Relocation of the Rich: Migration in Response to Top Tax Rate Changes from Spanish Reforms

David R. Agrawal   Dirk Foremny[*]

April 2018

Abstract

A Spanish reform granted regions the authority to set income tax rates, resulting in substantial tax differentials. Using administrative data, conditional on moving, taxes have a significant effect on location choice. A one percent increase in the net-of-tax rate for a region relative to others increases the probability of moving to that region by 1.7 percentage points. We estimate an elasticity of the number of top taxpayers with respect to net-of-tax rates of 0.85. The mechanical increase in tax revenue due to higher tax rates is larger than the loss in tax revenue from the net out-flow of migration. **JEL**: H24, H31, H73, J61, R23

High-income taxpayers may be literally "worth their weight in gold" (Wildasin 2009) to the government where they reside. As a means of tax avoidance, individuals may move in response to tax differentials resulting from residence based local income taxes. Tax avoidance typically arises when taxable income can be shifted in a way that it becomes subject to a favorable tax treatment (Piketty and Saez 2013), and mobile taxpayers might simply relocate their tax residence to reduce their income tax burden. As a result of tax induced mobility by high-income taxpayers, governments may be unable to engage in progressive redistribution (Epple and Romer 1991; Feldstein and Wrobel 1998) and tax competition may intensify (Wildasin 2006). Despite the policy importance of analyzing taxation in an open economy, most studies have analyzed avoidance responses of taxable income (Feldstein 1999) although a literature on tax-induced mobility has emerged.

We provide evidence on migration resulting from a major Spanish tax reform and fiscal decentralization. In the early 2000s, all *Autonomous Communities* (regions or states) in Spain had the same top marginal tax rate. In 2011, Spanish regions began changing their top marginal tax rates (MTR) in response to a reform that gave regions the authority to adjust rates and the corresponding tax brackets. In 2014, top marginal tax rates diverged across regions by as much as 4.5 percentage points. For an individual earning 300,000 Euros, this amounts to a tax differential of 10,000 Euros. These disparities in regional top tax rates led the popular press to dub low-tax regions such as Madrid as "tax havens" or one of several "paradises on Earth."

Research on migration requires linked data of individuals in the country of origin and destination, which is fairly complicated to obtain.[1] Exploiting sub-national variation is therefore an appealing alternative. However, personal income in most countries is taxed at the central level of government and only a few countries tax personal income at the regional or local level (i.e., United States, Canada – Milligan and Smart (2016), Swe-

---

[1] Kleven, Landais and Saez (2013) and Akcigit, Baslandze and Stantcheva (2016) are exceptions. They focus on selected sub-groups of the population for which access to individual income data linked across countries is not needed. Bakija and Slemrod (2004), Moretti and Wilson (2017) and Young et al. (2016) are state level examples.

den, Italy, Switzerland – Martínez (2016)). Given the expected mobility and avoidance responses, some of these countries only allow for small differentials across jurisdictions by limiting the tax-setting power of state and local governments. In countries with more substantial autonomy, regional personal income taxes have been implemented decades ago and large administrative data are not available for time periods before their implementation. Further, in the U.S., income taxes are often employment-based rather than residence-based, which means that for local moves, individuals may change jobs rather than residence to reduce taxes (Agrawal and Hoyt 2018).[2] The reform in Spain granted substantial autonomy to the regions on a purely residence-based tax system. Tax administration remains with the national authorities which facilitates access to individual micro-data that is available before and after the decentralization.

We use individual Social Security data for a sample of the population of Spain (excluding Navarre and Basque Country, which are not included in the data) from 2005 to 2014 to study the migration decisions of the rich in response to this unique fiscal decentralization. Our paper makes several contributions. First, we focus on all high income individuals rather than a select group of highly mobile individuals such as star scientists (Moretti and Wilson 2017; Akcigit, Baslandze and Stantcheva 2016), athletes (Kleven, Landais and Saez 2013), or foreigners subject to preferential taxation (Kleven et al. 2014; Schmidheiny and Slotwinski 2018). In terms of the scope of the sample, Young et al. (2016) are the closest to our paper and utilize population level U.S. tax return data for all millionaires in the United States over a thirteen year period. Exploiting state-to-state migration of millionaires and an empirical design comparing millionaire populations at state borders, Young et al. (2016) concludes that although taxes matter, it is with only very small economic significance.[3] Second, we study migration using a random sample

---

[2]Approximately 75 million people live in MSAs that cross state borders. Of these, two-thirds live in MSAs that have an employment-based component to income taxes.

[3]We make several contributions relative to Young et al. (2016). First, we study migration patterns of the rich (above 90,000 Euros) and not just the very rich (millionaires). We also study migration in a setting where regional taxes are purely residence-based;

2

of population level administrative data for a complete panel of all regions in a country; Young et al. (2016) have similar data and a single state analysis includes Young and Varner (2011). This administrative data provides us detailed information on industry and occupation that allows us to determine the generalizability of the prior literature focusing on specific occupations. We find significant effects of taxes on location choices, but an elasticity of the number of top taxpayers that is less than unity. We then contribute to the literature by using a theoretical model of revenues to simulate the implications for the fiscal authorities.

Discussion of taxing top incomes comes within the context of widening income inequality. One discussed policy response to widening inequality is changing top tax rates. Bonhomme and Hospido (2017) document that earnings inequality in Spain declined from about 1995 to 2007, but that it has risen dramatically since 2007 and is back to its 1995 level. Studying mobility in Spain is especially important given the implications for redistributive tax policy. The mobility response – especially of high income taxpayers – is critical to understanding whether increasing progressivity at the regional level is a viable policy response to increasing inequality. Mobility in response to more progressive tax policy could threaten the ability to engage in redistribution given that the optimal degree of redistribution will decline as the mobility elasticity increases (Mirrlees 1982).

The paper proceeds as follows: we use a 4% random sample of administrative data that is released publicly to study tax-induced mobility. We use individual Social Security and tax administration data that contains information on each taxpayer's income that is not top coded. These data also contain information on the taxpayers declared location of residence in addition to certain characteristics reported to the Social Security administration. These data do not contain tax rates, so we write our own tax calculator from regional tax codes that simulates average and marginal tax rates back to 2005.

We first conduct an aggregated region pair analysis. We calculate for each year in our data, the stock of top-taxpayers for every region pair combination in Spain and

in the United States, taxes may have a residence and employment-based component. Finally, we show heterogeneous effects by industry and occupation.

construct the log ratio of the stocks across pairs; in addition, we calculate the net-of-tax rate differential between each of the region pairs. We then show that higher individual income taxes reduce the stock of top-taxpayers after accounting for destination fixed effects, origin fixed effects, and year fixed effects. The stock elasticity is approximately 0.85. Tax policy is not set randomly and any state-specific unobservable that is correlated with taxes and migration may threaten our results. To deal with this, we show that migration effects follow tax changes and do not pre-date them such that there are no pre-trends in the periods prior to the reform. We also show a placebo test: pre-reform population stock changes show no correlation with post-reform tax differential changes.

Then, we turn to an individual level analysis that studies whether individuals are more likely to select low-tax regions, conditional on moving. Our empirical choice model exploits individual variation in tax rates across the fifteen Spanish regions; given we exploit person-specific tax rates, our model allows us to account for region by year fixed effects and individual characteristics that are allowed to vary by region in order to capture counterfactual wages in alternative regions. This approach has the advantage of allowing us to account for fixed characteristics of the mover that are constant across alternative regions, any sorting based on characteristics, as well as for other policy changes that affect all individuals in the top of the income distribution. A one percent increase in the net-of-tax rate for a region relative to others increases the probability of moving to that region by 1.7 percentage points. Although many things may matter for decisions on where to move, taxes appear to be important. These estimates suggest that the 0.75 percentage point average tax rate (ATR) differential between Madrid and Cataluña in 2013 increases the probability of moving to Madrid by 2.25 percentage points.

We then exploit the administrative data on occupation and industry to show that taxes play a stronger role for certain occupations and industries. Testing for heterogeneity across occupation and industry helps to inform the recent policy debate on the efficiency of tax schemes for top earners in specific occupations and industries. Several OECD countries have preferential tax schemes for foreigners in high-income occupations. We can shed light on the efficiency of these tax schemes. First, we replicate the result

in the prior literature for scientists and find that those in the "professional/scientific" and "health" industry have large and significant migration effects; entertainers including athletes have insignificant effects, likely due to sample size. Then, we look at other occupations and industries to determine if the estimates for the occupations studied in the prior literature generalize. Our results indicate that self-employed (a self-employed individual will only have income in the data for formal contracts with registered firms) and "higher-ability" occupations are more sensitive to taxes. Our industry-level data demonstrates substantial heterogeneity with the largest effects emerging in health, finance, real estate, and information, in addition to the scientific industries studied previously.

Our analysis comes with a caveat: our data do not allow us to disentangle a real move from a fraudulent move where the taxpayer changes residence to a second home without actually changing where they spend the majority of the tax year. In so much as this is possible, the presence of such evasion implies that mobility includes both real responses as well as tax evasion responses. From a tax revenue perspective, it does not matter if the move is a real response or simple misreporting; from a labor supply perspective, real moves may be more important.

As noted in Saez, Slemrod and Giertz (2012), absent both classic and fiscal externalities, the elasticity of taxable income (ETI) suggests that the revenue maximizing tax rate on top incomes may be as high as 80% with a broad income tax base. However, changes resulting from mobility across regions are not generally captured in these estimates and therefore understanding mobility has important implications for understanding the optimal top income tax rate. In order to interpret the elasticities that we estimate, we simulate a revenue maximization model incorporating migration. The model suggests that the effect of changes in taxes on revenue can be decomposed into a mechanical (tax rate) effect from higher taxes, a behavioral effect from changes in taxable income, and a migration effect. The last effect depends on the stock elasticity of migration. Using our stock elasticities, we find the mechanical effect dominates the other effects for all regions in Spain, which has important implications for how much additional revenue a region can raise [lose] by raising [lowering] its top tax rates. For the region of Madrid, its lower

rate relative to the central government's tax rate in 2014 results in revenue falling by 50 million Euro due to the mechanical effect of the lower tax rate. Using our mobility estimates, migration effects only contribute 9 million Euro more in revenue. For behavioral responses to offset the mechanical effect net of mobility effects, the elasticity of taxable income would need to be 1.40, which is well above reasonable estimates of it. We conclude that, in the short-run, migration does not pose a large threat to redistributive taxation.

1 Institutional Details

Spain consists of 17 autonomous communities (in Spanish: *comunidades autónomas*) which are comparable to states or regions in other countries. The autonomous communities are governed according to the Spanish constitution. Furthermore, the individual competences which each region assumes are regulated by a region specific organic law, known as Statute of Autonomy. Important for our purpose is that taxes are due at the place of residence (*residencia habitual*), which is declared in the local municipality of residence. Since 1994, the regions receive a share of the Personal Income Tax (*Impuesto sobre la Renta de las Personas Físicas*) as part of their revenues, but it was only in 1997 that partial autonomy over marginal tax rates was delegated to the regions (see Durán and Esteller 2005). Initial regional autonomy was quite limited as regional level marginal tax rates applied only to 15% of the tax base and thus autonomous communities had little interest in changing marginal tax rates. Instead, they focused on setting tax credits, mostly for housing and renting as well as some personal circumstances such as ascendants and decedents.[4] In 2007, following some reforms, the regional-level individual tax rates still had to complement the common tax brackets set by the central level, but they were applied to a larger share (35%) of their residents tax base. Therefore, in 2007, Madrid was the first autonomous community which changed marginal tax rates, followed by La Rioja and Valencia in 2008. These regions implemented top marginal tax rates which were slightly lower (less than 0.1 percentage point) than the tax schedule of the central government. Murcia followed in 2009, but returned to the common central scheme

---

[4]Tax credits predominantly lower the effective tax burden for the poor, as many fade out with income. See the appendix for details.

thereafter. These initial reforms resulted in very small differences in taxes across regions.

Another major wave of decentralization reforms followed this process in 2009 (last laws approved in July 2010) but regions could not exercise their new rights until 2011. Regions could now keep the revenues collected from half of the entire tax base in their territory. In addition, regions were also given the right to introduce new tax brackets on top of those implemented by the central government. In 2011, with both the ability to construct new brackets and marginal tax rates in hand along with added incentives to retain more of the tax revenue, several regions increased marginal tax rates substantially, while the ones which decreased them previously lowered their rates (Bosch 2010). Another reason for the immediate reaction of regional governments was that in 2011, the central government raised marginal tax rates substantially and regions used this event to increase simultaneously their own tax rates, or decrease them to counteract the national increase. In subsequent years, some further changes in regional top tax brackets were implemented, but the pattern of high versus low-tax regions as of 2011 generally persists.[5]

The regional tax changes are salient to top taxpayers. Tax forms compute an individual's average regional and central tax liability separately so that the individual sees both average tax rates. When filing taxes in April, taxpayers are asked to state their place of residence. A change of their address can be done online at the same page where individuals submit their tax declaration and becomes effective immediately.[6]

In Spain, the personal income tax is a dual tax which separates the income tax base and the capital income tax base. The reform *only* allowed regions to alter marginal tax

---

[5]A confounding factor could be the re-introduction of the wealth tax at the end of 2011. The decision was taken at the end of 2011, such that an immediate response in that year – and even one in 2012 is unlikely to happen. Further, the tax has been introduced as an explicitly temporary measure to reduce fiscal problems during the Great Recession. Not until the end of 2012 did the government announced that the tax will also be applied in the following year, again without establishing the tax permanently.

[6]We do not observe the location declared on the return. However, tax inspectors might check any change of the fiscal residence with the data from the local register.

rates of the labor income tax base, while the capital income tax base remained taxed under a common tax schedule. Given that we will use Social Security data to study migration, pure rentiers (capital income only) will be absent from the data. However, given that these rentiers face a common national tax rate on capital income, the decentralization of the labor income tax base is irreverent for these individuals. The reform did not affect corporate taxes, so we do not have to worry about any correlation with corporate taxes.

1.1 Descriptive Figures of the Reforms

In 2010, all regions in our data set have tax rates that are within 0.10 percentage points of each other. But, by 2014, substantial spatial variation had emerged. All tax rates increased over time in levels – although some decreased relative to the central government rate, which was changing over time. Figure 1 shows the changes for all regions and for all brackets. In order to ease interpretation, we show the tax changes relative to what the tax rate would be if the region had simply adopted the national tax rates in that year. Relative to this standard, some regions decreased their tax rates while others increased their tax rates. Immediately following the reform, top tax rates diverged by 5 percentage points. This pattern persisted with some changes to lower bracket tax rates.

Given that many tax brackets change (not just the top ones), we need to justify our focus on the top of the income distribution. The vertical line in figure 1 shows the cutoff for the top 1% of income. Notice that the top 1% – incomes above 90,000 Euro approximately — experienced the largest changes in tax rates. The tax differences for individuals even in the top 2 to 5% were relatively small across regions – this is because even if marginal tax rates differed across regions, average tax rates were relatively similar.

Regions that raise their tax rates by the largest amounts might be those regions where the mobility of top earners has been declining or where the stock of top earners is large or small. Regions had little information about how individuals would respond to tax changes given the immediate decentralization and tax changes following the reform. Simple correlations in appendix A.1 show that characteristics of the region related to the top 1% seem to have small effects on the tax changes. Larger correlates with the tax changes are political associations, debt, and income conditions.

Although the equilibrium tax rates may be a result of a rather arbitrary political process, this is not to say that the resulting equilibrium tax rates are as good as random. In particular, the resulting tax rates following the policy decentralization may be a function of unobservable characteristics. While this is not something that we can rule out entirely, we provide evidence that this does not appear to be the case. First, we do not find pre-trends in the populations of the regions; changes in the population stocks occur after the tax changes. Second, we show that post-reform tax rate changes do not predict pre-reform populations or migration flows suggesting that regions do not set taxes based on pre-reform characteristics. Of course, other unobservable changes in state policy or state shocks may exist. While we cannot rule this possibility out with aggregate data, we then turn to individual level data where we can control for state by year shocks.

2 Description of Data

We use panel data from 2005-2014 from Spain's Continuous Sample of Employment Histories (*Muestra Continua de Vidas Laborales*, MCVL). The data is provided by the Ministry of Employment and Social Security *(Ministerio de Empleo y Seguridad Social)*. This administrative data matches individual microdata from social security records with data from the tax administration (*Agencia Tributaria,* AEAT), and official population register data (*Padrón Continuo*) from the Spanish National Statistical Office (INE).[7] The Social Security administration publicly releases an approximately 4% non-stratified random sample (over 1 million observations each year) of the population of individuals which had any relationship with Spain's Social Security system in a given year due to work, receiving unemployment benefits, or receiving a pension. These data have been previously in applied work on labor and urban economics (Bonhomme and Hospido 2017; De la Roca and Puga 2017). Individuals from Navarre and the Basque Country do not appear in the data because these regions operate independent fiscal systems.[8] If an individual is

---

[7]The tax return data often reports the location of work. Social Security data contains residence information based off local registers.

[8]We exclude the individuals living there from our analysis. We treat people moving to those two regions as people leaving the sample for any other reason (moving abroad,

in the data, they remain in the data as long as they have contact with the Social Security ministry, but new observations enter each year so that it remains representative. Self-employed individuals that make contact with the Social Security system do appear in our data, however, we only observe income for them if they have a relationship with a registered firm, as the firm remits taxes on their behalf. Self-employed individuals that do not have contracts with firms do not appear in the data; however, we believe that at the very top of the income distribution most self-employed individuals provide services to firms such that they will be covered in our sample. Nonetheless, even for self-employed individuals with some contracts with firms, we may mismeasure their true income.

From 2005 to 2014, the Social Security data are matched to income from tax data. These income tax data are valuable because they are not subject to censoring; Social Security contributions are censored and do not contain some portions of income that are important for high income tax payers (see Bonhomme and Hospido 2017). Given we will focus on top income taxpayers it is important we have income data that is not censored and contains all sources of income. The observational unit of the raw data is based on each contact an individual had within a given year with Social Security. We define the main work affiliation in each year as the one which was active for the longest time span since starting work. We aggregate this data at the individual level to obtain a panel data set which sums all individual income sources in a given year for a given tax payer.[9]

We define a change of location if an individual changed his or her residence between t and $t-1$.[10] Residence data of the current year is updated using the residence of April in $t+1$, which ensures that this period overlaps with the tax year as tax declarations are

_____

etc.). However, we do include people moving from those two regions to another region in Spain as we observe their income in the new destination and know the origin from the social security database. We furthermore exclude Ceuta and Melilla, two autonomous cities (not autonomous communities) on continental Africa.

[9]Only some of the sample are reported as "married" with a substantial fraction declaring "other." Two individuals may move in our data when they are a common household.

[10]The transition matrix of movers in are given in appendix A.2.

due in April to June. As an example, an individual would be characterized as a mover in 2012 if he was living in a different region between April 2012 and April 2013 compared to his residence between April of 2011 and April 2012. In this way, his 2012 income is the relevant one for tax purposed in the region he moved to.[11] While residence information is available at much smaller spatial units, in this paper, we will define a "mover" as an individual that relocates *across* (not within) regions. One reason for this is that we only observe municipality codes for individuals living in sufficiently large cities, which means that many within region moves remain unobserved to us.

We construct taxes using the sum of all reported income by different employers within each year which is subject to the personal income tax (labor income, reported self-employed income, and income in-kind). Given this information and other attributes, we simulate average and marginal tax rates for each individual in each year for each region using the information in the tax code provided by official documents.[12] This simulation takes into account the variation of marginal tax rates, their brackets, and basic deductions and tax credits for ascendants, decedents, and disabilities. We do not take into account any further region specific deductions or tax credits. However, given that we focus on high income individuals, this would almost never affect the marginal tax rate and the average tax rate only to a negligible amount as those omitted policies are targeted to low income individuals. We use the tax calculator to simulate the tax rate in the region of

---

[11]Registration is mandatory within three month and municipalities have an incentive to register citizens because they receive transfers allocated on a per capita base (Foremny, Jofre-Monseny and Solé-Ollé 2017). An alternative location variable which is available in this data-set are the region the firm provides when they remit taxes for the individual. This does not have any legal effect on tax declarations, but rather corresponds to the address on file with the employer. We observe 57% of movers have firms reporting the same state from the registrar data. Adding the observations for which the province declared by a firm coincides with the residential province before moving increases the share to 96%. This indicates that there is a lag of updating the firm database.

[12]We use the tax laws to write a tax calculator similar to TAXSIM. See appendix A.3.

residence and the tax rates in all counterfactual alternative regions.

Summary statistics of our data are given in appendix A.4. Unique to our setting is detailed data on occupation and industry. The appendix also shows descriptives concerning the top occupations and industries of movers in the top 1%.

3 Aggregate Analysis

For simplicity, consider a two region economy where $r = o, d$ indexes the two regions, which we call origin and destination for simplicity.[13] The utility of top income individuals living in region r in period t is given by:

$$V_{r,t} = \alpha u(c_{r,t}) + \pi v(g_{r,t}) + \mu_r - \gamma \rho(N_{r,t}) \tag{1}$$

where $c_{r,t}$ is private consumption of the individual, $g_{r,t}$ is public services consumption, μ_r is the value of other amenities that are specific to living in the region. The function ρ is a disutility that depends on population. The $\rho(N_{r,t})$ function allows us to indirectly bring in housing markets into the problem: a region becomes less attractive, the larger is its population perhaps because housing prices increase. In particular, fewer people in a region mean the cost of housing will be lower which raises utility relative to a region with more people and higher housing costs.[14] In particular, this congestion cost it is an alternative mechanism to get to a spatial equilibrium even without formally modeling housing price adjustments. Following the standard in the literature, we assume that the separable functions u, v, and ρ each take on the log functional form.

Each individual supplies a fixed unit of labor so that given the nature of the problem, an agent consumes all after-tax income: $c_{r,t} = (1 - \tau_{r,t})w_{r,t}$ where $\tau_{r,t}$ is the

---

[13]For lack of a better term, we refer to one region d as the destination and the other o as the origin. Given these are stocks and not flows there is no origin or destination per se. This verbiage will help us talk about the model without refering to arbitrary regions.

[14]Suppose both regions were ex ante identical and private and public consumption are the same in both regions. Then an individual who moves from region o to d will, all else equal, realize a lower level of utility in region d because after the move $N_{d,t} > N_{o,t}$.

tax rate on wages $w_{r,t}$. In practice, $\tau_{r,t}$ is not a single rate but rather is the average tax rate on wages. If the tax system exhibits any progressivity, then $\tau_{r,t}$ will be a function of $w_{r,t}$; thus, a progressive tax system would require estimation to use an average tax rate. To see this, a progressive tax system would be given by the tax function $T(w_{r,t})$ and so consumption would be $c_{r,t} = w_{r,t} - T(w_{r,t}) = (1 - atr_{r,t})w_{r,t}$.

To close the model, assume production in any given region is given by $f(N_{r,t})$ and satisfies the standard properties $f_{N_{r,t}} > 0$ and $f_{N_{r,t}} < 0$; the price of output is normalized to one Euro. With mobility, the equilibrium in the labor market requires the wage rate equal the marginal product of labor $w_{r,t} = f_{N_{r,t}}(N_{r,t})$. Assuming that production is given by $A_r N_{r,t}^\theta \bar{K}_r^\vartheta$ where A_r are fixed productive amenities in the region and $\bar{K}_r$ is the land/capital stock that is fixed in the short run. Then we have in each r that $w_{r,t} = \frac{A_r \bar{K}_r^\vartheta}{N_{r,t}^\theta}$.

A locational equilibrium requires for all $r = o, d$ that $V_{o,t} = V_{d,t} = \overline{V}$. Setting $V_{o,t} = V_{d,t}$, taking logs of the equilibrium wage equation and substituting implies

$$ln(\frac{N_{d,t}}{N_{o,t}}) = \frac{1}{\theta + \frac{\gamma}{\alpha}} ln\left(\frac{1 - \tau_{d,t}}{1 - \tau_{o,t}}\right) + \frac{\pi}{\alpha(\theta + \frac{\gamma}{\alpha})} ln\left(\frac{g_{d,t}}{g_{o,t}}\right) + \zeta_d - \zeta_o \qquad (2)$$

where ζ_o and ζ_d are defined to include the fixed productive amenities, fixed capital resources and consumption amenities across regions defined above. The above equation characterizes the equilibrium in the model. Notice that the endogenous adjustment of wages can be obtained as $\frac{dln(w_{r,t})}{dln(1-\tau_{r,t})} = \frac{dln(N_{r,t})}{dln(1-\tau_{r,t})} \times \frac{dln(w_{r,t})}{dln(N_{r,t})} = -\theta \frac{1}{\theta + \frac{\gamma}{\alpha}}$, which allows for the possibility of less than full capitalization of wages. This expression clearly highlights the role of the congestion cost and the parameter γ.

3.1 Methods

We estimate the pairwise *equilibrium condition* derived in (2). Denote the net-of-tax rate with respect to the average tax rate by $1 - atr_{d,t}$ $[1 - atr_{o,t}]$ in the destination [origin] region. We calculate the average tax rate for a representative taxpayer in the top 1% of the income distribution. We estimate for the working-age population:

$$ln(\frac{N_{d,t}}{N_{o,t}}) = \beta[ln(1 - atr_{d,t}) - ln(1 - atr_{o,t})] + \zeta_d + \zeta_o + \zeta_t + \delta ln\left(\frac{g_{d,t}}{g_{o,t}}\right) + X_{do,t}\phi + \varepsilon_{do,t} \quad (3)$$

13

where β captures the effect of taxes on population stocks, which is a function of the structural parameters in (2).[15] As suggested by theory, we include origin fixed effects that capture amenities (both for households and firms) in the region of origin and destination fixed effects that capture such amenities in the destination region. These fixed effects also capture any time invariant policies of the regions over our sample. Time fixed effects are included in the model to capture any aggregate shocks. As suggested in (2), we control for region-level spending changes across the regions. These spending controls are designed to capture the effect of any changes in services that may make a region more attractive following a tax change. In particular, we control for differentials on basic public services, social protection programs, public programs, general spending, and transportation infrastructure. In some specifications we include a vector $X_{do,t}$, where we control for time varying, region-pair specific shocks including economic shocks, demographic shocks, and regional amenities. These controls help facilitate identification given that the tax changes are not likely random. Given the set of fixed effects and covariates, identification requires that, absent tax changes, region-pair stocks are fixed over time.

Notice (2) leads to a structural interpretation of the estimated coefficient in the locational equilibrium: β is the effect of tax rate changes including their indirect effects through changes in the regional wages, i.e. the effect taking all fixed regional characteristics (amenities) and public services as given except for tax rates *and* wages.[16] Given this interpretation of β, the capitalization into wages does not pose a threat, but other unobservable wage shocks that are correlated with tax changes would be problematic. To deal with this, we control for time-varying regional economic conditions in $X_{do,t}$.

Theory implies to estimate the equilibrium condition using the ratios of populations and taxes rather than the level of the region's own population. In particular, this pairwise ratio is useful because we have a small number of regions, which implies the

---

[15] There were some very small tax differentials that existed prior to 2011. We set these differentials to zero prior to 2011. If we include them, the coefficient is almost unchanged.

[16] If we derived it from a McFadden location choice model without assuming spatial equilibrium through wages, then, (endogenous) wages would need to be controlled for.

number of people in a given region depends on the entire vector of net-of-tax rates in all of the regions. Thus, tax changes in region $r' \neq r$ will have a *non*-zero effect on $N_{r,t}$. This is a purpose of estimating the stocks in pairwise ratios. Of course, this pairwise estimation complicates treatment of the standard errors and interpretation of β. We cluster the standard errors three ways to account for correlation over time within region-pairs and to account for the correlation of errors within both origin and destination by year pairs.

Estimating the location equilibrium condition in ratios influences the interpretation of β. Given we allow the tax rate of a given region to influence the population of other regions, the estimating equation delivers the elasticity of the ratios, $N_{d,t}/N_{o,t}$. Differentiating (3) with respect to the net-of-tax rate in region d, yields:

$$\beta = \frac{\mathrm{d}ln(N_{d,t})}{\mathrm{d}ln(1-atr_{d,t})} - \frac{\mathrm{d}ln(N_{o,t})}{\mathrm{d}ln(1-atr_{d,t})} \equiv \eta - \mu \tag{4}$$

where η is the stock elasticity of the population in region d with respect to its own net-of-tax rate and μ is the cross-elasticity of region o's population with respect to region d's net-of tax rate. Given $\eta > 0$ and $\mu < 0$ are opposite signed, we can conclude that our estimate of β will over-estimate the elasticity of the stock of a given region. However, as the number of regions becomes large, then it is likely $\mu \approx 0$. In our setting, with fifteen regions, we expect μ to be non-zero, but relatively close to zero and thus β acts as a reasonable approximation to the stock elasticity. We verify this is true by estimating the model in levels, which assumes a large number of regions and zero cross-price effects.

Some notes concerning the empirical model are in order. First, in our baseline specifications, we utilize net-of-average tax rates. To construct the average tax rate, following Moretti and Wilson (2017), we simulate taxes in all years and regions for a representative taxpayer in the top 1% holding fixed (across regions and time) income and any inputs to our tax calculator so that variation in the rate is only due to statutory changes. As noted above, the use of the average tax rate is theoretically grounded. However, some of the prior literature has presented results using the top marginal tax rate (Kleven, Landais and Saez 2013; Akcigit, Baslandze and Stantcheva 2016) as a good approximation of the average tax rate. We also present results using the top marginal

tax rate in each region, but note it is not a good approximation to the average tax rate in our setting. The top marginal tax rate will be correlated with the average tax rate because regions that raised the top tax rate were also generally regions that raised rates in lower income brackets (see figure 1). Although our preferred specification uses the (theoretically grounded) average tax rate, the top marginal tax rate may be salient when determining the tax liability in alternative regions; although individuals know the average tax rate in their region, they are unlikely to be able to calculate this across regions.

We estimate a stock model rather than a flow model. First, the stock elasticity is the parameter of interest in the revenue simulations we will conduct. Second, estimation in a flow model raises selection concerns because we do not observe migration between some regions due to our 4% sample and because a flow model would miss international migration and to the Basque country and Navarre; the stock model will not. Appendix A.5 discusses additional justification in detail. The stock model avoids all of these issues and in our opinion, provides a more accurate measure of all tax-induced migration. Limitations of the aggregate analysis resulting from non-random setting of tax rates are addressed in section 4 where we control for time varying region-specific shocks.

3.2 Results

Given the simple panel data setting, we present our baseline results visually. To do this, we regress the stock ratio on the fixed effects and controls and then predict the residuals. We then regress the net of average tax rate variable on the fixed effects and predict the residuals. We then bin the residuals into equally sized bins and fit a line of best fit through these data. Figure 2 shows the baseline results; panel A present the results using the average tax rate while panel B shows the marginal tax rate. We present all results with and without covariates to see if our identifying assumption is reasonable. Because all tax variables are in terms of the net-of-tax rate, when individuals keep more on the Euro in region d relative to region o, they are more likely to move to (or stay in) region d and the stock increases in d relative to o. As the net-of-tax differential increases, we see that it is consistent with $\beta > 0$. The addition of covariates does not meaningfully change the slope, but does reduce the noise. The bottom panel shows the regression using the marginal

tax rate. Although the vertical axis is identical to the upper panel, the horizontal axis is more disperse because differences in top marginal tax rates are larger than average tax rates. Thus, the slope of the line of best fit remains positive but is flatter because a one percent change in the net of (marginal) tax rate will change tax liability less.

We present point estimates of β and standard errors for the aggregate analysis in table 1. If we assume that the cross-elasticity is small, the specification without controls suggests the stock elasticity is approximately 0.92. Our estimates of the elasticity are stable and are not statistically different with or without other covariates. With covariates, it rises just above unity. This estimate of the elasticity is higher than the estimates in Moretti and Wilson (2017), who obtain a stock elasticity of 0.45; Akcigit, Baslandze and Stantcheva (2016) estimate an elasticity of 1 for foreign star scientists. It is larger than Young et al. (2016) who find very small effects. The elasticity with respect to the top marginal net-of-tax rate is 0.65, but the average tax rate is the theoretically relevant one.

A concern with this model is that we may overestimate the stock elasticity resulting from migration because of taxable income responses. In particular, we might worry that in regions lowering their tax rates, taxable income may rise resulting in more people moving into the top 1% of the income distribution in that region. To address this concern, we implement several robustness checks. In table 1 we show the results are robust to controlling for the (endogenous) ratio of taxable income reported by the top 1% in the region pairs. Second, and more preferably, we also adjust the population stocks accounting for movement within the income distribution. Let $\Delta_{r,t}$ be the number of people who move in/out of the top percentile of the income distribution in region r. This number is calculated as the number of people who are in the top 1% this year but were not in it last year minus the number of people who were in the top 1% last year but are not this year. Then we calculate an adjusted stock ratio using $\widetilde{N_{r,t}} = N_{r,t} - \Delta_{r,t}$. We then run all specifications using $ln(\widetilde{N_{d,t}}/\widetilde{N_{o,t}})$ as the dependent variable so that we are exploiting variation in the stock of people in the top 1% adjusted for any yearly churn in the income distribution. After doing this, with covariates, we estimate an elasticity of 0.88, which falls slightly suggesting our prior estimates may capture some taxable income

responses.[17] We also estimate a taxable income elasticity directly by regressing the share of total earned income (working age individuals, excluding pensions and unemployment benefits) in a region earned by the one percent on the net-of-tax rate, region and year fixed effects. This regression estimates a small insignificant elasticity suggesting we are identifying mobility effects in our stock analysis. Appendix A.6 shows the small taxable income response. This is consistent with Rubolino and Waldenström (2017), which estimates a taxable income elasticity for the top percentile in Spain of 0.05. The results of these exercises suggest that we are not identifying taxable income responses. However, to further address this issue, we will subsequently turn to an individual analysis.

We also test for a heterogeneous effect of the tax reform on other lower income groups. In particular, we consider whether these lower income groups respond to the average tax rate for top income taxpayers. In figure 3 panel A we present (using the same scale as the upper panel in the prior figure) results for the top 5% (excluding the top 1%) and the top 10% to 5% of the income distribution. A mildly positive pattern emerges in response to the average tax rate differentials for the top 5% and the slope is declining as income declines. So, what makes lower income households less responsive to the tax differentials? They have smaller tax rate differentials due to progressivity and the top 1% average tax rates may not be relevant for them unless they anticipate income growth, they may care relatively less about tax rates and more about public services because of non-homothetic preferences, or they may have lower moving probabilities because of relatively higher moving costs and less job opportunities. Thus, these results should be interpreted as an analysis of heterogeneity rather than a placebo test given that the differences along these dimensions cannot be ruled out.

---

[17]Assuming the cross-elasticity is small, to interpret the magnitude of the elasticity, the average region has a stock of 565 taxpayers in the top 1% of the income distribution (multiplying by 25 yields population estimates). The average net-of-tax rate implies that a 1% change results in a 0.54 percentage point change. The net-of-tax differential between regions is on average 1.2 percentage points or over 2 times a 1% change in regional taxes. This implies a change in the stock of top taxpayers by 11 people.

We also conduct an exercise in the spirit of a placebo test using the pre-reform period. To do this, we take our measured tax differentials from 2011-2014 and lag them into the pre-reform period. We then match these post-reform tax rates to the years prior to the reform. We show in figure 3 panel B that no significant correlation between the pre-reform stock ratio changes of the top 1% and post-reform tax rate changes exists.

Given identification is based on a large reform, we wish to show a trend break in the patterns following the 2011 reform. For each region pair d and o, if the net-of-tax differential increases in region d relative to region o, we classify that pair as one where the net-of-tax rate increases. Panel A in figure 4 shows the raw averages of the stock ratio for pairs where the region d gets to keep more income (taxes fall) relative to region o. For these observations, the stock ratio increases following the reform and remains higher. To do this formally, we implement an event study approach by estimating

$$ln(\frac{N_{d,t}}{N_{o,t}}) = \overline{ln(\frac{1-atr_d}{1-atr_o})}[\sum_{y=-6}^{-2} \pi_y 1(t-t^*=y) + \sum_{y=0}^{3} \beta_y 1(t-t^*=y)] + \zeta_o + \zeta_d + \zeta_t + X_{od,t}\phi + \varepsilon_{do,t}$$

(5)

where $\overline{ln(\frac{1-atr_d}{1-atr_o})}$ is the average log differential of the net-of-tax rates in the post-reform years. Then, $1(t-t^*=y)$ are indicator variables relating to the time since the reform happened in $t^* = 2011$. As such, π_y show the evolution of the stock ratios prior to the reform and the β_y show the evolution following the reform. Multiplying by $\overline{ln(\frac{1-atr_d}{1-atr_o})}$ captures the intensity of the treatment and allows us to jointly estimate the effect of relative increases and decreases in one specification. Panel B of figure 4 shows no clear pre-trends – if anything, a slight downward trend – but an immediate level increase in the stock ratio following the reform; given the large jump on impact this may suggest tax evasion rather than real moves. The regions that lowered rates allowing residents to keep more, saw an immediate increase in the stock of top income taxpayers. To reduce noise in the post-reform period, the generalized event study design can also be presented as a simple dif-in-dif controlling for trends; the results are given in appendix A.6. The event study is reassuring because population changes do not pre-date the tax reform. Next, we exploit individual data to control for all time varying regional characteristics and have

19

proxies for counterfactual incomes to deal with confounding unobservables.

4   Individual Analysis: Where to Move?

Although the aggregate analysis is appealing in its simplicity, we cannot rule out the possibility of unobservable time-varying region-specific covariates that are correlated with taxes and populations, i.e. the tax rates may not be random. To address this concern, we now use data for an individual tax-payer i who moves in year t. An individual enters the sample if the individual is in the top 1% in year t and relocates *across* regions between t and $t-1$ (and only is in the sample in the year of move).[18] Subsequently, we will refer to an individual that moves regions as a "mover". We denote the alternative residential options (regions within Spain) as j in our model. In particular, we focus on working-age movers in our estimating sample from 2006 to 2014; the vast majority of individuals move only once over the course of our sample but some individuals move multiple times in which case they appear in our data multiple times. For most individuals there will only be one time observation (but still J region observations). A "move" is a time-specific move which is indexed by (i, t) and the choice set for each move is indexed by j. One justification for focusing on movers is that tax rates are likely a function of all individuals' location decisions. Because movers are a relatively small share of the population, it is likely that the equilibrium tax rates selected following the fiscal decentralization are driven by the large share of the stayers – reducing endogeneity concerns (Schmidheiny 2006; Brülhart, Bucovetsky and Schmidheiny 2015). Also, as noted in Schmidheiny (2006), "Households do not daily decide upon their place of residence. There are specific moments in any individual's life when the decision about where to live becomes urgent.... Limiting the analysis to moving households therefore eliminates the bias when including households that stay in a per se sub-optimal location because of high monetary and psychological costs of moving. However, the limitation to moving households introduces a potential

---

[18]Moves within a region are not included in our sample; we do not include these moves because some regions are composed of a small number of provinces and we cannot observe moves within the same province unless a municipality code is available. We only observe municipality codes for individuals living in sufficiently large cities.

selection bias when the unobserved individual factors that trigger the decision to move are correlated with the unobserved individual taste for certain locations."

Focusing on movers leaves us with a sample of 893 moves in the top 1%, of which approximately 330 are in the post-reform period, resulting in 13,395 move-region observations in our dataset. Note that the aggregate stock analysis allowed "stayers" and "moving" residents to respond to taxes; in this section we study the effect of taxes conditional on moving regions. In appendix A.7, we also present results for the full sample of movers and stayers; the appendix also tests for differences in the characteristics of movers and stayers and finds no significant difference in average incomes prior to the move.

The dependent variable $d_{i,t,j}$ is coded one for the chosen region of residence for a given move (i,t) and zero for all other region that are not selected. In its most complex form, we estimate the following linear probability model where the *person-specific* tax rates from our tax calculator are denoted by $\tau_{i,t,j}$:

$$d_{i,t,j} = \beta ln(1 - \tau_{i,t,j}) + \alpha_{i,t} + \iota_{t,j} + \zeta_j \mathbf{x}_{i,t} + \gamma \mathbf{z}_{i,t,j} + \varepsilon_{i,t,j}. \tag{6}$$

Because we use moves pre- and post-reform, and because all taxes for a given individual are the same in the pre-reform period, the pre-reform period helps us to pin down other explanatory variables.[19] Our model contains individual move dummies denoted $\alpha_{i,t}$, alternative region by year dummies $\iota_{t,j}$, individual characteristics interacted with region dummies $\zeta_j \mathbf{x}_{i,t}$, and move specific covariates that vary across the choice regions $\mathbf{z}_{i,t,j}$. We will discuss each of these components in turn.

Move dummies. We estimate our model using a linear model for which we wish to highlight two important properties. First, predicted choice probabilities over all regions will add up to 1 for an individual i moving in year t. Second, an increase in the tax rate in one region does theoretically increase the probability of choosing this region while it decreases the probability of choosing any other region. Along both of these points, the

---

[19]In all that follows, the very small tax differentials in the pre-reform period are not utilized. Including these differentials in the regression yields almost identical results.

presence of $\alpha_{i,t}$ for each move is critical because it forces the predicted probabilities to sum to one and therefore an increase in one region will lower the probability of the alternative regions; this fixed effect adjusts the predicted probability from all other covariates by capturing the average deviation from the average probability.[20] However, the predicted probability of a mover selecting any one region need not be bounded between zero and one; even with this, the predicted probabilities across the choice regions will sum to one. The inclusion of $\alpha_{i,t}$ also forces identification of our parameter of interest to come from variation in tax rates across regions for a given move. In this way, we exploit the income tax differential across regions for a given tax payer which relocated. In the subsequent paragraphs, we discuss each of the components of this regression.[21]

Taxes. The theoretically appropriate tax rate, the "true-ATR", would be the actual effective average tax rate facing a given individual in a given region, which is a function of the individual's (possibly counterfactual) income in that region and the tax system in that region. Because this counterfactual is not observable to us, we use a "simulated-ATR" measure, which is the average tax rate for a given income level in a given region; this simulated tax rate is a function of the (assumed to be constant) income level across regions and the tax schedule in that region. In the baseline specification, the

---

[20]For ease of notation, we prove this for a single covariate denoted by $x_{i,t,j}$. The sum of the predicted probabilities for a move (i,t) is given by $\sum_j(\widehat{\beta}x_{i,t,j} + \widehat{\alpha_{i,t}}) = \widehat{\beta} \cdot J \cdot \overline{x_{i,t}} + J \cdot \widehat{\alpha_{i,t}} = J \cdot [\widehat{\beta}\overline{x_{i,t}} + \widehat{\alpha_{i,t}}]$ where the bar denotes an average over the j's. Given we have J alternatives and only one region can be chosen: $\overline{d_{i,t}} = \frac{1}{J}$. The linear model implies that the estimated fixed effects, $\widehat{\alpha_{i,t}}$, are given by $\widehat{\alpha_{i,t}} = \overline{d_{i,t}} - \widehat{\beta}\overline{x_{i,t}} \Rightarrow \overline{d_{i,t}} = \widehat{\beta}\overline{x_{i,t}} + \widehat{\alpha_{i,t}}$. Plugging in the fixed effects into the sum of probabilities and using $\overline{d_{i,t}} = \frac{1}{J}$, shows that $\sum_j(\widehat{\beta}x_{i,t,j} + \widehat{\alpha_{i,t}}) = J \cdot \overline{d_{i,t}} = J \cdot \frac{1}{J} = 1$. This, then, implies that an increase in the probability of selecting one region must lower the probability of the alternative regions.

[21]The linear model assumes that the effects are constant. Very small regions with very low baseline probabilities, will experience the same effect in percentage points as large regions with high baseline probabilities. The inclusion of region by year fixed effects controls for characteristics like jurisdiction size that influence the baseline probabilities.

22

simulated net-of-tax rate is *person-specific* (not for a representative taxpayer). Specifically, because counterfactual wages are not observed to us, we use our tax calculator to construct the individual's average tax rate by assuming her income is constant across the regions. However, in practice and as shown by our theory which allows for wage capitalization to arise in spatial equilibrium (recall (2)), income may differ across the regions. In particular, a given individual is more likely to move to a high-income region all else equal. Given that taxes are progressive, by assuming that income is constant we will overestimate counterfactual wages (because we observe them in the selected, likely higher wage, region) and therefore overestimate counterfactual tax rates. This raises measurement error concerns, because the average simulated tax rates depend on the assumed to be constant income across regions and not the true counterfactual wages that may differ across regions. However, as noted in Kleven, Landais and Saez (2013), the (individual's) marginal tax rate, the tax rate facing a given individual in a given region, proxies for the exogenous component of the ATR because it is independent of earnings and allows us to implement an IV strategy discussed below.[22] This also has the advantage of reducing measurement error in the average tax rate that might result from elements of the tax code not captured by our tax calculator. The variation in the marginal tax rate, across regions for the top 1% relative to other income groups is verified by looking at the within mover variation in appendix A.7. The variation across different regions a taxpayer could choose increases substantially from 2011 onward, and is more pronounced for the top 1%.

Wage controls and sorting. Equilibrium wage differentials across regions may also be important to the choice of region. Although we assume wages are equal across regions to calculate taxes, wages may differ across regions and influence migration decisions irrespective of taxes. To control for unobservable counterfactual wages, we include location specific dummies interacted with characteristics of the mover (age, age squared, male, and education). Denote the vector of characteristics interacted with region specific

---

[22]We cannot use the top marginal tax rate as an approximation for ATR because most individuals do not have income well into the top bracket. Thus, we use the *mover's* marginal tax rate based on the bracket for her income as an instrument.

dummies $\zeta_j \mathbf{x}_{i,t}$. This allows the returns to education and the skill premium to vary by region; by allowing observables to vary by region, these observable characteristics are used to account for unobservable counterfactual wages. Although motivated by wage differentials, this parameterization also can be interpreted as capturing any sorting of specific types of individuals to particular regions. For example, if high-educated or older individuals have preferences for locating in a region, the specification will capture this.

Public services and regional shocks. Public services across regions matter. The inclusion of region by year dummies ($\iota_{j,t}$) captures any time varying policies – such as changes in public services – that are constant across all individuals in the top 1%.[23] In general, however, we note that tax increases on the rich are not likely to change public services for the rich as these taxpayers are net payers into the tax system. Thus, in addition to accounting for regional policies, $\iota_{j,t}$ also account for time-varying amenities or economic shocks in the alternative regions that affect individuals in the top 1%. Unlike in the aggregate analysis, inclusion of these dummies is possible because we exploit mover-specific income to calculate tax rates rather than income for a representative taxpayer.

Other controls. Finally, moving costs between regions, which could be thought of as higher γ in our theory, also matter. To capture these moving costs, we include in a vector $\mathbf{z}_{i,t,j}$ a dummy variable that equals one if the region is the place of birth for the individual. We also include a dummy variable for the region of the principle workplace of the individual and a dummy variable that equals one if the individual had their first job in that region. Following a standard gravity model of migration, we also include the log of distance between the region of prior residence and each of the alternative regions. This captures the fact that nearby regions have lower moving costs because they allow individuals to maintain their social and family network. Acknowledging the region of residence prior to moving plays a special role as it cannot be selected by a mover, we also include a dummy variable that equals one if the individual previously lived in the region. Note that these covariates can enter the regression even though they are time-invariant

---

[23]These also capture the possible effect of wealth taxes across regions and time.

because they vary over the alternative j regions for a given individual i in year t.[24]

One important issues is the treatment of standard errors in this model. In particular, the dummy variables in (6) are related over the different regions j as only one option can be chosen for a given move. Further, tax policy is set by the region. First, we cluster over moves to resolve the first issue. Second, we cluster over destination region by year clusters to account for tax law in a region influencing all movers. In particular, it allows for correlation in errors across all movers for a particular region-year that may result, perhaps, from the common elements of the tax code that affect all individuals. We show the results for other treatments of standard errors (see appendix A.7).

Two additional selection issues arise. In terms of identification, selection concerns may arise because the sample excludes two Spanish regions because they operate autonomous fiscal systems; we treat these two regions as being "international" so that any concerns about moves to these regions are similar to concerns about moves abroad (these two regions also have their own languages, so moving to these regions may involve similar costs of moving abroad). This issue is common to prior studies of domestic migration that do not observe migration abroad. An additional concern arises because individuals appear in our estimating sample only when they move and thus the time dimension is unbalanced. If someone exits the top 1% and then moves, they would not appear in our sample. This type of exit from the sample would only be a concern if the individual exits the top 1% for reasons due to the income tax. In particular, this would result in us overestimating [underestimating] our effects if an individual that reduced their taxable income (enough to drop out of our sample) was also an individual that then elected to stay or move to a high-tax [low-tax] region. However, as discussed in the aggregate analysis, appendix A.6 shows the taxable income response following the reform was small.

4.1 Baseline Results

Table 2 shows estimation of (6) using the average tax rate for each individual. Column (1) shows the results including $\alpha_{i,t}$ fixed effects and region by year fixed effects. The coefficient, 0.588, is the expected sign: a higher net-of-tax rate implies a higher probability

---

[24]For example, the place of birth dummy is one in the region of birth and zero in others.

of migrating to that region because the individual can keep more of what is earned. In terms of the magnitude, this coefficient implies that a one percent increase in the net-of-tax rate raises the probability of choosing a given destination by 0.59 percentage points. This represents a substantial increase in the probability of moving to a region, which if random would be 1/15. Subsequent columns of the table add various controls discussed above. This helps with the precision of our estimates, but the coefficient on the net-of-tax rate are stable – if anything, slightly increasing after accounting for these controls. Given that these variables richly control for counterfactual wage differences across regions and any possible sorting, this is very reassuring. Critical to our analysis is that column (2) containing no wage controls and column (7) containing a complete set of wage controls ($\mathbf{x}_{i,t} \times \iota_j$) are not statistically different. The covariate adjusted coefficient is 0.90.[25]

We focused on the sample of movers across regions where the role of taxes is salient. We estimate a model using both stayers and movers in the top 1% and find a smaller coefficient of 0.08. This is consistent with Akcigit, Baslandze and Stantcheva (2016) who estimate elasticities of domestic scientists around 0.03, but near 1 for foreigners.

4.2 IV Approach

Kleven, Landais and Saez (2013) adopt a grouping estimator a to construct the average tax rate by year $\times$ country $\times$ foreign $\times$ quality. In their specification, quality serves a similar role as the, assumed to be constant, income level in our tax simulator. But, in our setting, by assuming that income is constant we may overestimate counterfactual wages (because we use income from the selected, likely higher wage region) and therefore overestimate counterfactual tax rates. This raises measurement error concerns because the simulated-ATR is a noisy measure of true-ATR in the non-chosen regions. To resolve this issue, we instrument for the average tax rate with the mover-specific marginal tax rate, which is independent of earnings conditional on being in the same tax bracket. Given most individuals in the top 1% only have a fraction of income taxed at their marginal tax rate, the relationship with the average rate is not as close to one as for superstars. For this reason, the MTR cannot simply be used as a proxy for the true-ATR.

---

[25]Appendix A.8 shows that our results are robust to using a nonlinear model.

Table 3 shows results using this IV strategy (appendix A.7 shows the reduced form with the marginal rate). As expected, the IV reduces attenuation bias concerns; the results are larger in magnitude. The first stage coefficient is the expected sign, and is less than one given that a one point increase in the marginal rate raises the average rate by less than one point. The instrument is strong given changes in marginal rates at the top of the distribution are generally correlated with the pattern of changes lower in the income distribution. After instrumenting, a one percent increase in the net-of-tax rate increases the probability of moving to a region by 1.452 percentage points in our baseline specification and 1.731 percentage points in our most comprehensive estimating equation. The intuition is clear: absent an instrument our tax sensitivity is estimated off of simulated net-of-tax differentials that are noisy measures of the counterfactual tax rate because we assumed that income is constant across regions, which due to attenuation bias results in underestimating the true coefficient of interest. Our instrument resolves this issue. These estimates suggest that the 0.75 percentage point differential in the average tax rate at mean income between Madrid and Cataluña in 2013 increases the probability of moving to Madrid by 2.25 percentage points. The effect of Madrid's 2014 tax cut of 0.38 percentage points further increases the probability of moving to Madrid by another 1.14 percentage points. Effects are likely larger at higher income levels in the top 1% because the average tax rate differential at 300,000 Euros is over two percentage points.

We also estimate our model using the top 2% and top 3% of taxpayers. The effects fall off substantially and become insignificant after the top 1% of the income distribution. This may be due to the variation in tax rates in figure 1, which shows the divergence of tax rates is strongest – and most likely to overcome migration costs – for the top 1%. The divergence of tax rates across regions is not large outside of the top 1% and thus these "lower" income individuals are not much influenced by the reform unless they expect to see large income increases in the future. To formally test if the semi-elasticity varied across the income distribution (Lehmann, Simula and Trannoy 2014), we would need equally salient tax changes for lower income groups exceeding their moving costs.

With respect to our instrument, the marginal tax rate may not be independent

of earnings if individuals change brackets. Thus, the best instrument would account for the individual being in the same tax bracket for all regions. To analyze the importance of this, we remove any individuals from the analysis that are within $\chi\%$ above or below all tax bracket thresholds for all regions. We show results for $\chi = \{1, 2.5, 5\}$. Thus, for $\chi = 1$, we remove anyone who is 3,000 Euro above or below the top tax bracket of 300,000 Euro, 1,750 Euro above or below the 175,000 Euro threshold, etc. Appendix A.7 shows that for $\chi = 1$ or $\chi = 2.5$, the results remain similar: a one percent change in the net-of-tax rate changes the probability of moving to a given region by approximately 1.8 percentage points. For $\chi = 5$, the results increase because the large cutoff removes a substantial fraction of "lower" income individual in the top percentile.

Given the IV only resolves measurement error concerns relating to counterfactual wages, one concern may be that the variation in tax rates across regions is not random. We conduct an exercise in the spirit of a placebo test to verify that post-reform tax rates are not correlated with unobservable characteristics that predict migration. We use the post-reform data to construct a placebo measure of tax rates in the pre-period. To do this, for each individual i and region alternative j in the Social Security data, we construct the mean tax rate in the post reform data (2011 to 2014). We use this tax rate as an explanatory variable to explain pre-reform moves to see if these post-reform tax rate differentials have any effect on the decisions pre-reform.

In table 4 column (3), we first show results using post-reform migration to show that taking the mean tax rate for each individual yields a very similar coefficient (2.05 versus 1.73 previously). In column (4), we restrict the sample to individuals that moved in the pre-reform period between 2005 and 2010 and implement the IV approach for these individuals using the placebo tax rates. The coefficient falls to 0.09. Post-reform tax rates are not correlated with unobservable factors that may have influenced pre-reform migration. This suggests to us the post-reform tax rates were not set in a way that was correlated with the observable migration patterns in the pre-reform period.

4.3 Heterogeneity

Although we have identified significant location choice effects, we have yet to determine if

the location choices reflect real moves or simply tax evasion by misreporting the primary residence (perhaps to a second home). In order to shed light on this, we explore whether the tax changes have heterogeneous effects across different types of people by interacting the tax rates with indicator variables for various groups. Appendix A.7 presents the results by age, whether the individual has children, gender and education.

In general, we do not find statistically significant differences across groups. This could be a result of characteristics not affecting the probability of where to move, but rather the ability to move per se. However, characteristics may also matter for where to move. One category that does have economically meaningful differences in the point estimates relates to education status; individuals with a higher education have a stronger influence of taxes on their location choices. This is consistent with these higher educated individuals having less job constraints and thus having a larger feasible set of regions to choose from. High educated households might also be more likely to seek the advice of a investment or tax consultant that might give advice on low-tax residential location.

In appendix A.7 we show results by job characteristics.We wish to see if the individual moves are driven by employment shocks or changes in the locations of firms. We focus on individuals that had a "non-voluntary stop" of their main contract in the previous year or in the year of the move, individuals where the headquarter of the firm of their main contract moved, and individuals that changed their contract. We find similar point estimates across all of the categories, however, one category is usually insignificant. Given that the estimates are not statistically different from each other, we conclude that the increases in the probability of moving to a region are not driven by firm-side responses.

4.4   Occupation and Industry

With the exception of Young et al. (2016), who focus on millionaires, the prior literature has been unable to answer the question whether policymakers can take the estimates derived for star scientists and athletes and apply these elasticities to the top of the income distribution more generally. The Spanish data we have access to is unique in that occupation and industry are reported in the data; this is not information that would be easily available when using U.S. tax return data. We test the generalizability of focusing

on star scientists and athletes using these data. Although the number of athletes and star scientists are too small to focus on these groups specifically (tables A.19 and A.18), we can aggregate to broader occupation/industry categories that allow us to study the heterogeneity. This section helps to inform the recent policy debate on the efficiency of tax schemes for top earners in specific occupations. Several OECD countries have preferential tax schemes for foreigners in high-income occupations. By focusing on heterogeneity by occupation and industry, we can shed light on the efficiency of these schemes.

To determine if some industries or occupations are more responsive, we estimate (6) with an interaction of $ln(1 - \tau_{i,t,j})$ with dummy variables for occupation or industry categories. We show the results in figure 5 with precise point estimates in appendix A.9. When looking at the result for occupation, we identify the strongest effect for self-employed occupations; they are twice as large as all other occupations.[26] This is consistent with these individuals being able to change their residence because their work location may also be flexible. Most of the other three broad categories have smaller degrees of responsiveness to each other. Although we have grouped the occupations based on skill, the occupation categories do not follow a natural hierarchy. Thus, we switch to industry classifications, where we use the one digit industry groupings in the data.

Figure 5 shows substantial heterogeneity by industry. We find the largest (and statistically significant) effects in the health, real estate, information, financial, and professional/scientific industries. Even within these groups, the effects in the health industry are three times larger than the financial industry. This heterogeneity may result, for example, from lower moving costs because of ease in relocating jobs. To compare this to the prior literature, athletes would fall under the category of arts and entertainment, while scientists could be under health or professional/scientific, which exhibit a very high degree of tax-induced mobility to lower tax regions. Our general takeaway from these results is that the responsiveness to taxes varies substantially depending on occupation

---

[26]This result should be interpreted with the caveats discussed previously. Self employed are only included if they have a relationship with a registered firm. We verify in appendix A.9 that the self-employed have a majority of their income from non-labor income.

and industry. Thus, these results provide a cautionary tale; the prior literature focusing on star scientists and athletes may not generalize to other occupations/industries.

5   Interpretation and Revenue Implications

An important policy question is how this reform affects tax revenue. For simplicity, consider a nonlinear tax schedule $T(y^i)$ where individual i earns income y^i, which is endogenous to the tax system because of taxable income responses. To proceed, assume that the top tax rate above the income bracket $\bar{y}$ is linear and given by $\bar{\tau}$. Define N, which is a function of taxes because of potential migration responses, as the stock of individuals above $\bar{y}$. To characterize the revenue maximizing top tax rate, follow the approach of Piketty and Saez (2013), maintaining all of their assumptions. Holding fixed taxes below $\bar{y}$ and perturbing $\bar{\tau}$ by $d\bar{\tau}$ and aggregating across individuals – letting y denote the average income in the top bracket – yields the total effect of top tax rate changes on revenue. The tax change will have three effects: a mechanical effect as a result of the change in the tax rate, an effect resulting from migration, and taxable income responses. Totally differentiating tax revenue, the change in tax revenue R can be decomposed into:

$$dR = \underbrace{[N(y - \bar{y})]\,d\bar{\tau}}_{\text{mechanical}} \underbrace{-\epsilon a \left[N(y - \bar{y}) \frac{\bar{\tau}}{1 - \bar{\tau}} \right] d\bar{\tau}}_{\text{taxable income}} \underbrace{-\eta N(y - \bar{y}) \left[\frac{T(y)}{y - T(y)} \right] d\bar{\tau}}_{\text{mobility}} \quad (7)$$

where $\eta = \frac{dN}{N} \frac{(y - T(y))/y}{d[(y - T(y))]/y}$ is the stock elasticity (with respect to the average tax rate), $\epsilon = \frac{dy}{y} \frac{1 - \bar{\tau}}{d(1 - \bar{\tau})}$ is the elasticity of taxable income and $a = \frac{y}{y - \bar{y}}$ is the Pareto parameter. This is a partial equilibrium analysis: it abstracts from spillovers from the presence of top earners to the income of, and thus revenues from, lower taxpayers; it ignores any other revenue effects obtained through other taxing instruments; and it assumes no horizontal or vertical fiscal externalities. One important limitation is that the elasticity of taxable income is calibrated rather than estimated. Setting (7) to zero and solving for ϵ, we can obtain the critical value of the elasticity of taxable income necessary for a government to maximize revenue:

$$\widetilde{\epsilon} = \frac{1 - \eta \left(\frac{T(y)}{y - T(y)} \right)}{a \left(\frac{\bar{\tau}}{1 - \bar{\tau}} \right)}. \quad (8)$$

We estimate the revenue effects holding fixed the central government's tax rate at its 2014 level. We ask the question: at 2014 regional tax rates, how much does revenue change relative to if the region had simply mimicked the central government's tax rate on its tax base in 2014? Thus, for regions that raised their tax rate relative to the central government's tax rates, the mechanical effect is positive, but both the taxable income and mobility effect will be negative. For regions that decreased their tax rates relative to the central government tax rate, the effects will be opposite in sign. The Pareto parameter is estimated using income data and the mobility elasticity using aggregate analysis. We estimate confidence bands using the parametric bootstrap. We assume the ETI is 0.15, which is slightly lower than the mid-point in the literature because the part of the tax base we analyze excludes capital income. Appendix A.10 details the simulation assumptions.

Figure 6 panel A shows the change in revenue resulting from the regional tax rates as a percent of total personal income tax revenue from all residents. The precise revenue effects with confidence bands are given in appendix A.10. The figure shows that in all circumstances, the mechanical effect of higher or lower tax rates is always the same sign as the total effect on tax revenue after accounting for all behavioral effects. This means that governments are on the left side of the Laffer curve: raising tax rates relative to the central government rate increases tax revenue in the regions. Madrid lowering its tax rate relative to the central tax rate corresponds to a decline in tax revenue from the top 1%. This lower tax rate results in revenue falling by 50 million Euro. However, taxable income only rises by 4 million Euro and the additional new high taxpayers contribute 9 million Euro more. Thus, the behavioral effect from migration is only 18% of the mechanical effect. The total change in revenue from the reform lowers tax revenue by approximately 0.42 percent of total personal income tax revenues in the region of Madrid. Although the stock elasticity is "large," if taxable income responses are small, progressive taxation remains a feasible means of raising revenue in the short-run. Of course, the calculation of the revenue effects has all the partial equilibrium caveats discussed above.

One limitation of this is that the elasticity of taxable income is calibrated rather than estimated. Figure 6 panel B shows the value of the ETI that is necessary for the

region's deviation from the national tax rate to result in the taxable income and mobility response exactly offsetting the mechanical effect. The figure indicates that this ETI must be between 1.02 and 1.34. These values are well outside the range of the best available estimates of this elasticity, which range from 0.12 to 0.40 (Saez, Slemrod and Giertz 2012), and suggest that our revenue conclusions are not driven by the calibration.

6 Conclusion

We find that income tax changes result in a stock elasticity less than unity. In revenue terms, the behavioral effects induced by tax rate changes have a smaller effect on tax revenue than the mechanical effect resulting from a higher or lower tax rate. Although the migration response is significant, the taxable income responses are likely small meaning that the elasticity of the tax base is well below unit elastic. Although the recent economics literature has seen an increase in research on migration, we are the first study to use population-representative administrative data in a country where taxes are purely residence based. Our revenue simulations suggest that changes in the stock of top taxpayers has minimal tax base effects. Thus, our results, at least in the short run, are consistent with Epple and Romer (1991) who show that local redistribution is feasible with migration, but in contradiction to Feldstein and Wrobel (1998) who show the opposite. In the long run, mobility is likely to rise given demographic shifts and technological innovations, which may impose added constraints on redistributive policy.

References

Agrawal, David R. and William H. Hoyt. 2018. "Commuting and Taxes: Theory, Empirics, and Welfare Implications." *The Economic Journal* forthcoming.

Akcigit, Ufuk, Salome Baslandze and Stefanie Stantcheva. 2016. "Taxation and the International Mobility of Inventors." *American Economic Review* 106(10):2930–2981.

Bakija, Jon and Joel Slemrod. 2004. "Do the Rich Flee from High State Taxes? Evidence from Federal Estate Tax Returns." NBER Working Paper 10645.

Bonhomme, Stéphane and Laura Hospido. 2017. "The Cycle of Earnings Inequality: Evidence from Spanish Social Security Data." *Economic Journal* 127(603):1244–1278.

Bosch, Núria. 2010. "The Reform of Regional Government Finances in Spain." *IEB's World Report on Fiscal Federalism* .

Brülhart, Marius, Sam Bucovetsky and Kurt Schmidheiny. 2015. "Taxes in Cities: Interdependence, Asymmetry, and Agglomeration." *Handbook of Regional and Urban Economics* .

De la Roca, Jorge and Diego Puga. 2017. "Learning by Working in Big Cities." *Review of Economic Studies* 84:106–142.

Durán, José María and Alejandro Esteller. 2005. Descentralización Fiscal y Política Tributaria de las CCAA: Un Primera Evaluación a Través de los Tipos Impositivos Efectivos en el IRPF. In *La financiación de las comunidades autónomas: Políticas tributarias y solidaridad interterritorial*, ed. Núria Bosch and José María Durán.

Epple, Dennis and Thomas Romer. 1991. "Mobility and Redistribution." *Journal of Political Economy* 99(4):828–858.

Feldstein, Martin. 1999. "Tax Avoidance and the Deadweight Loss of the Income Tax." *Review of Economics and Statistics* 81(4):674–680.

Feldstein, Martin and Marian Vaillant Wrobel. 1998. "Can State Taxes Redistribute Income?" *Journal of Public Economics* 68:369–396.

Foremny, Dirk, Jordi Jofre-Monseny and Albert Solé-Ollé. 2017. "Ghost citizens: Using Notches to Identify Manipulation of Population-based Grants." *Journal of Public Economics* 154:49–66.

Kleven, Henrik J., Camille Landais, Emmanuel Saez and Esben A. Schultz. 2014. "Migration, and Wage Effects of Taxing Top Earners: Evidence from the Foreigners' Tax Scheme in Denmark." *Quarterly Journal of Economics* 129:333–378.

Kleven, Henrik Jacobsen, 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.

Lehmann, Etienne, Laurent Simula and Alain Trannoy. 2014. "Tax Me If You Can! Optimal Nonlinear Income Tax Between Competing Governments." *Quarterly Journal of Economics* 129(4):1995–2030.

34

Martínez, Isabel Z. 2016. "Beggar-Thy-Neighbour Tax Cuts: Mobility after a Local Income and Wealth Tax Reform in Switzerland." Working Paper.

Milligan, Kevin and Michael Smart. 2016. "An Estimable Model of Income Redistribution in a Federation: Musgrave Meets Oates." UBC Working Paper.

Mirrlees, James A. 1982. "Migration and Optimal Income Taxes." *Journal of Public Economics* 18(3):319–341.

Moretti, Enrico and Daniel Wilson. 2017. "The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists." *American Economic Review* 107(7):1859–1903.

Piketty, Thomas and Emmanuel Saez. 2013. "Optimal Labor Income Taxation." *Handbook of Public Economics* 5:391–474.

Rubolino, Enrico and Daniel Waldenström. 2017. "Trends and Gradients in Top Tax Elasticities: Cross-country Evidence, 1900-2014." CEPR Discussion Paper 11935.

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.

Schmidheiny, Kurt. 2006. "Income Segregation and Local Progressive Taxation: Empirical Evidence from Switzerland." *Journal of Public Economics* 90:429–458.

Schmidheiny, Kurt and Michaela Slotwinski. 2018. "Behavioral Responses to Local Tax Rates: Quasi-Experimental Evidence from a Foreigners' Tax Scheme in Switzerland." *Journal of Public Economics* .

Wildasin, David E. 2006. "Global Competition for Mobile Resources: Implications for Equity, Efficiency, and Political Economy." *CESifo Economic Studies* 52(1):61–110.

Wildasin, David E. 2009. Public Finance in an Era of Global Demographic Change: Fertility Busts, Migration Booms, and Public Policy. In *Skilled Immigration Today: Prospects, Problems, and Policies*, ed. Jagdish Bhagwati and Gordon Hanson.

Young, Cristobal and Charles Varner. 2011. "Millionaire Migration and State Taxation of Top Incomes: Evidence from a Natural Experiment." *National Tax Journal* 64:255–284.

Young, Cristobal, Charles Varner, Ithai Lurie and Rich Prisinzano. 2016. "Millionaire

Migration and the Demography of the Elite: Implications for American Tax Policy."
American Sociological Review 81(3):421–446.

Figure 1: Tax Rate Changes Relative to Central Government Tax Rate (2011 & 2014)
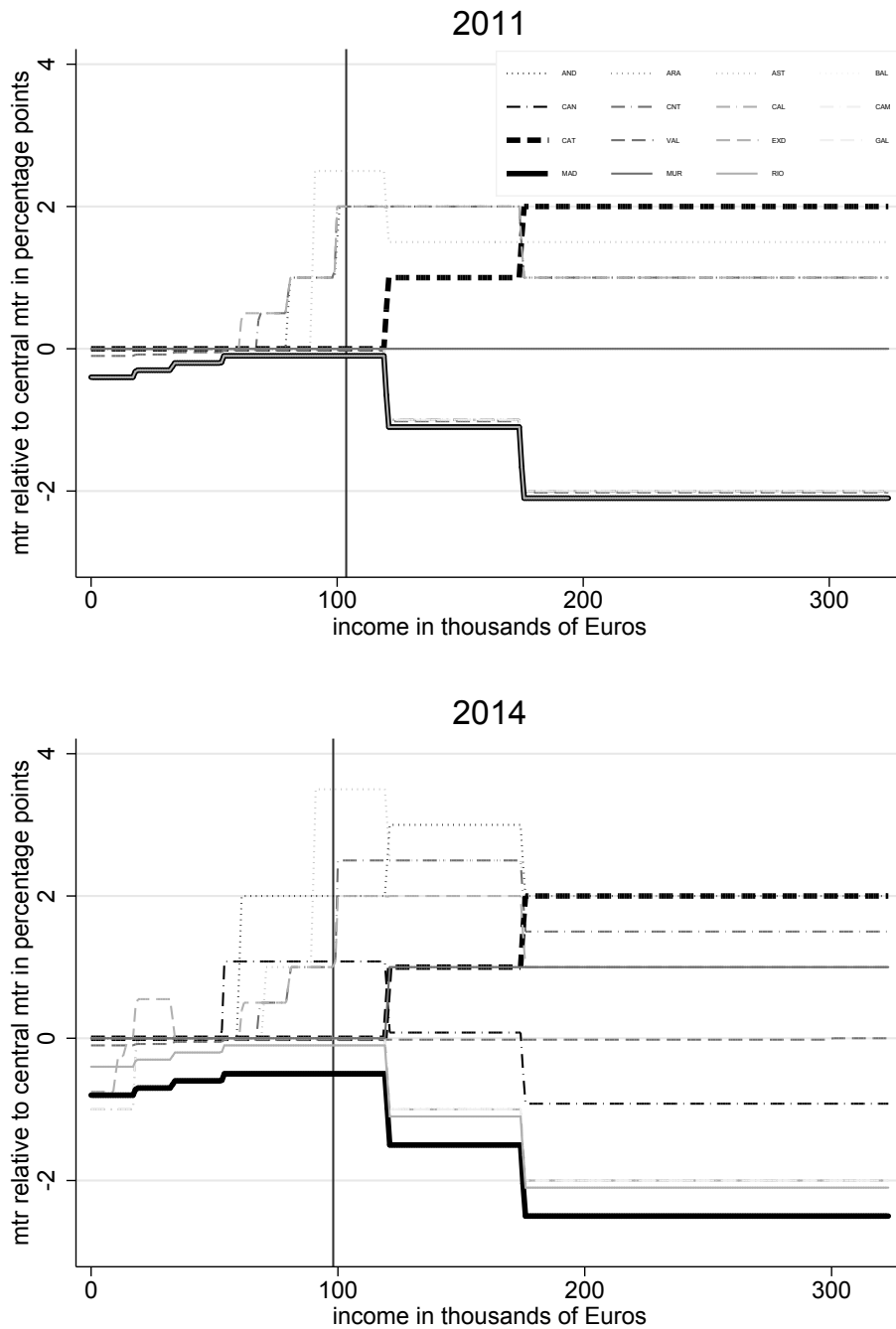
The graph shows the deviation in the marginal tax rates for each region relative to what the tax rate would be if the region just copied the central government's tax rate in that same year. The first figure shows the deviations in 2011 immediately following the reform and the second figure shows the deviations in 2014. The vertical line shows the income cutoff for the top 1% in each year.

Figure 2: Effect of Taxes on the Stock Ratio

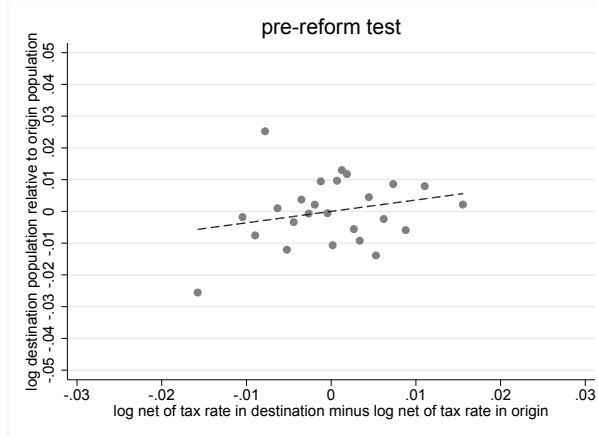PANEL A: ATR RESULTS

PANEL B: MTR RESULTS
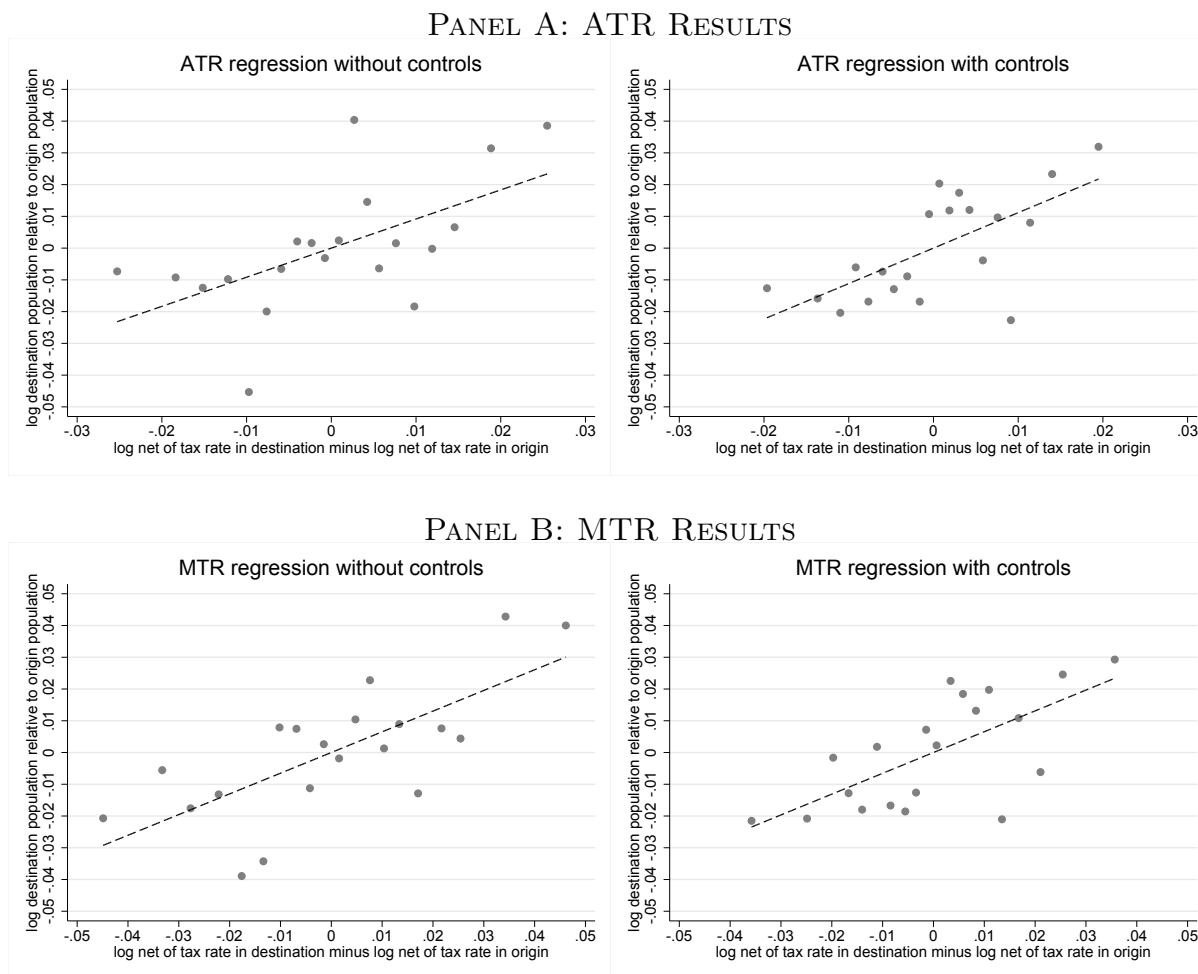
Panel A shows a visual representation of the regression of the stock ratios on the net-of-average-tax differentials, while panel B presents the regression of the stock ratios on the net-of-marginal-tax differentials. The left figure of each panel excludes controls while the right figure of each panel includes controls in these regressions. To construct the figures, we regress the log of the stock ratio in region d (called "destination" for convenience) relative to region o (called "origin" for convenience) on the fixed effects, government spending controls, and controls to obtain the residuals. The tax variable is residualized in a similar manner. The figure plots the residuals in equal sized bins and shows the line of best fit. The provides both a non-parametric and linear estimate of the conditional expectation function. The theoretically expected effect is positive as keeping more money in the destination state means tax rates are lower, which implies more migration to the destination, increasing the stock.

Figure 3: Heterogeneity Using Lower Income Groups and Placebo Test

PANEL A: HETEROGENEITY BY INCOME GROUP
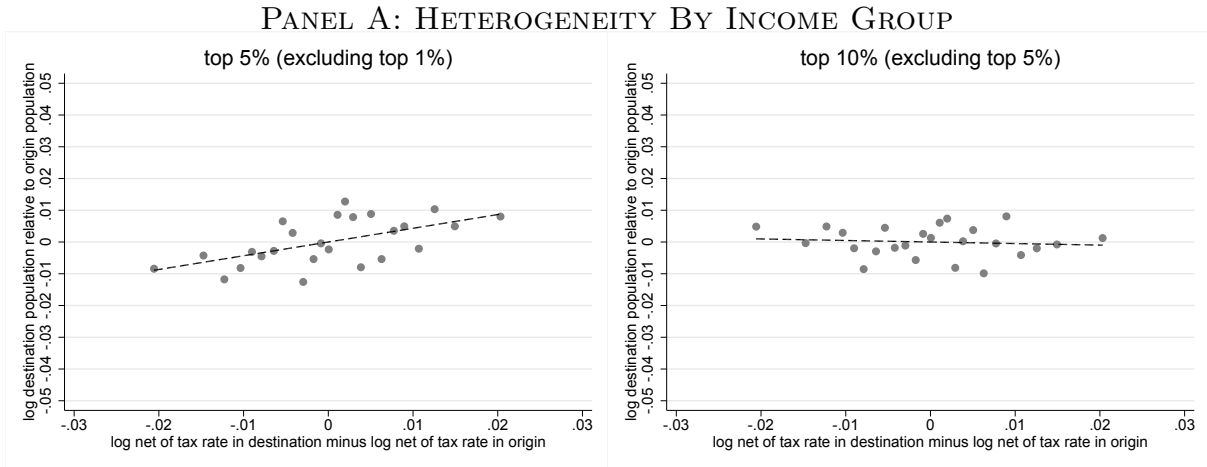
PANEL B: EFFECT OF POST-REFORM TAXES ON PRE-REFORM STOCKS

The figures in panel A are comparable to the covariate adjusted ATR results of figure 2 except they focus on lower percentiles of the income distribution. To remain comparable, we use the net-of-average-tax variable for the top 1% to explain these lower-income stock ratios. The only difference from the prior figure is that we use the stock ratios of individuals outside the top one percent (the top 5% excluding the top 1% and the top 10% to 5%). Panel B is comparable to the covariate adjusted ATR results of figure 2 except for the use of only pre-reform moves and the time lag of tax rates. We take post-reform tax rates and lag them into the pre-reform period to see if they predict pre-reform population stocks. In panel B, the theoretically expected effect is zero given this is a placebo test. All figures use the same vertical/horizontal scale as panel A in figure 2.

Figure 4: Event Study of Stock Ratio Using the 2011 Reform

PANEL A: RAW AVERAGES

PANEL B: FORMAL EVENT STUDY
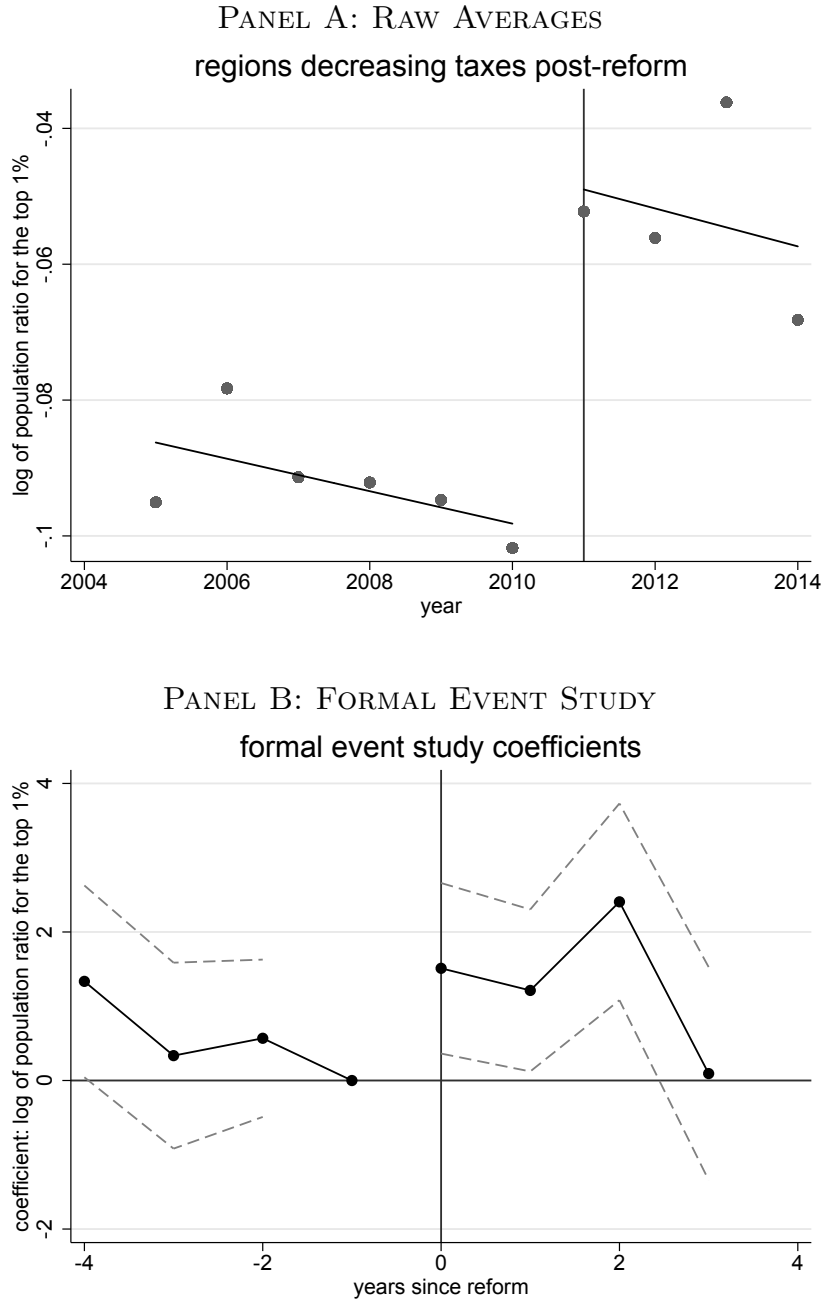
Panel A shows the raw averages of the stock ratios in region pairs where the net-of-tax rate increases in region d relative to region o following the reform. The stock ratio is the population in the top 1% of region d relative to region o. Panel B presents a formal event study where each year indicator is scaled by the average tax differential in the post-reform period between the region pair. Thus, the vertical axis can be interpreted as the effect on the stock ratio of a one percent increase in the net-of-tax rate in the destination region relative to the origin region. We plot 90% confidence bands.

Figure 5: Effects by Occupation and Industry

PANEL A: OCCUPATION
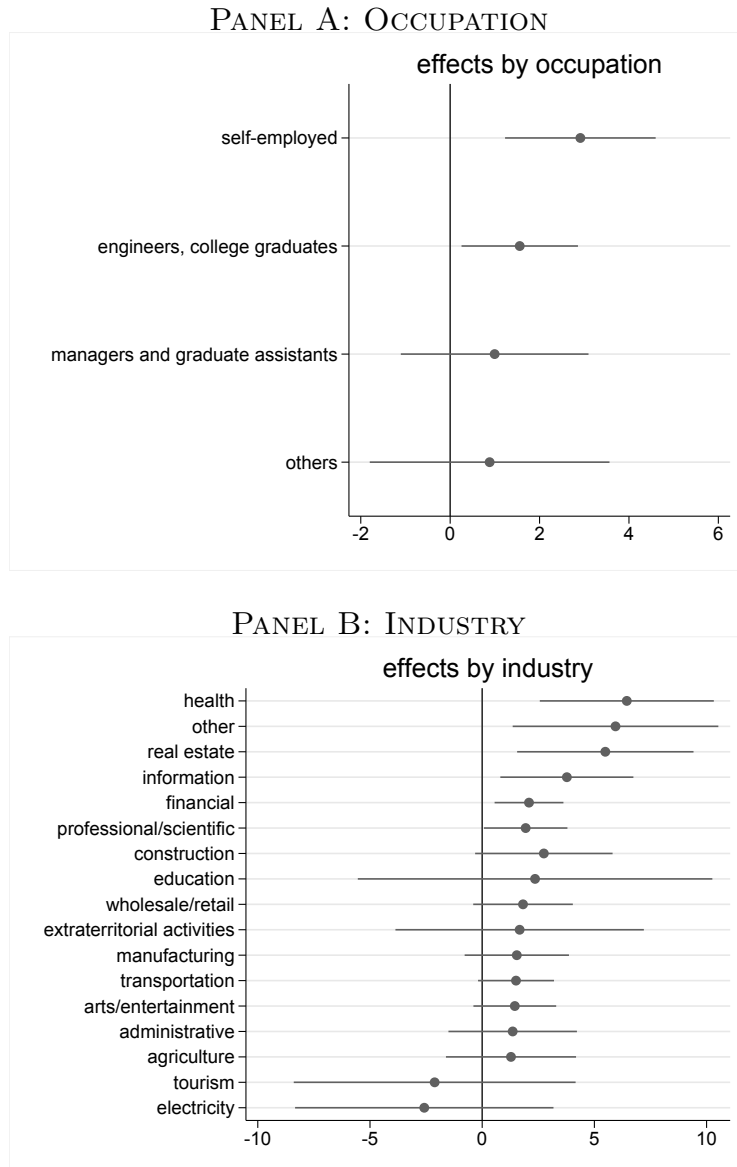
PANEL B: INDUSTRY

This figure shows the effects by occupation and industry with 90% confidence intervals around the point estimates. In Panel A, we show the effects by occupation (sorted by significance and then magnitude). Self-employed are those individuals that have their longest contract with a registered firm. The "other" category includes all occupations not in the previous three groups; these occupations include non-graduate assistants, administrative officers, subordinates, administrative assistants, first and second class officers, third class officers and technicians, and labourers. In the Panel B, we show the effects by industries. Detailed industry descriptions are in the appendix.

Figure 6: Revenue Effects and ETI Simulations

PANEL A: REVENUE EFFECTS
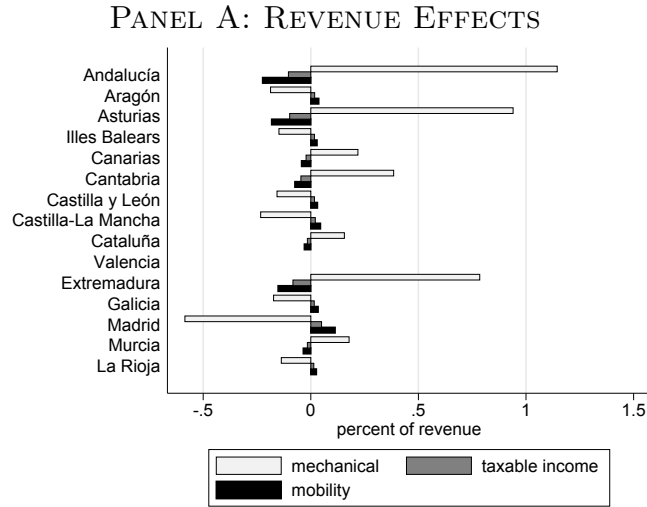
PANEL B: ETI THAT WOULD RESULT IN NO CHANGE IN TAX REVENUE
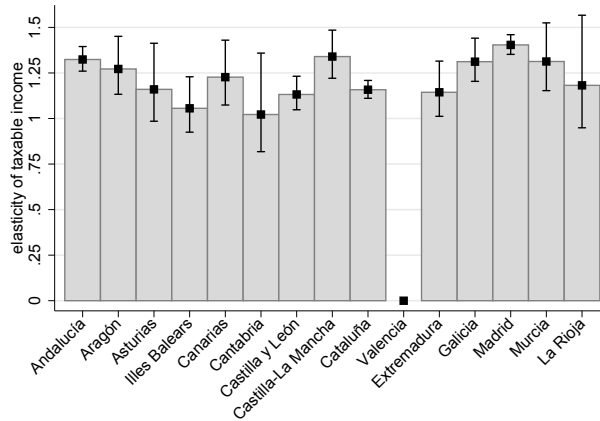
Panel A of the figure shows the breakdown of changes in revenue, as a percent of total personal income tax revenue raised. We show the effect of the region's post-reform tax change relative to if the region had followed the central government tax rate. Thus, when the mechanical effect is negative this corresponds to the region lowering its tax rate; if it is positive, it corresponds to the region raising its tax rate. We focus on tax changes on incomes above 90,000 Euros with an ETI of 0.15. We estimate the Pareto parameter and the mobility effect is taken from our estimate of the stock elasticity. In panel B, the bars show the critical elasticity of taxable income for which the regional deviations from the central government tax schedule would have resulted in the region observing no change in revenue. We obtain confidence intervals using the parametric bootstrap. Valencia simply mimicked the central government and therefore is not included.

42

Table 1: Aggregate Analysis: Effect on Stock Ratios

	Average Tax Rate				Marginal Tax Rate			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$ln[(1 - atr_{d,t})/(1 - atr_{o,t})]$	0.917*	1.116**	1.129**	0.878*	0.652**	0.656**	0.669**	0.556**
	(0.537)	(0.545)	(0.549)	(0.500)	(0.288)	(0.300)	(0.303)	(0.267)
origin FE	Y	Y	Y	Y	Y	Y	Y	Y
destination FE	Y	Y	Y	Y	Y	Y	Y	Y
year FE	Y	Y	Y	Y	Y	Y	Y	Y
govn't spending	Y	Y	Y	Y	Y	Y	Y	Y
controls	N	Y	Y	Y	N	Y	Y	Y
specification			income	churn			income	churn
observations	1050	1050	1050	1050	1050	1050	1050	1050

The dependent variable is the log of the stock ratio, which is the number of individuals in the top 1% in region d relative to region o. The log of net-of-average-tax rate differential is the ratio of the net-of-tax rate in region d relative to region o and uses the average tax rate in the first four columns and marginal tax rate in the last four columns. Controls include demographic, economic, and amenity variables. The expected sign is positive. The last two columns in each set address potential taxable income responses by controlling for the income ratio in the top 1% or by adjusting the stock for the number of people that transition in/out of the top 1% relative to the prior year. The estimates represent the elasticity of the ratio. Standard errors allow for three-way clustering (region pair, origin-year, destination year). *** p<0.01, ** p<0.05, * p<0.1

Table 2: Individual Analysis: Average Tax Rates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$ln(1 - atr_{i,t,j})$	0.588	0.714**	0.894***	0.712**	0.767**	0.714**	0.904***
	(0.420)	(0.343)	(0.336)	(0.337)	(0.336)	(0.343)	(0.332)
place of origin		-0.797***	-0.765***	-0.797***	-0.796***	-0.796***	-0.766***
		(0.061)	(0.060)	(0.061)	(0.061)	(0.061)	(0.060)
place of birth		0.207***	0.206***	0.207***	0.207***	0.207***	0.206***
		(0.022)	(0.021)	(0.022)	(0.022)	(0.022)	(0.021)
place of first work		0.185***	0.177***	0.186***	0.186***	0.185***	0.177***
		(0.020)	(0.020)	(0.020)	(0.020)	(0.019)	(0.020)
work place		0.288***	0.261***	0.288***	0.287***	0.287***	0.261***
		(0.018)	(0.021)	(0.018)	(0.018)	(0.018)	(0.021)
ln(distance)		-0.075***	-0.072***	-0.075***	-0.075***	-0.075***	-0.072***
		(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
move FE	Y	Y	Y	Y	Y	Y	Y
j by year FE	Y	Y	Y	Y	Y	Y	Y
j by edu	N	N	Y	N	N	N	Y
j by age	N	N	N	Y	Y	N	Y
j by age^2	N	N	N	N	Y	N	Y
j by male	N	N	N	N	N	Y	Y
controls	N	Y	Y	Y	Y	Y	Y
observations	13,395	13,395	13,395	13,395	13,395	13,395	13,395
moves	893	893	893	893	893	893	893
R^2	0.122	0.278	0.302	0.279	0.280	0.279	0.304

In all specifications, the estimating sample uses pre- and post-reform moves in the top 1% of the income distribution. Each move has fifteen observations: one for each possible alternative region. The dependent variable equals one if the region is selected. This table uses the person-specific net-of-average-tax rate as the independent variable. All standard errors are clustered two-ways: region-year clusters and move (i, t) clusters. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Individual Analysis: Average Tax Rates with IV

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$ln(1 - atr_{i,t,j})$	1.452	1.542*	1.711**	1.510*	1.614**	1.534*	1.731**
	(0.948)	(0.788)	(0.791)	(0.789)	(0.792)	(0.787)	(0.797)
place of origin		-0.797***	-0.766***	-0.797***	-0.796***	-0.797***	-0.766***
		(0.061)	(0.060)	(0.061)	(0.061)	(0.061)	(0.060)
place of birth		0.207***	0.206***	0.207***	0.207***	0.207***	0.206***
		(0.022)	(0.021)	(0.022)	(0.022)	(0.022)	(0.021)
place of first work		0.185***	0.177***	0.186***	0.186***	0.185***	0.177***
		(0.020)	(0.020)	(0.020)	(0.020)	(0.019)	(0.020)
work place		0.288***	0.261***	0.288***	0.287***	0.287***	0.261***
		(0.018)	(0.021)	(0.018)	(0.018)	(0.018)	(0.021)
ln(distance)		-0.075***	-0.072***	-0.075***	-0.075***	-0.075***	-0.072***
		(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
move FE	Y	Y	Y	Y	Y	Y	Y
j by year FE	Y	Y	Y	Y	Y	Y	Y
j by edu	N	N	Y	N	N	N	Y
j by age	N	N	N	Y	Y	N	Y
j by age^2	N	N	N	N	Y	N	Y
j by male	N	N	N	N	N	Y	Y
controls	N	Y	Y	Y	Y	Y	Y
observations	13,395	13,395	13,395	13,395	13,395	13,395	13,395
moves	893	893	893	893	893	893	893
R^2	0.122	0.278	0.302	0.279	0.280	0.278	0.304
first stage coefficient	0.392***	0.392***	0.392***	0.392***	0.391***	0.392***	0.391***
	(0.015)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
F-statistic	735.1	740.2	745.6	740.2	795.6	740.9	792.9

This table uses the person-specific net-of-average-tax rate and instruments for it with the person-specific net-of-marginal-tax rate. The bottom panel shows the first stage. We present the F-statistic for instrument strength. The treatment of standard errors, the estimating sample and variables are as defined in Table 2. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Placebo Test

	(1)	(2)	(3)	(4)
	Average Tax Rate		Average Tax Rate with IV	
	Post-	Pre-	Post-	Pre-
	Reform	Reform	Reform	Reform
$ln(1 - atr_{i,t,j})$	1.273***	0.286	2.051***	0.093
	(0.414)	(0.356)	(0.687)	(0.469)
move dummies	Y	Y	Y	Y
j by year dummies	Y	Y	Y	Y
$\zeta_j \mathbf{x}_{i,t}$	Y	Y	Y	Y
controls: $\mathbf{z}_{i,t,j}$	Y	Y	Y	Y
observations	4,965	6,180	4,965	6,180
number of moves	331	412	331	412
F-statistic			797.9	509.1

This table shows results of a placebo test verifying that post-reform tax rates have no significant effect on pre-reform migration patterns. The post-reform sample is restricted to migrants in the post-reform period, while the pre-reform sample is restricted to individuals moving in the pre-reform period. To do this, we construct the mean tax rate for each alternative and individual in the post-reform period 2011-2014. Columns (1) and (3) show that even using this mean tax rate, rather than year specific rates, we can obtain similar results as in table 2 and 3. Columns (2) and (4) then use migration decisions in the period 2005-2010, but using the mean tax rates constructed from the period 2011-2014. Column (1) and (2) use the net-of-average-tax rate. Column (3) and (4) use the net-of-average-tax rate and instruments for it using the net-of-marginal-tax rate. All standard errors are clustered two-ways: region-year clusters and move (i,t) clusters. We present the F-statistic as a test of instrument strength. *** p<0.01, ** p<0.05, * p<0.1