-mail: seth.giertz@utdallas.edu). Ramezani: University of Texas at Dallas, EPPS, 800 W Campbell Rd, GR31, Richardson, TX 75080-3021 (e-mail: rasoul.ramezani@utdallas.edu).

** We would like to thank Jeffrey Thompson for feedback regarding our data construction. We would also like to thank Mehmet Tosun and Indiana University SPEA seminar participants for helpful comments on earlier versions of this paper.

1) Introduction

For several decades, following Tiebout (1956), Musgrave (1971) and Oates (1972), a consensus emerged among economists that the ability of state and local governments to redistribute income was quite limited. Musgrave (1971) succinctly summarizes this view: “Policies to adjust the distribution of income among individuals must be conducted on a nationwide basis ... [R]egional measures are self-defeating, as the rich will leave and the poor will move to the more egalitarian-minded jurisdictions.” A sub-national policy aimed at redistribution is not just ineffectual, but, as Musgrave notes, “leads to severe locational inefficiencies.”

While this view was generally accepted by the economics profession, arguments were made for some redistribution at the state or local level. For example, Pauly (1973) argues that redistribution is partly a local public good whose value varies across jurisdictions. Thus, it follows that sub-national governments should weigh the local benefits from redistribution against the costs. If wage rates do not fully capitalize, areas placing a high value on redistribution may want to pursue such policies at the margin – although to a lesser degree than if factors of production were not very mobile.¹

Reflecting on developments in fiscal federalism, Oates (1999) maintains that states have a limited ability to redistribute income, reiterating a view he articulated in his seminal 1972 book, *Fiscal Federalism*. However, by 1999, the theory had only undergone limited rigorous empirical testing. The most direct test of this hypothesis, Feldstein and Wrobel (FW 1998), had

¹ More recently, a similar view is explored by Gordon and Cullen (2012), who begin by assuming that the state and federal governments are pursuing redistributive policies with the objective of maximizing social welfare. Given a choice between state-level policies or a federal policy that deviates from the state’s preferred choice, they conclude that some state redistribution is likely optimal even with high migration elasticities.

just been published and strongly supported the view of full capitalization – i.e., an inability to redistribute at the sub-national level.

Over the past two decades, a number of papers have examined the relationship between state and local taxes and either inequality or migration – where migration is presumed the key factor driving wage capitalization. Studies using individual-level data, such as Leigh (2008), Young and Varner (2011), and Young et al. (2016), have found very modest migration responses to state tax changes, implying that states have considerable leeway in redistributing incomes.²

Those arguing that taxes are an important factor behind interstate migration often point to the persistently more rapid employment growth in low-tax states (with less progressive tax structures). Also consistent with the view that taxes drive migration is Haughwout et al. (2004). Although not explicitly examining migration, they do find the tax base of several large US cities to be quite responsive to taxes. They conclude that three of the four cities that they examine are close to maximizing local tax revenues (Houston, New York City, and Philadelphia).³ They further find that city-level wage and income taxes have considerable negative effects on employment.

In other work, Kleven et al. (2013 and 2014), and Moretti and Wilson (2017) find substantial migration responses among star athletes and scientists with respect to locational

² In a replication of Young and Varner (2011), Cohen, Lai, and Steindel (2015) contend that migration responses were in fact larger, at least for very high-income taxpayers targeted by the 2004 New Jersey tax increase.

³ The revenue-maximizing tax rate is a function of both the effective tax rate and the elasticity of the tax base (Saez, et al., 2012).

variation in tax rates. However, these results may not be generalizable to upper-income groups as a whole.⁴

Changes to redistributive tax policies not only lead high-income groups to reassess their location decisions. Lower-income groups have an incentive to respond as well. This is implied by Tiebout (1956) and is behind “race to the bottom” fears associated with leaving welfare relief to the states. However, estimated migration responses to tax or welfare policies targeting low-income groups are not always strong. In reviewing this literature, Brueckner (2000) reports modest migration responses. Although, he does find that, when making policy, states behave as though low-income groups are quite responsive to taxes and benefits.

While the empirical evidence is mixed, the preponderance of recent studies suggests that interstate migration responses to taxes are modest, especially for those outside of the very top of the income distribution. However, it may be that responses are large but attenuated over the relatively short time horizons of most studies. Because of the fixed costs of moving, or frictions more generally (see Chetty, 2012), long-run or so-called “structural” elasticities may be larger. In many instances, it may take many years, or even decades, to reach a new equilibrium. Likewise, the literature on frictions suggests that short-term responses should be larger and adjustment quicker when frictions are lower – for example, when moving costs are lower and policy differences are more salient. This view is consistent with following observation from Glaeser (2011):

⁴ Another strand of the tax-migration literature focuses on taxes and location decisions of retirees. States often craft tax policy under the presumption that retirees are very sensitive to taxes. However, panel-data evidence suggests that location decisions for retirees are not responsive to state taxes (Conway and Rork, 2006 and 2012) or are only modestly responsive (Bakija and Slemrod, 2004).

East St. Louis provides an extreme example of the urban poverty paradox, whereby public policy that helps the poor in one area can lead to a massive concentration of poverty. East St. Louis lies across the Mississippi River, in Illinois, from St. Louis, Missouri. *In 1989, the annual Aid to Families with Dependent Children payment was 20 percent higher in Illinois than in Missouri.* If you were out of work, it made sense to move to Illinois, and so in 1990 the poverty rate in East St. Louis was 43 percent...Since welfare reform in 1996, the gap in welfare payments has essentially disappeared, and the poverty rate gap between St. Louis and East St. Louis has narrowed considerably.

-- Edward Glaeser (2011-01-18T23:00:00+00:00). Triumph of the City (Kindle Locations 1561-1566). Penguin Group USA, Inc. Kindle Edition. Emphasis added.

In this paper, we reexamine the ability of US states to have a persistent effect on the distribution of income. Following FW, we focus on wage capitalization. We measure the relationship between pre-tax money wages and average individual tax burdens. By including state (or cohort)⁵ fixed effects, we control for differences in cost-of-living across states. Fixed effects is a “within” estimator. Thus, this approach measures the effects of tax redistribution from within-state variation in average tax rates. We begin with FW’s general framework for individual cross sections, extending the analysis to include many more years. Data are from the Current Population Survey (CPS) for years 1997-2015, in conjunction with information from other sources, such as NBER’s TAXSIM and the Consumer Expenditure Survey (CEX). From these data, a pseudo panel is created in order to control for unobservables that may result in heterogeneous wage growth across groups, a phenomenon that could otherwise result in spurious estimated tax responses.

⁵ Cohort fixed effects are used with pseudo-panel data.

As with FW, our key coefficient measures the percent change in the gross wage associated with a one-percent increase in the average net-of-tax (or after-tax) rate. The average after-tax rate equals one minus the average tax rate – in other words, the ratio of after-tax earnings to gross earnings. An estimate of 0 suggests that taxes have no effect on gross wages – thus behavioral responses do not counteract attempts to alter the distribution of income. An estimate of -1 suggests that tax changes are fully offset by changes to the gross wage – thus behavioral responses counteract attempts to alter the distribution of income. Examining annual cross-sections, we find a range of about 0 to -0.47 for men and -0.12 to -0.34 for women. Our preferred estimates from the pseudo-panel data, for prime-age workers, are -0.67 for men and -0.49 for women. Adding controls for differing secular wage trends result in gross-wage elasticities with respect to the average net-of-tax rate of -0.48 for men and -0.43 for women. In sum, we find substantial, albeit partial, wage capitalization, consistent with tax-induced migration responses. However, in contrast to FW, these results suggest that money wages far from fully adjust to offset changes to state-level attempts at redistribution. While this is consistent with the notion that states have some margin for redistributing income, it is not definitive, since it is possible that other prices, such as for housing, also respond to state policies towards redistribution.⁶

Given the substantial frictions associated with moving, focusing on capitalization offers insights that cannot be gleaned from short-term migration responses. With respect to location decisions, the costs to reoptimizing are substantial and vary over time. Thus, the benefits from a more favorable tax situation may not outweigh the costs of relocating. However, a series of tax

⁶ For example, in examining tax variation across Swiss cantons, Kirchgassner and Pommerehne (1996) report capitalization effects for both wages and rents. However, these effects were only strong enough to partially negate canton-level tax redistribution.

changes, over time, may push taxpayers over the threshold, where moving makes economic sense. Family considerations also attenuate short-run migration responses. Thus, families may not move shortly after a tax change. But, a job offer down the road, which one would not have previously considered, may now be enticing under the new tax regime. Likewise, some may decide to relocate, but hold off until doing so is less burdensome, such as after children have completed high school. While these factors seem likely attenuate short-run responses, Chetty (2012) shows that frictions more generally could either result in over- or under-estimates of the true structural response. For example, a series of tax changes may result in small responses. Whereas, one additional modest tax change may push taxpayers beyond the tipping point, where the gains from more favorable taxation outweigh the costs of relocating, resulting in a large response.

The rest of the paper is organized as follows. Section 2 presents a simplified model of tax capitalization. Section 3 presents the data, including background on tax imputations, variables, and pseudo-panel construction. Section 4 focuses on the econometric approach. Section 4 presents our empirical findings and Section 6 concludes.

2) Regional Economies and Locational Indifference

A central tenet of regional economics is that the free mobility of factors of production yields an equilibrium characterized by locational indifference. That is, individuals migrate towards areas that afford them a higher quality of life, just as capital flows to areas where its (risk-adjusted) return is highest. Equilibrium is achieved when there is no incentive to move – i.e., where individual utility or returns to capital are maximized. Of course, those whose preferences deviate substantially from the norm may have a clear locational preference – such as those with unique family ties etc. But, the marginal individual should be locationally indifferent.

As FW emphasize, full capitalization does not necessarily imply that changes to *money wages* fully offset redistributive state and local policies. Rather, full capitalization implies that *real wages* adjust to counteract such policies. The real wage is the money wage adjusted for cost of living, where cost of living may vary not only by locality but also by income type. For example, the effects of a tax change could be capitalized into land or house prices, where high- and low-income housing are not substitutable. FW express this equilibrating condition such that

$$(1 - t_H)W_H/P_H = \omega_H, \text{ and} \quad (1)$$

$$(1 - t_L)W_L/P_L = \omega_L. \quad (2)$$

W is the money wage, t is the average tax rate, and P is the price level, where all of these variables are denoted separately for high- (H) and low- (L) skilled individuals. Adjusting money wages for taxes and price levels yields real net wages (ω_H and ω_L), which are the same across states. For locational indifference to hold, equations (1) and (2) should also account for locational amenities, in addition to taxes. In a regression context, state (or cohort) fixed effects make explicitly accounting for amenities unnecessary, so long as the amenities are valued similarly by the different income groups.

Locational indifference is an axiom of the Rosen-Roback model used to measure quality of life, and the value of amenities, across cities. The model, based on Rosen (1979) and Roback (1982), assumes that migration drives local wages and cost of living (through house or land values) such that, in equilibrium, these prices exactly offset the value placed on amenities. For example, high wages and low house values will predominate in areas with few amenities. And, in-migration will drive down wages and drive up house prices in high-amenity areas. Albouy (2016) uses this model to estimate the value placed on amenities for a large number of US MSAs. In earlier work, Albouy (2009) finds that federal taxes have a differential impact across

locations, even after controlling for individual characteristics. Albouy (2016) argues that federal taxes contribute to locational inefficiencies and that incorporating federal taxes into his model improves quality-of-life measures. Note, the Rosen-Roback model is undermined if individuals are unresponsive to taxes, or to other differences in desirability across locations.

Figure 1 presents average gross wages by state, broken down by gender and by education. We focus on two broad education groups: Those with a high-school degree or less and those with at least a four-year college degree. In calculating these numbers, we assume that state age distributions conform to the national age distribution. Furthermore, we assume that the educational distribution within each (of the two) broader education groups is the same across states.⁷ Thus, differences in wages across states are not due to compositional differences with respect to age or educational attainment. Our aim is to highlight differences in wages due to location-specific factors – however, we cannot rule out differences in unobservables. For example, human capital levels may vary across states, even after controlling for years of schooling (or degree attainment).

Figure 1 shows that, for a given education group, the gross wage does in fact vary considerably across states. Furthermore, the wage gap – defined as the wage for the higher education group minus the wage for the lower education group – also varies across states. Gross wages for males with at least a college degree range from more than \$40/hour in Connecticut and New Jersey to less than \$25/hour in Montana. For males with a high school degree or less, the range is \$21/hour (Alaska) to \$15/hour (Arkansas and Washington, DC). For females, wage rates are monotonically lower, but the pattern is the same. Casual empiricism suggests a strong

⁷ For example, in making calculation for college graduates, we hold the breakdown between those whose highest degree is a bachelor's and those with graduate degrees constant.

positive relationship between cost of living and wages. In general, states with more progressive tax systems tend to have higher wage premiums, consistent with the notion that capitalization effects are offsetting attempts at redistribution.

Wage patterns across states may reflect the capitalization of amenities. This likely reflects factors such as cost-of-living, weather, or aesthetics. Positive amenities are valued across the wage distribution – although the intensity of preferences may vary with income. By contrast, the within-state wage gap between education groups is more likely to reflect capitalization in response to redistributive policies, where such policies represent an amenity to one group and a disamenity to another.

The states in Figure 1 are ordered from highest to lowest male wage gap. The magnitudes of the wage gap can be inferred from comparing the wage levels within each state. For men, this gap ranges from \$8/hour in Montana to \$24/hour in New Jersey. For women, the gap ranges from \$8 in Wyoming to \$17 in Washington, DC.

In sum, the patterns presented in Figure 1 are consistent with the notion that location-specific amenities are, at least partially, capitalized into pre-tax wage rates. Furthermore, the pattern is consistent with wage capitalization counteracting state-level attempts at redistribution. However, this analysis is only suggestive. Many factors underlie the variation observed in Figure 1 and some relationships may be spurious. For example, states with greater wage inequality may respond with more progressive tax schedules and, within our broad education categories, there is surely great variation in worker productivity.

[Figure 1 about here]

3) Data

Our core data are from the Current Population Survey (CPS) and span years 1997-2015. We first construct individual cross sections for these years. Later, we stack these data, forming a repeated cross-sectional dataset. Finally, we use these data to form a pseudo panel. The CPS includes detailed information on income and hours, as well as demographic variables. We also use the Consumer Expenditure Survey (CEX) to impute household consumption in order to estimate sales tax liabilities. The CEX data are further used to impute property tax liabilities both for homeowners and implicit liabilities borne by renters.⁸ Finally, we use NBER's TAXSIM model (Feenberg and Coutts, 1993) to impute federal and state income tax rates and liabilities.

3-1) Current Population Survey

The CPS is a household survey produced by the US Census Bureau in conjunction with the US Bureau of Labor Statistics. It is a nationally representative sample that currently includes around 54,000 households – or, approximately 106,000 individuals ages 15 and older.⁹ Administrative tax data are often considered the gold standard for research on taxation. The Statistics of Income's confidential tax panel, for example, comprises of roughly 250,000 tax returns per year and heavily oversamples very high-income taxpayers. While IRS tax data offer advantages, they also have some drawbacks. For example, the CPS includes variables not available in the tax data, such as detailed information on hours worked, fringe benefits, some nontaxable income sources, and demographic information.

Recent research has shown the CPS to be reliable for tax studies. For example, Burkhauser, et al. (2012) show that when income and tax units are defined in the same way, income shares from CPS very closely match those estimated by Piketty and Saez (2003) using

⁸ The CPS also includes property tax liabilities but only for a subset of households (and not for renters).

⁹ See Current Population Survey, 2015 Annual Social and Economic (ASEC) Supplement.

IRS tax data. Tax data are still superior in some respects because of their much larger sample size, which includes nearly all filers at the very top of the income distribution. However, with the CPS, Burkhauser, Larrimore, and Simon (2012) are able to examine the impact to changing family structures on income inequality. By contrast, with tax data, it is difficult to focus on anything other than the tax unit. Another important advantage of the CPS is that the data are publicly available. SOI's tax panel is only accessible to handful of federal employees and to very few external researchers who apply to the Statistics of Income's Joint Statistical Research Program.¹⁰

While tax data are ideal for estimating relationships between reported income and tax rates (Saez, Slemrod, and Giertz, 2012), they may not be ideal for other measures. Because tax returns are official documents, they are likely prepared with care. However, taxpayers generally have a strong incentive to underreport income and overstate deductions. Responses to CPS interviews, on the other hand, have no bearing on one's tax liability. Thus, while there may be psychic reasons to misrepresent one's income, these incentives are likely smaller than on tax returns. Thus, for our study, which examines the effect of taxes on pre-tax wages, CPS wages are arguably less biased than IRS data. Additionally, tax returns do not include hours information, and thus lack information needed to calculate average wage rates.

The version of the CPS that we use is compiled by the Center for Economic and Policy Research (CEPR). CEPR cleans the variables, including making adjustments for definitional changes over time, and links the March CPS with information included in other months of the

¹⁰ Some versions of IRS data are publicly available. This includes the SOI's Public Use File, which includes annual cross sections of tax data. Some information is redacted or blurred to protect the identity of filers. As a result, these data do not include identifiers needed to link these filers across years.

CPS. Each household in the CPS is interviewed in four consecutive months. Next, the household is not interviewed for eight consecutive months. Finally, the household is interviewed for four more consecutive months. I.e., each household is interviewed for four months in one year and then in the same four months the following year. For our purposes, the two most important months of the CPS are March and the last month of each four-month stint in the sample. Those in the last of four consecutive months in the sample are called an Outgoing Rotation Group (ORG).

The March CPS is especially important because it includes data on total income and its components (i.e., wages and salary, self-employed income, and unearned income – including dividends, unemployment benefits, Social Security benefits, pension benefits, etc.). It also includes work experience, including employment status, number of weeks worked, hours worked per week, occupation, and industry. Note that data on employment and income refer to the preceding year while the demographic data refer to the year of the survey. The ORG files are important because they include detailed labor force information, some of which is not available in the March files.

As mentioned earlier, we use micro level data from the CEX to impute the taxable consumption (and so state sales tax) and local property tax liabilities for the CPS sample.

3-2) Variables

Following FW, our baseline sample includes all full-time workers who reported positive earnings (including wages and salary plus self-employment income). A full-time worker is defined as one who works more than 35 hours per week. Moreover, the baseline sample is restricted to workers between 24 and 64 years old.

In order to estimate the extent to which state and local taxes are capitalized into wages, we regress the average hourly gross wage, which equals total earnings divided by total hours

worked, on net-of-tax rate (*NATR*). We also control for individual demographic characteristics by including dummies for education level, race, metropolitan area, and filing status. Other controls include work experience and number of children (broken down by age). Education dummies are based on the highest level of attainment (including less than high school, high school, some college, four-year college, and advanced degrees). Experience is approximated by age minus years of education minus six. Race is represented by a vector of dummy variables for: white, black, Hispanic, and other races.

Our key explanatory variable is net-average tax rate (*NATR*), which equals one minus the average tax rate (*ATR*).¹¹ *ATR* includes five components:

$$ATR = FATR + SATR + FIATR + PATR + SLATR.$$

Where, *FATR*, *SATR*, and *FIATR* are federal, state, and FICA payroll tax share of total income, respectively. *PATR* and *SLATR* are property and sales taxes, respectively, as a share of total income.

The first three components of the *ATR* are imputed into the CPS data using NBER's TAXSIM model. TAXSIM applies the federal and state income tax laws to the taxpayer's income (and other characteristics) in order to calculate the payroll tax liabilities. The March CPS includes all information that are required for TAXSIM, such as filing status, state of residence, number of dependents, annual earned income, annual unearned income (including dividends income, Social Security benefits, unemployment compensation benefits etc.).

¹¹ The average tax rate is used instead of the marginal tax rate because, following FW and Thompson (2011), individuals are assumed to maximize their post-tax wage (controlling for amenities and cost of living) because it is generally prohibitively costly to shift an hour of earnings across states.

State property and sales tax payments are imputed using the micro-level CEX data. Specifically, we run repeated cross-sectional OLS regressions both for local property taxes paid and for taxable consumption. Each of these variables are regressed against income and exogenous characteristics (such as age, number of children, education, and a metropolitan dummy). We then merge the predicted values from these regressions into our CPS-based dataset. Data on state sales-tax rates are then used to impute sales tax liabilities.¹² This exercise is done both for homeowners and renters, assuming that renters bear property taxes implicitly through rent payments.¹³

3-3) Pseudo-Panel Data Construction

We construct the pseudo-panel dataset from the repeated cross-section (RCS) data by first grouping observations within each cross section into groups with shared exogenous and time-invariant characteristics. Specifically, we grouped observations, by year, into bins based on four traits: gender, education level, birth cohort (based on five-year intervals), and the state of residence. Within each year, group assignments are mutually exclusive.

Note the trade-off between the number of observations per cohort and the number of cohorts. Previous studies (such as Deaton (1985), Moffitt (1993), Verbeek and Nijman (1992), and Verbeek and Vella (2005)) examine the asymptotic properties of pseudo-panel data constructed from large cross-sectional samples (such as March CPS). They find that the fixed effect regression model yields consistent estimates if either the number of cohorts are fixed as cohort size approaches infinity or as the number of cohorts approaches infinity while the cohort size is fixed. Moreover, the cohort fixed effects, which equal the average of cohort-member

¹² Don Bruce graciously provided us pre-2000 data on state sales-tax rates.

¹³ For detailed discussion regarding tax imputations, see an appendix is available from the authors.

individual fixed effects, become time-invariant (Deaton, 1985). Deaton (1985) further reports that if cohorts are very different in size, each observation to be weighted by the square root of the cohort size.

For each cohort-year, we calculate averages for all of the relevant variables. For $\ln(NATR)$, we calculate the cohort's weighted mean, using total family income as the weight. This is because low- and high-income individuals may be grouped together. Income-weighting provides a more meaningful average tax rate that is independent of the distribution of taxes within a cohort. In addition, Browning, Deaton, and Irish (1985) note that, for variables in log (or square) form, one should calculate the average of the logged (or squared) values, not the log (or square) of the averages. For indicator variables, we calculate proportions. Summary statistics of pseudo-panel data are presented in Table 1 for males and females, separately. The pseudo-panel dataset includes 2,510 male cohorts (39,619 observations) and 2,458 female cohorts (38,301 observations). The average cohort size is 16 for men and 13 for women. The average hourly wage is \$24.87 for men and \$17.93 for women. The average tax rate – i.e., state, local, and federal taxes as a share of total income – is 30.35 percent for men and 28.85 percent for women. The number of children under 18 averages 0.9 and is almost the same for men and women.

[Table 1 about here]

4) Methodology

4-1) The Feldstein and Wrobel (FW) Model

FW used two cross sections of March CPS data (1983 and 1989) to separately test whether the local and state redistributive taxes lead to a more equal distribution of post-tax income. To do this, they regressed the log of the hourly gross wage ($\ln(wage_i)$) on the log of the average net-

of-tax rate – represented by $\ln(1 - ATR_i)$ or $\ln(NATR_{is})$ – a vector of individual characteristics, X_i , and state fixed effects, φ_s , such that

$$\ln(wage_{is}) = \alpha + \beta \ln(1 - ATR_{is}) + X_i' \gamma + \varphi_s + \varepsilon_{is}. \quad (3)$$

The ATR is endogenous because, in a progressive tax system, the average tax rate increases with the earnings (of which the wage is a key component). I.e., an increase in the wage, all else equal, causes the ATR to rise. Thus, an OLS regression should yield upward-biased estimates of wage capitalization – i.e., a more negative estimate of β . To address this, FW employed 2SLS, using a predicted average tax rate, based on predicted income, as an instrument for ATR . We return to this endogeneity issue in section 4-4.

4-2) Adapting FW for Repeated Cross-Sections (RCS)

RCS analysis allows for many more observations in the regression, increasing the precision of the model. Furthermore, pooling many years of data will tend to average out unobserved random shocks hitting subsets of the sample in particular years. In order to implement the RCS approach, we first stack all cross-sections to form a pooled dataset. Next, we estimate a single regression such that

$$\ln(wage_{ist}) = \alpha + \beta \ln(NATR_{ist}) + X_{it}' \gamma + \varphi_s + \lambda_t + \varepsilon_{ist}, \quad (4)$$

where, λ_t represents year fixed effects.¹⁴ Year fixed effects control for national shocks, including the impact of federal tax changes, so long as the effects are symmetric across states. State fixed effects, φ_s , control for time-invariant factors that may be correlated with gross wages and explanatory variables.

¹⁴ Because we are not following the same individuals over time, we cannot include individual fixed effects in this equation.

4-3) Adapting FW for Pseudo-Panel Data

With cross-sectional analysis we cannot control for unobservable fixed effects, which may be correlated with earnings. For example, skill is only partially captured by education and experience. Thus, not subtracting individual fixed effects from the error term may bias estimates (Moffitt, 1993).

Deaton (1985) developed the pseudo-panel approach, allowing researchers to control for fixed effects with only repeated cross sections.¹⁵ Under this approach, in a given year, data are sorted into bins where observations within a bin have common exogenous characteristics, which are constant over time. Groups should be constructed over characteristics that lead to maximum inter-group and minimum intra-group variation. Groupings over common characteristics (e.g., birth cohort, gender, race etc.) are now the unit of observation, which can be followed over time. Moreover, for each bin, the mean values of the variables represent the value for a random sample drawn from a population with identical characteristics. Although members of a given group change over time, panel techniques applied to the cohort-based data are sound, so long as each cross section is randomly selected.

Equation (5) represents our baseline pseudo-panel regression,

$$\overline{\ln(wage_{ct})} = \alpha + \beta \overline{\ln(NATR_{ct})} + \bar{X}'_{ct}\gamma + \bar{\mu}_c + \lambda_t + \bar{\epsilon}_{ct}. \quad (5)$$

The dependent variable, $\overline{\ln(wage_{ct})}$, is the mean log wage for cohort c at year t .¹⁶ Explanatory variables include cohort means of the log net-of-tax rate at year t , $\overline{\ln(NATR_{ct})}$, a vector of demographics (i.e., cohort means), \bar{X}_{ct} , year dummies, λ_t , and cohort fixed effects, $\bar{\mu}_c$.

¹⁵ For an application of this approach related to tax policy, see Blundell, Duncan, and Meghir (1998).

¹⁶ As noted previously, the accepted practice is to take the average of the logged values, as opposed to the log of the averages (Browning, Deaton, and Irish, 1985).

In principle, cohort fixed effects may vary over time because the composition of individual characteristics may change from year to year. However, as cohort size increases, year-to-year variation in the composition of the cohort diminishes, implying that cohort fixed effects become time invariant (Verbeek, 2008). Thus, fixed effects under a traditional panel is based on the assumption that unobserved individual factors do not vary over time. Here, with a pseudo panel, it is further assumed that the composition of the cohort is not varying over time (Deaton, 1985). Pseudo-panel data also has advantages over traditional panel data. In particular, traditional panels, in most cases, suffer from attrition bias or nonrandom sampling. Pseudo-panel analysis can sidestep this problem because it is based on independent random sampling from each cross section.

Pseudo-panel data also allows for controls for secular divergence in the wage distribution. Piketty and Saez (2003) and Piketty, Saez, and Zucman (2018) document long-run divergence in the U.S. distribution of income. In many instances, labor income plays a major role in this divergence. One possible factor is skill-biased technological change, where those with higher wages, who are also more likely to have higher average tax rates, have faster wage increases than low-wage groups (Card and DiNardo, 2002). This issue has hampered panel-data analysis in the elasticity of taxable income (ETI) literature (Saez, Slemrod, and Giertz, 2012). As labor income is a key component in our wage-rate calculations, divergence in the income distribution could also hamper our analysis.

In light of this, we add controls for income divergence similar to those used in the ETI literature. To implement this, we switch from the fixed-effect model to one estimated in first differences. This allows us to include lags of base-year wages in our model. Our other key variables are now presented as changes from base-year t to year $t + 1$ (Δ_{t+1}) such that

$$\Delta_{t+1} \overline{\ln(wage_{ct})} = \beta \Delta_{t+1} \overline{\ln(NATR_{ct})} + \sum \psi_k \overline{\ln(wage_{ct-k})} + \Delta_{t+1} \bar{X}'_{ct} \gamma + \lambda_t + \Delta_{t+1} \bar{\epsilon}_{ct}. \quad (6)$$

In equation (6), $\overline{\ln(wage_{ct-k})}$ is added to the right-hand side, but this variable is not differenced. The intuition for this variable is that wage growth is partly a function of the wage in earlier periods. Weber (2014) showed that the base-year value (i.e., $k = 0$) is generally endogenous. Note that the dependent variable is also a function of the base-year wage. Because the wage is likely serially correlated, lags of the base-year wage may also be endogenous. Weber shows that longer lags or instrumenting for the wage can resolve this endogeneity problem, at least in the context of the ETI. Our proposed instruments are discussed later in this section.

4-4) Identification Issues: Endogeneity of the NATR and Base-Year Income

As noted earlier, in a progressive tax system the income tax rate is positively correlated with income, which is equal to the hourly wage times total hours worked plus non-labor income. This causes (in equations (3) to (6)), the regression error term, to be correlated with the log net average tax rate, $\ln NATR$, biasing downwards the estimated coefficient (i.e., making it more negative). In order to address the endogeneity issue, FW produced a predicted net average tax rate, \widehat{NATR} , to serve as an instrumental variable (IV) for the $NATR$. \widehat{NATR} equals one minus the predicted average tax rate. And, the predicted average tax rate is obtained by running taxpayers' predicted incomes through the TAXSIM model, where predicted income is based on a regression of income against exogenous characteristics.¹⁷

¹⁷ Specifically, predicted income equals predicted wage income plus actual capital income. The wage income component is predicted by running cross-sectional OLS regressions of earnings on exogenous wage determinants from the right-hand side of equation (3).

Predicted income is highly correlated with actual income, thus, we expect $\ln(\widehat{NATR})$ to be highly correlated with $\ln(NATR)$. That is, the IV is relevant. Moreover, since the predicted average tax rate \widehat{ATR} is calculated based on exogenous wage determinants and a small amount of non-wage income, we expect $\ln(\widehat{NATR})$ to be exogenous.

With our pseudo-panel analysis, the (log of the) base-year wage, used to control for secular divergence in wages, is also likely endogenous. With first-differences, the base-year wage is a component of the dependent variable – i.e., wage-rate growth. All else equal, an unusually high base-year wage *causes* the dependent variable to push down the dependent variable. As Weber (2014) shows, in the context of taxable income, this correlation shows up in the error term. Thus, we also include as instruments lags of base-year income. The idea is to include lags whose correlation with base-year income is fairly strong but whose relationship to contemporaneous wage growth is weak.

5) Results

Employing several different techniques, we find strong support for partial wage capitalization. That is, when instrumenting for endogeneity, our estimates of the elasticity of the pre-tax wage with respect to net-of-tax rate are robust and statistically different from both full capitalization and no capitalization. Our estimates from individual cross-sections show considerable year-to-year variation and average -0.22 (or 22-percent capitalization) for men and -0.30 (30-percent capitalizations) for women. Controlling for individual-specific effects and exogenous variation in wage growth, our preferred estimates are -0.48 for men and -0.43 for women. These findings are at odds with FW, who report full capitalization, as well with Leigh (2008) who finds that tax redistribution does not affect pre-tax wages. However, our results are consistent with the

literature in urban and regional economics on location theory combined with significant optimization frictions that attenuate or delay the capitalization process.

5-1) Cross-Sectional Results

First, we estimate wage-capitalization rates separately for years 1997 to 2015 via 2SLS. This most closely resembles the approach employed by FW for years 1983 and 1989. Estimated capitalization rates, reported in Table 2, range from -0.07 to -0.47 for men and -0.12 to -0.44 for women. These estimates are usually statistically different from 0 and always statistically different from -1. The mean and median estimated wage elasticities from the individual cross sections are about -0.23 for men and -0.30 for women. (For each gender, the mean and median are almost the same.) This supports the notion of tax capitalization, but it is far from full capitalization. Without instrumenting, the average elasticity estimate is -3.61 for men and -2.87 for women, suggesting that endogeneity bias dominates any causal effects.

Note that FW's estimates for 1983 and 1989 differ greatly. For year 1983, their core estimates are -0.62 and -0.92 for men and women respectively while using 1989, these estimates are -2.08 and -1.12, suggesting over-capitalization, especially for men.¹⁸ In fact, their estimate of -2.08 is not only statistically different from 0 but also statistically different from -1. FW find much greater rates of wage capitalization than what we find using their approach.

[Table 2 about here]

Why might our estimated capitalization rates be considerably lower than those reported by FW? One possibility is that capitalization rates are lower in recent years than they were during the 1980s. It could be that people are less likely, or slower, to respond to incentives to

¹⁸ For 1989, FW do not include sales taxes. They do show that their 1983 estimates are robust to this exclusion.

move than in the past. A number of studies find that interstate migration rates within the US have generally been declining since the 1980s (see Molloy, Smith and Wozniak, 2011).

5-2) Repeated Cross-Section Results

For the repeated cross-section analysis, we stack all individual CPS cross sections. Using the stacked dataset, we estimate equation (4) via 2SLS. To underscore the importance of instrumenting, the first two columns of Table 3 are estimated via OLS as opposed to 2SLS. Here, the endogeneity of the tax rate yields an implausibly high estimate for the rate at which taxes are capitalized into gross wages. Columns (3) and (4) show that controlling for endogeneity results in an elasticity of the gross wage with respect to the net-of-tax rate of -0.27 for men and -0.31 for women.¹⁹ (The first-stage F-statistics rejects the weak-instrument hypothesis with 99-percent confidence.) In contrast to FW, these results suggest far less than full capitalization.

With respect to our control variables, the wage rate increases nonlinearly with experience. All else equal, the first 10 years of experience is associated with a 27-percent increase in the wage for men and a 22-percent increase for women (relative to 0 years of experience). This relationship between experience and gross wages is lessened as the baseline level of experience increases. That is, as experience increases, the quadratic component of experience (experience squared), which has an estimated coefficient of -0.04, becomes more important. Conditional on observable characteristics, the hourly wage is higher for white workers than for black, Hispanic and “other races”. An additional child under 6 years of age is associated with a 3 percent higher hourly wage. For men, this effect is smaller when the added child is age 6 to 18, and for women

¹⁹ Note, state and time fixed effects are included in the regressions but are not reported in Table 3.

is slightly negative.²⁰ Finally, wages for metropolitan residents (again, conditional on observables) are 13 to 14 percent greater than their non-metropolitan counterparts.

[Table 3 about here]

5-3) Pseudo-Panel Results

A drawback of repeated cross-section analysis is that, because the same people are not followed over time, one cannot control for individual specific effects – e.g., via individual fixed effects or first differences. Likewise, individual earnings histories are not available for constructing instruments. However, traditional panel-data analysis is often compromised because of nonrandom attrition and nonresponse. As discussed earlier, the pseudo-panel approach allows us to exploit the desirable features of panel data without some of the drawbacks. To that end, we created groups using four criteria: gender, birth cohort, education, and state of residence. The baseline pseudo-panel results are presented in Table 4. Again, the OLS-based estimates of wage capitalization (in columns (1) and (2)) appear to be biased upwards (implying capitalization rates in excess of 200 percent), consistent with tax-rate endogeneity. Controlling for this endogeneity results in estimates of -0.67 for men (column (3)) and -0.49 for women (column (4)). For comparison, employing first differences, in place of fixed effects, yields somewhat lower rates of wage capitalization – specifically, estimates of -0.40 for men and -0.36 for women. For each of the four estimates from columns (3)-(6), we strongly reject null hypotheses of both no capitalization and of full capitalization.

Estimated coefficients on the demographic controls are consistent with our expectations. The patterns are similar to those found in the repeated cross-sectional analysis, although the

²⁰ Note that this represents just the correlation, not necessarily a causal relationship.

magnitudes are sometimes different. For example, the relationship between number of children and the wage rate is stronger (i.e., more positive) here.

[Table 4 about here]

5-3-1) Accounting for Secular Trends in Wages

While fixed effects control for individual-specific (or cohort-specific) factors, they do not control for unobservables whose influence on wage growth varies (or has a trend component) over time. To address this possibility, and as detailed in the previous section, we include the base-year wage (in log form), or lags of this variable, to control for exogenous divergence in wages. Table 5 presents elasticities of the gross wage where the wage and NATR are measured in first differences. For each differenced observation, the base year is represented by t and the later year $t + 1$.

In columns (1) and (2) of Table 5, we add the log of the base-year wage and only instrument for the NATR. (Aside from the addition of the log of the base-year wage, these specifications are identical to those from columns (5) and (6) in Table 4.) Here, estimated capitalization rates are much smaller and for men are not statistically different from zero. The drop in the degree of estimated wage capitalization could result from endogeneity of the base-year wage. See Weber (2014).

To further investigate this possible endogeneity, we add as instruments three lags of the base-year wage in columns (5) and (6). This yields estimates of -0.49 for both men and women. These results are modestly larger (in absolute magnitude) than those reported in Table 4 (columns (5) and (6)), which do not include the base-year controls. However, a Sargan-Hansen test rejects the null hypothesis that the instruments are jointly valid – i.e., at least one instrument is not exogenous. In columns (7) and (8) of Table 5, we include just the two- and three-year wage lags, in addition to the NATR instrument. Here our estimated elasticities are very similar to

those in columns (5) and (6), but a Sargan-Hansen test no longer rejects the null hypothesis that the instrument are exogenous. In sum, including the log of base-year income with no additional instruments leads to estimates that are consistent with endogeneity bias. Instead, including longer lags (columns (3) and (4)) or using longer lags to instrument for base-year income (columns (5)-(8)) yields estimates that are similar to each other but very different from those in columns (1) and (2). Furthermore, these estimates are modestly larger (in absolute magnitude) from those in Table 4 (columns (3) and (4)).

[Table 5 about here]

5-3-2) Heterogeneous Costs of Location Decisions

The pseudo-panel regression, specified in equation (5), implicitly assumes that the costs of interstate migration are homogeneous across cohorts and, thus, gross-wage responses to tax changes are also assumed to be the same for all individuals. However, the migration literature suggests that relocating is costly and that these costs vary with age, experience, and education. Migration costs are a type of optimization friction that is likely to attenuate migration responses (and tax capitalization effects), at least over the short run. For example, Topel (1986) finds (using CPS data from 1977 to 1980) that a positive shock to local labor demand generates a wage differential (controlling for location-specific amenities) across locations for workers in occupations with high moving costs. Kennan and Walker (2011) also conclude that, in addition to expected income differentials, interstate migration is influenced by moving costs. According to Kennan and Walker, younger and more educated people are more likely to move. Cross-state migration patterns in our data corroborate these finding. For example, moving across Table 6, holding education constant, annual cross-state migration decreases with age. And, moving down Table 6 shows that, for each age group, cross-state migration increases with education.

[Table 6 about here]

Moving may be relatively less expensive for young people – where age is highly correlated with experience. Using the CPS data of 1988 to 2002, Thompson (2011) concludes that pre-tax wages for high-educated young workers are quite responsive to changes to state tax progressivity. This may be because young (i.e., less experienced) people likely have fewer belongings to move than do older people and they are less likely burdened by the transaction costs associated with selling a house. Nonpecuniary costs may also play a role. The young are less likely to be married or to have children, which may make moving easier. It may also be true that moving costs increase with income but at a decreasing rate. If so, a given percent wage increase is more likely to cover the costs of moving for a higher-educated worker (with a relatively high wage) than for a lower-wage worker (with a relatively low wage). Likewise, even if moving costs did not vary by age, younger workers, with less experience, may still be more likely to move, since they have greater remaining work years. This makes it more likely that an increased wage will, over time, be enough recoup moving costs.

Of course, the degree or speed of wage capitalization (in response to tax differentials) also depends on the substitutability of tasks between groups with different levels of education and of different ages. For example, if paralegals and lawyers are perfect substitutes (although not one for one), capitalization rates for lawyers will be much lower than if the two are not substitutable. Also relevant are labor-supply responses among those not changing locations. For example, local labor-supply responses – such as entry or exit from the labor market – among less-educated, younger people may have capitalization implications for that segment of the labor market. And, some location-specific factors may limit migration and tax capitalization for some occupations. This could include some occupations tied to the location of natural resources or industries with long-established agglomeration economies.

We allow for heterogeneity in wage capitalization rates by modifying equation (5), such that

$$\overline{\ln(wage_{ct})} = \alpha + \beta_0 \overline{\ln(NATR_{ct})} + \beta_1 \overline{\ln(NATR_{ct}) \cdot exp_{ct}} + \overline{\ln(NATR_{ct}) \cdot edu_{ct}} \beta_2 + \bar{X}'_{ct} \gamma + \bar{\mu}_c + \lambda_t + \bar{\varepsilon}_{ct}. \quad (7)$$

Where the log net-of-tax rate is interacted with experience and education dummies, which are used as proxies for migration costs. We hypothesize that β_1 is positive and that the vector β_2 is smaller (or more negative) for higher levels of education. This is because, as discussed above, older (more experienced) workers are often found to be less mobile, while highly-educated workers are often found to be more mobile. As presented in Table 7, the estimated coefficient on the interaction of the net-of-tax rate and experience is close to zero and statistically insignificant for both men and women. In contrast to our hypothesis based on Topel, this suggests little elasticity heterogeneity by experience. Since we are measuring changes to gross wage rates, and not migration directly, this does not rule out variation in migration responses. For example, a set of workers with different levels of experience may share the same occupation – even if their skills sets differ somewhat. Tax migration responses from only the younger, less experienced, members for this occupation could offset the relative lack of mobility among the older, more experienced, members. Thus, it is possible for tax capitalization to vary little with experience but, at the same time, for migration responses to vary by experience. With respect to education, we generally find a positive relationship between education and wage capitalization, as hypothesized. In fact, for those with college degrees, capitalization rates are approximately 100 percent for males and over 80 percent for females. For males, this capitalization rate drops substantially for those with advanced degrees.

[Table 7 about here]

Table 8 presents tax elasticities for the gross wage by level of schooling, based on the estimates from Table 7. Experience, for these estimates, is assumed to equal the sample mean of 24 years for men and 23.6 years. The elasticity of the gross wage increases with education (except for men with advanced degrees) and is usually higher for men, conditional on level of schooling. For those males with more than a high school degree, but without an advanced degree, elasticity estimates range from -0.83 and -1.16. The null hypothesis of full capitalization for these groups of males cannot be rejected.

[Table 8 about here]

6) Conclusion

The ability of subnational governments to redistribute income was once a settled issue, but it appears so no more. A resolution rests on the responsiveness of location decisions to subnational taxes. On this issue, widespread disagreement looms large among the media, politicians, and economists. Leader in states such as Texas, Nevada, and Tennessee argue that low levels of both taxation and progressivity are key to economic growth. Those in this camp will cite employment numbers to support their position. Former Texas Governor Rick Perry embodied this view, actively pursuing California businesses by traveling to the state and touting Texas's favorable tax environment. Meanwhile, leaders in states such as California and New York tout the resilience of their economies and defend their high taxes. In recent years, California has raised its top state income tax rate to 13.3 percent, while, for four straight years, one house of the New York State Legislature has passed some version of universal health care, which would be funded by steeply progressive new taxes. The divide among economists is less extreme but is still pronounced. For example, in influential papers, Feldstein and Wrobel (1998) find evidence consistent with substantial tax responsiveness, whereas Leigh (2008) finds evidence of virtually no responsiveness. Even when focusing on top earners, no consensus emerges. For example, Moretti

and Wilson (2017) find that location decisions of star scientist in the US are quite responsive to taxes. Likewise, Kleven, Landais, Saez, and Schultz (2014) find large migration responses from a Danish policy offering preferential tax rates for high-earning immigrants. However, Young, Varner, Lurie, and Prisinzano (2016) report very small migration responses when using administrative data for “all million-dollar income-earners in the United States over 13 years”.

In this paper, we shed light on this issue, finding support for neither of the extremes. We employ a wage-capitalization model, which is capable of capturing responses that are diffuse over time. Thus, we can capture responses that might be overlooked in models focusing on migration responses made over a several year window. We apply the wage-capitalization model to pseudo-panel data capable of controlling for individual effects. We further include controls for secular divergence in the wage distribution. We find strong evidence of substantial, but partial, wage capitalization. Overall, we find that about 50 percent of state and local redistributive policies are dissipated via wage capitalization. Depending on the specification, this estimate ranges from about 40 to 70 percent, with estimates consistently statistically different from both full capitalization and no capitalization. When allowing for heterogeneous responses, we generally find greater levels of capitalization for those with more education. Close to full capitalization is found for those whose highest level of education is a four-year college degree.

We do not find these results particularly surprising. On one hand, wage capitalization is consistent with the basic economic principle that resources tend to flow to where they are most remunerative. This notion also underlies the standard urban model from urban economics as well as location-based models of quality of life. On the other hand, full capitalization seems extreme. Both pecuniary and nonpecuniary costs to migration are substantial. In some cases, investments in human and physical capital is location specific. Even when it would otherwise make economic sense to relocate an industry, agglomeration effects from past decisions persist. Krugman (1991)

makes the case that this was a major explanation for the persistence of the US Manufacturing Belt, which remained dominant for many decades after inherent advantages from natural resources had been eliminated. Furthermore, focusing on wage capitalization may overstate the ability of subnational governments to redistribute income. In particular, tax-capitalization effects may also occur through housing markets. This is especially true for local housing markets with more inelastic supply and where housing markets are more segmented by income.

References

- Albouy, David, 2009. "The Unequal Geographic Burden of Federal Taxation." *Journal of Political Economy*, 117(4): 635–667.
- Albouy, David, 2016. "What Are Cities Worth? Land Rents, Local Productivity, and the Total Value of Amenities." *Review of Economics and Statistics*, 98(3): 477–487.
- Bakija, Jon, and Joel Slemrod, 2004. "Do the Rich Flee from High State Taxes? Evidence from Federal Estate Tax Returns." Working Paper No. 10645. National Bureau of Economic Research, Cambridge, MA.
- Blundell, Richard, Alan Duncan, and Costas Meghir, 1998. "Estimating Labor Supply Responses Using Tax Reforms." *Econometrica*, 66(4): 827–861.
- Browning, Martin, Angus Deaton, and Margaret Irish, 1985. "A Profitable Approach to Labor Supply and Commodity Demands over the Life-Cycle." *Econometrica*, 53(3): 503–544.
- Brueckner, Jan K., 2000. "Welfare Reform and the Race to the Bottom: Theory and Evidence." *Southern Economic Journal*, 66(3): 505–525.
- Burkhauser, Richard V., Shuaizhang Feng, Stephen Jenkins, and Jeff Larrimore, 2012. "Recent Trends in Top Income Shares in the United States: Reconciling Estimates from March CPS and IRS Tax Return Data." *Review of Economics and Statistics*, 94(2): 371–388.
- Burkhauser, Richard V., Jeff Larrimore, and Kosali I. Simon, 2012. "A "Second Opinion" on the Economic Health of the American Middle Class." *National Tax Journal*, 65(1): 7–32.
- Card, David, and John E. DiNardo, 2002. "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles." *Journal of Labor Economics*, 20(4): 733–783.
- Chetty, Raj, 2012. "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." *Econometrica*, 80(3): 969–1018.

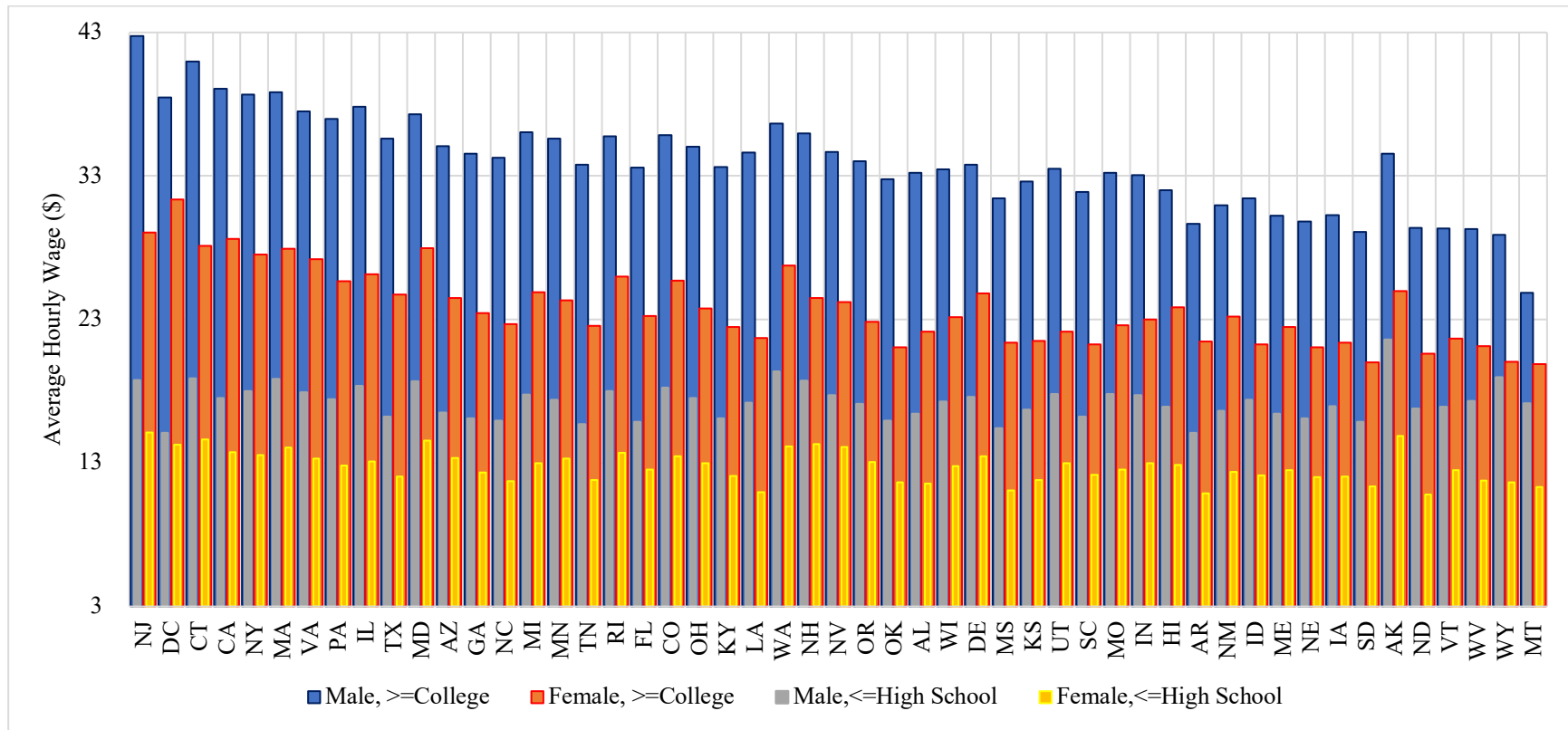
- Cohen, Roger, Andrew E. Lai, and Charles Steindel, 2015. “A Replication of “Millionaire Migration and State Taxation of Top Incomes: Evidence from a Natural Experiment” (National Tax Journal 2011).” *Public Finance Review*, 43(2): 206–225.
- Conway, Karen S., and Jonathan C. Rork, 2006. “State “Death” Taxes and Elderly Migration — The Chicken or the Egg?” *National tax Journal*, 59(2): 97–128.
- Conway, Karen S., and Jonathan C. Rork, 2012. “No Country for old men (or Women) — Do State tax Policies Drive Away the Elderly?” *National tax Journal*, 65(2): 313–356.
- Deaton, Angus, 1985. “Panel Data from Time Series of Cross-Sections.” *Journal of Econometrics*, 30(1): 109–126.
- Feenberg, Daniel, and Elizabeth Coutts, 1993. “An Introduction to the TAXSIM Model.” *Journal of Policy Analysis and Management*, 12(1): 189–194.
- Feldstein, Martin, and Marian Vaillant Wrobel, 1998. “Can State Taxes Redistribute Income?” *Journal of Public Economics*, 68(3): 369–396.
- Glaeser, Edward, 2011. *Triumph of the City*. Penguin Group USA, Inc.
- Gordon, Roger H., and Julie B. Cullen, 2012. “Income Redistribution in a Federal System of Governments.” *Journal of Public Economics*, 96(11–12): 1100–1109.
- Haughwout, Andrew, Robert Inman, Steven Craig, and Thomas Luce, 2004. “Local Revenue Hills: Evidence from Four U.S. Cities.” *Review of Economics and Statistics*, 86(2): 570–585.
- Kennan, John, and James R. Walker, 2011. “The Effect of Expected Income on Individual Migration Decisions.” *Econometrica*, 79(1): 211–251.
- Kirchgassner, Gebhard, and Werner W. Pommerehne, 1996. “Tax Harmonization and Tax Competition in the European Union: Lessons from Switzerland.” *Journal of Public Economics*, 60(3): 351–371.

- Kleven, Henrik J., Camille Landais, and Emmanuel Saez, 2013. "Taxation and International Migration of Superstars: Evidence from the European Football Market." *American Economic Review*, 103(5), 1892-1924.
- Kleven, Henrik J., Camille Landais, Emmanuel Saez, and Esben Schultz, 2014. "Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners' Tax Scheme in Denmark." *Quarterly Journal of Economics*, 129(1): 333–378.
- Krugman, Paul R., 1991. *Geography and Trade*, Leuven University Press: Leuven, Belgium and The MIT Press Cambridge, Massachusetts and London, England.
- Leigh, Andrew, 2008. "Do Redistributive State Taxes Reduce Inequality?" *National Tax Journal*, 61(1), 81–104.
- Moffitt, Robert, 1993. "Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross-Sections." *Journal of Econometrics*, 59(1), 99–123.
- Molloy, Raven, Christopher L. Smith, and Abigail Wozniak, 2011. "Internal Migration in the United States." *Journal of Economic Perspectives*, 25(3): 173–196.
- Moretti, Enrico, and Daniel J. Wilson, 2017. "The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists." *American Economic Review*, 107(7): 1858–1903.
- Musgrave, Richard A., 1971. "Economics of Fiscal Federalism." *Nebraska Journal of Economics and Business*, 10(4): 3–13.
- Oates, Wallace E., 1972. *Fiscal Federalism*. Harcourt Brace Jovanovich, New York, NY.
- Oates, Wallace E., 1999. "An Essay on Fiscal Federalism." *Journal of Economic Literature*, 37(3):1120–1149.
- Pauly, Mark V., 1973. "Income Redistribution as a Local Public Good." *Journal of Public Economics*, 2(1): 35–58.

- Piketty, Thomas, and Emmanuel Saez, 2003. "Income Inequality in the United States, 1913–1998." *Quarterly Journal of Economics*, 118 (1): 1–41.
- Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman, 2018. "Distributional National Accounts: Methods and Estimates for the United States." *Quarterly Journal of Economics* 133(2): 553–609.
- Roback, Jennifer, 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy*, 90(6): 1257–1278.
- Rosen, Sherwin, 1979. "Wages-Based Indexes of Urban Quality of Life." In P. Mieszkowski and M. Straszheim (Eds.), *Current Issues in Urban Economics*, pp. 74-104. Baltimore: Johns Hopkins University Press.
- Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz, 2012. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature*, 50(1): 3–50.
- Thompson, Jeffrey P., 2011. "Costly Migration and the Incidence of State and Local Taxes." Working Paper No. 251. Political Economy Research Institute.
- Tiebout, Charles M., 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy*, 64(5): 416–424.
- Topel, Robert H., 1986. "Local Labor Markets." *Journal of Political Economy*, 94(3): 111–143.
- Verbeek, Marno, 2008. "Pseudo-Panels and Repeated Cross-Sections." In L. Matyas and P. Sevestre (Eds.), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, Chapter 11, pp. 369-383, Berlin:, Springer.
- Verbeek, Marno, and Theo Nijman, 1992. "Can Cohort Data be Treated as Genuine Panel Data?" *Empirical Economics*, 17(1): 9–23.

- Verbeek, Marno, and Francis Vella, 2005. “Estimating Dynamic Models from Repeated Cross-Sections.” *Journal of Econometrics*, 127(1): 83–102.
- Weber, Caroline E., 2014. “Toward Obtaining a Consistent Estimate of the Elasticity of Taxable Income Using Difference-in-Differences.” *Journal of Public Economics*, 117: 90–103.
- Young, Cristobal, and Charles Varner, 2011. “Millionaire Migration and State Taxation of Top Incomes: Evidence from a Natural Experiment.” *National Tax Journal* 64 (2, Part 1): 255–284.
- Young, Cristobal, Charles Varner, Ithai Z. Lurie, and Richard Prisinzano. 2016. “Millionaire Migration and the Taxation of the Elite: Evidence from Administrative Data.” *American Sociological Review*, 81(3): 421–446.

Figure 1. Pre-Tax Wages across States, by Gender and Education



Notes: Each state is represented by two sets of overlapping bars. Each bar represents the average gross wage for one of four gender-education combinations. Education is divided into two broad groups: “high school degree or less” and “at least college degree”. States are ordered from highest to lowest male wage gap (the height of the blue bar minus the height of the gray bar). The sample includes only full-time workers (i.e., weekly hours between 35 to 70) who reported positive earnings and are between 24 to 64 years old. The year spans 1997 to 2015. The gender-education-state combination’s average wage is weighted by the share of five age groups (i.e., 24-29, 30-39, 40-49, 50-59, 60-64) relative to the national level.

Table 1. Summary Statistics, Pseudo-Panel Data

Gender	Variable	Mean	Std. Dev.	Min	Max	
Male	Wage (\$)	24.87	16.92	0	833.33	
	Average Tax Rate (percent)	30.35	6.5	-29.48	65.75	
	Experience (years)	24	12.06	0	49	
	Race:					
	White	0.73	0.28	0	1	
	Black	0.08	0.14	0	1	
	Hispanic	0.13	0.23	0	1	
	Number of children:					
	Under 6	0.3	0.38	0	5	
	Between 6 and 18	0.63	0.53	0	6	
	Over 18	0.06	0.12	0	2	
	Filing status:					
	Joint filer	0.68	0.24	0	1	
	Household Head	0.04	0.09	0	1	
	MSA	0.75	0.26	0	1	
	Cell size	16	16.6	1	162	
	Cohorts	2,510				
	Observations	39,619				
	Female	Wage (\$)	17.93	10.05	0	261.83
		Average Tax Rate (percent)	28.85	8.21	-0.32.05	59.06
Experience (years)		23.62	12.04	0	49	
Race:						
White		0.69	0.3	0	1	
Black		0.13	0.19	0	1	
Hispanic		0.12	0.22	0	1	
Number of children:						
Under 6		0.23	0.34	0	5	
Between 6 and 18		0.63	0.59	0	6	
Over 18		0.06	0.13	0	2	
Filing status:						
Joint filer		0.56	0.25	0	1	
Household Head		0.15	0.18	0	1	
MSA		0.76	0.27	0	1	
Cohort size		13	12.95	1	129	
Cohorts		2,458				
Observations		38,301				

Notes: The summary statistics are based on pooled pseudo-panel data for years 1997 to 2015. The sample includes only full-time workers (i.e., weekly hours between 35 to 70) who reported positive earnings and are between 24 to 64 years old. The pseudo-panel groups are created based on four characteristics: gender, five-year-interval birth cohorts, highest education attainment level (less than high school, high school, some college, college, and advanced degree), and the state of residence. Groups with fewer than 7 time periods are dropped. In the table, the dropped race is “other races” and the dropped filing status is “single filer.” The average tax rate includes income (federal and state), FICA, sales, and property taxes. Experience is approximated by age minus years of education minus 6.

Table 2. Elasticity of the Gross Wage on the Net-of-Tax Rate: Individual Cross-Sections

Year	Men		Women	
	$\hat{\beta}$	Std. Error	$\hat{\beta}$	Std. Error
1997	-0.23	(0.13)	-0.42**	(0.11)
1998	-0.24*	(0.12)	-0.38**	(0.09)
1999	-0.07	(0.12)	-0.32**	(0.09)
2000	-0.16	(0.11)	-0.25**	(0.09)
2001	-0.07	(0.12)	-0.18	(0.10)
2002	-0.20*	(0.10)	-0.44**	(0.07)
2003	-0.47**	(0.08)	-0.40**	(0.07)
2004	-0.26**	(0.08)	-0.18*	(0.07)
2005	-0.17	(0.09)	-0.30**	(0.07)
2006	-0.15	(0.11)	-0.19*	(0.08)
2007	-0.29**	(0.10)	-0.30**	(0.07)
2008	-0.23**	(0.08)	-0.36**	(0.07)
2009	-0.33**	(0.11)	-0.35**	(0.07)
2010	-0.46**	(0.12)	-0.34**	(0.06)
2011	-0.08	(0.08)	-0.28**	(0.07)
2012	-0.27**	(0.08)	-0.29**	(0.07)
2013	-0.10	(0.06)	-0.12*	(0.06)
2014	-0.37**	(0.09)	-0.28**	(0.08)
2015	-0.12	(0.07)	-0.27**	(0.06)
Mean	-0.22	(0.14) ²¹	-0.30**	(0.10)

Notes: These elasticities are obtained from cross-sectional regressions of log wage on log net-of-tax rate, ln (NATR), individual demographics (such as experience, experience squared, education, race, number of children age 0-6, 6-18, and above 18, filing status, metropolitan residence status, and the state fixed effects. The sample includes only full-time workers (i.e., weekly hours between 35 to 70) who reported positive earnings and ages 24 to 64 years old. NATR equals 1 minus average tax rate. All estimated coefficients are statistically from -1 at 0.01 significance level. Robust standard errors shown in parentheses. * p<0.05, ** p<0.01.

²¹ The standard error of the mean of cross-sectional estimates is obtained by $\sqrt{\frac{1}{T-1} \sum_{t=1}^T \left(\hat{\beta}_t - \frac{1}{T} \sum_{t=1}^T \hat{\beta}_t \right)^2}$.

Table 3. Elasticity of the Gross Wage on the Net-of-Tax Rate: Repeated Cross-Sections

Dependent Variable:	(1)	(2)	(3)	(4)
ln(gross wage)	Men	Women	Men	Women
ln(NATR)	-3.42** (0.01)	-2.75** (0.01)	-0.27** (0.02)	-0.31** (0.02)
experience (in 10s)	0.15** (0.00)	0.10** (0.00)	0.27** (0.00)	0.22** (0.00)
experience-squared (in 100s)	-0.02** (0.00)	-0.01** (0.00)	-0.04** (0.00)	-0.04** (0.00)
# of children under 6	0.17** (0.00)	0.12** (0.00)	0.03** (0.00)	0.03** (0.00)
# of children 6-18	0.09** (0.00)	0.04** (0.00)	0.02** (0.00)	-0.01** (0.00)
# of children over 18	0.08** (0.00)	0.03** (0.00)	0.00 (0.00)	-0.03** (0.00)
High school	0.06** (0.00)	0.13** (0.00)	0.23** (0.00)	0.26** (0.00)
Some College	0.12** (0.00)	0.22** (0.00)	0.38** (0.00)	0.44** (0.00)
College	0.25** (0.00)	0.38** (0.00)	0.65** (0.01)	0.72** (0.01)
Advanced	0.41** (0.00)	0.55** (0.00)	0.91** (0.01)	0.96** (0.01)
Black	-0.09** (0.00)	0.03** (0.00)	-0.18** (0.00)	-0.07** (0.00)
Hispanic	-0.11** (0.00)	-0.11** (0.00)	-0.22** (0.00)	-0.15** (0.00)
Other races	-0.09** (0.00)	-0.06** (0.00)	-0.12** (0.00)	-0.05** (0.00)
Joint filer	0.35** (0.00)	0.09** (0.00)	0.16** (0.00)	0.005* (0.00)
Household Head	0.38** (0.00)	0.43** (0.00)	0.10** (0.01)	0.03** (0.00)
MSA	0.03** (0.00)	0.05** (0.00)	0.13** (0.00)	0.14** (0.00)
Constant	0.59** (0.01)	0.73** (0.01)	1.55** (0.01)	1.34** (0.01)
Observations	634,117	492,162	633,020	491,916
R-squared	0.429	0.391	0.275	0.283
2SLS	NO	NO	YES	YES
State FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Elasticity= -1	NO**	NO**	NO**	NO**

Notes: Each specification regresses the log wage on log net-of-tax rate, ln (NATR), individual demographics, and state and year fixed effects for years 1997 to 2015. The sample includes only full-time workers (i.e., weekly hours between 35 to 70) who reported positive earnings and are between 24 to 64 years old. NATR equals 1 minus average tax rate. The dropped categories for education, race, and filing status dummies are less than high school, white, and single filer, respectively. Robust standard errors shown in parentheses. * p<0.05, ** p<0.01.

Table 4. Elasticity of the Gross Wage on the Net-of-Tax Rate: Pseudo Panel

Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)
ln(wage) or Δln(wage)	Men	Women	Men	Women	Men	Women
ln(NATR)	-3.02** (0.04)	-2.01** (0.03)	-0.67** (0.08)	-0.49** (0.06)		
Δln(NATR)					-0.40**	-0.36**
Observations	39,618	38,301	39,618	38,301	36,681	35,028
R-squared	0.58	0.51	0.52	0.47	0.09	0.07
Number of cohort	2,510	2,458	2,510	2,458	2,510	2,458
2SLS	NO	NO	YES	YES	YES	YES
FE	YES	YES	YES	YES	NO	NO
FD	NO	NO	NO	NO	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Elasticity=-1	NO**	NO**	NO**	NO*	NO*	NO*

Notes: Columns 1 to 4 report FE pseudo-panel regressions of log wage, $\ln(\overline{wage})$ on the log net-of-tax rate, $\ln(\overline{NATR})$, demographics, and cohort and year fixed effects for years 1997 to 2015. Estimates on the demographic variables are suppressed. Columns 5 and 6 are estimated in first differences. The sample includes only full-time workers (i.e., weekly hours between 35 to 70) who reported positive earnings and are between 24 to 64 years old. NATR equals 1 minus the average tax rate. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$.

Table 5. Elasticity of the Gross Wage on the Net-of-Tax Rate: Pseudo Panel with Controls for Secular Wage Divergence

Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta \ln(\text{gross wage})$	Men	Women	Men	Women	Men	Women	Men	Women
$\Delta \ln(\text{NATR})$	-0.14 (0.10)	-0.15* (0.07)	-0.48** (0.11)	-0.42** (0.09)	-0.49** (0.11)	-0.49** (0.08)	-0.48** (0.11)	-0.43** (0.09)
Observations	36,681	35,028	28,926	27,159	28,729	26,817	28,926	27,159
Cohorts	2,510	2,458	2,506	2,445	2,488	2,406	2,506	2,445
Controls: $\log(\text{wage})$	base-year	base-year	2, 3 lags of base-year	2, 3 lags of base-year	base-year	base-year	base-year	base-year
$\log(\text{wage})$ Instruments	--	--	--	--	1, 2, 3 lags of base-year	1, 2, 3 lags of base-year	2, 3 lags of base- year	2, 3 lags of base- year
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Sargan-Hansen P-value	--	--	--	--	0	0	0.79	0.39
Weak Identification F- statistics	102	3360	621	2526	159	790.00	208	840

Notes: All specifications are run as 2SLS regression in first difference of the log wage, $\Delta \overline{\ln(\text{wage})}$ for years 1997 to 2015. All specifications include the first difference of the log predicted net-of-tax rate as an instrument for $\Delta \overline{\ln(\text{NATR})}$. The sample includes only full-time workers (i.e., weekly hours between 35 to 70) who reported positive earnings and are between 24 to 64 years old. Estimates on the demographic variables are suppressed. The NATR equals 1 minus average tax rate. All estimated elasticities are statistically different from -1 at the 0.01 significance level. Robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$.

Table 6. Annual Interstate Migration Rates (percent), by Education and Age Groups

Education	Age				
	24-29	30-39	40-49	50-59	60-64
< High School	3.31%	1.85%	1.22%	1.04%	0.92%
High School	3.55	1.99	1.34	1.05	0.89
Some College	3.81	2.18	1.40	1.15	1.22
College	6.26	2.86	1.51	1.26	1.18
Advanced	9.02	4.18	1.91	1.35	1.48

Notes: The rate of annual interstate migration is calculated based on March CPS data from 1997 to 2015. The sample includes only full-time workers (i.e., weekly hours between 35 to 70) ages 24 to 64 years old with positive earning.

Table 7. Elasticity of the Gross Wage with Heterogeneous Migration Costs, Pseudo-Panel Regressions

Dependent Variable:	(1)	(2)
ln(gross wage)	Men	Women
ln(NATR)	-0.35** (0.12)	-0.20* (0.10)
ln(NATR)*experience	-0.04 (0.03)	0.02 (0.02)
ln(NATR)*High School	-0.48** (0.11)	-0.48** (0.10)
ln(NATR)* Some College	-0.38** (0.12)	-0.61** (0.11)
ln(NATR)* College	-0.71** (0.13)	-0.62** (0.13)
ln(NATR)*Advanced	0.23 (0.17)	-0.59** (0.14)
Observations	39,618	38,301
R-squared	0.52	0.47
Number of cohorts	2,510	2,458
2SLS	YES	YES
Cohort FE	YES	YES
Year FE	YES	YES

Notes: Fixed effect regressions regress log wage, $\ln(wage)$, on the log net-of-tax rate, $\ln(NATR)$, interactions of $\ln(NATR)$ with experience and education dummies, cohort means of demographics, cohort and year fixed effects for years 1997 to 2015. Estimates on the demographic variables are suppressed. The “less than high school” dummy is dropped as the reference group. The sample includes only full-time workers (i.e., weekly hours between 35 to 70) ages 24 to 64 with positive earnings. The NATR equals 1 minus average tax rate. Robust standard errors shown in parentheses. * $p < 0.05$, ** $p < 0.01$.

Table 8. Elasticity of the Gross Wage on the Net-of-Tax Rate, by Level of Schooling

	< High School	High School	Some College	College	Advanced
Men	-0.45**	-0.93**	-0.83**	-1.16**	-0.22
	(0.09)	(0.11)	(0.12)	(0.13)	(0.17)
Elasticity= -1	NO**	YES**	YES**	YES**	NO**
Women	-0.15	-0.63**	-0.76**	-0.76**	-0.74**
	(0.08)	(0.09)	(0.09)	(0.11)	(0.12)
Elasticity= -1	NO**	NO**	NO**	NO**	NO**

Notes: The elasticity of the gross wage for each level of schooling equals the linear summation of the reference coefficient and the coefficients on the relevant interaction terms (with level of education), and all evaluated the sample mean for experience. * p<0.05, ** p<0.01.