Working Papers

Income Taxes, Sorting, and the Costs of Housing: Evidence from Municipal Boundaries in Switzerland

Christoph Basten
Maximilian von Ehrlich
Andrea Lassmann

CESifo
Center for Economic Studies & Ifo Institute

Income Taxes, Sorting, and the Costs of Housing: Evidence from Municipal Boundaries in Switzerland

Abstract

This paper provides novel evidence on the role of income taxes for residential rents and spatial sorting. Drawing on comprehensive apartment-level data, we identify the effects of tax differentials across municipal boundaries in Switzerland. The boundary discontinuity design (BDD) corrects for unobservable location characteristics such as environmental amenities or the access to public goods and thereby reduces the estimated response of housing prices by one half compared to conventional estimates: we identify an income tax elasticity of rents of about 0.26. We complement this approach with census data on local sociodemographic characteristics and show that about one third of this effect can be traced back to a sorting of high-income households into low-tax municipalities. These findings are robust to a matching approach (MBDD) which compares identical residences on opposite sides of the boundary and a number of further sensitivity checks.

JEL-Code: H710, H240, R210, R380, C140.

Keywords: local taxation, income taxation, housing prices, income sorting, boundary discontinuity design, spatial differencing.

Christoph Basten
ETH Zurich / Switzerland
cc.basten@gmail.com

Maximilian von Ehrlich
University of Bern / Switzerland
Maximilian.vonEhrlich@vwi.unibe.ch

Andrea Lassmann
ETH Zurich / Switzerland
lassmann@kof.ethz.ch

1 Introduction

Location decisions of individuals depend among others on neighborhood attributes as well as on the bundle of taxes and public goods offered by local governments. Tiebout (1956) pointed out that mobile citizens weigh the costs and benefits of public goods provided at the local level and thereby induce competition among municipalities and regions that ultimately yields an efficient provision of public goods and a sorting of households according to their preferred bundles. Even with a relatively small number of mobile, well informed individuals who *vote by feet* the mechanism should reflect in housing prices. A marginal difference in local taxes between two municipalities providing the same quality of public goods capitalizes in the price for housing as residents have to be compensated for an increase in taxes by a decline in housing prices. Apart from this direct effect, tax differentials are likely to induce a sorting of households across jurisdictions because rich households have a relatively stronger preference for low taxes than poor households if taxes are progressive or preferences are non-homothetic (see Schmidheiny, 2006a). Accordingly, low-tax jurisdictions are characterized by higher incomes which raises housing prices further.

The empirical relevance of these intuitive mechanisms plays an important role for the local provision of public goods and the optimal design of fiscal federalism. On the one hand, a high elasticity of housing prices with respect to local taxes implies a high degree of mobility and consequently, competition among local governments. This should prevent a misuse of public funds. On the other hand, a high tax elasticity points to the limitations of fiscal decentralization because local income taxes may not be used for redistributional policies and expenditures for goods producing inter-jurisdictional spillovers may be too low (see e.g. Gordon, 1983). Moreover, as pointed out by Hilber (2011) the extent to which fiscal variables capitalize into local housing prices has important implications for the evaluation of a wide range of policies. For instance, a high degree of capitalization may have adverse consequences for intergovernmental aid or tax deductions.

This paper's contribution is twofold: first, we identify the causal effect of local income taxes on the costs of housing and second, we isolate the role played by sociodemographic sorting. Our analysis draws on a unique microgeographic dataset which covers detailed information on the universe of 3.5 million rental advertised residences over the period 2005–2012 and across almost all Swiss municipalities. Due to the large degree of regional autonomy which allows municipalities to charge different income taxes, Switzerland

is often referred to as a prime example for fiscal federalism. We match the apartment level information with census data about sociodemographic characteristics of individual households.

Estimating the response of housing prices to tax changes is complicated by the fact that capitalization of taxes is conditional on the quality of local public goods and services which cannot be measured in a satisfactory way. Moreover, the spatial income distribution and the level of local taxes are determined simultaneously because a larger tax base allows for lower tax rates. This is particularly relevant as (income) sorting may arise for various other reasons than taxes. For instance, households being similar in terms of sociodemographic characteristics such as income, education, or cultural background tend to cluster because of environmental and neighborhood amenities as well as social interactions (Bayer, Ferreira, and McMillan, 2007; Bayer and McMillan, 2012). Likewise, the distance to the central business district (CBD) matters because different types of households value commuting costs differently. Accordingly, sociodemographic variables are intrinsically correlated with unobservable location attributes as well as local taxes. As a consequence the identification of the tax elasticity is hardly convincing in conventional hedonic regression.

We address these endogeneity problems using a boundary discontinuity design and compare residences in a very close neighborhood on both sides of a municipal boundary. Under the assumption that the distribution of potentially confounding variables changes continuously, we can hold local amenities as well as household preferences over neighborhood attributes and public goods constant and use the jump in the income tax across municipal boundaries to identify the marginal willingness to pay for living in a low-tax community. The precise georeferencing of households and apartments allows us to analyze the variation in rents as well as in sociodemographic and apartment characteristics within a narrow distance band from the municipal boundaries. Features of geographical attractiveness certainly vary continuously just as the crime rate, and the population density. Likewise, the access to infrastructure and other non-excludable public goods and services are continuous functions of distance. We embed our approach in a sorting framework and quantify the degree of income sorting caused by tax differences. This enables us to to disentangle the direct capitalization of taxes in housing prices from the indirect effect caused by sociodemographic sorting.

Our results provide strong support for a capitalization of income taxes in rents and housing prices. We show that the average elasticity of annual rents with regard to annual income taxes is about 0.26. However, conven-

3

tional hedonic regressions would suggest a much higher rate of about 0.54. Accordingly, unobservable preferences about local public goods and neighborhood characteristics that offset the effect of higher taxes are decisive and reduce the estimated elasticity by almost one half. Further accounting for sociodemographic characteristics, we show that roughly one third of the higher rents in low-tax municipalities can be attributed to a sorting a high-income households.

The remainder of the paper is organized as follows. Section 2 reviews the literature on taxes and housing prices before we introduce the institutional background in Section 3. Section 4 presents the data and in Section 5 we derive an empirical model for inference about the effect of local income taxes on rents. We describe our key results in Section 6 and perform numerous sensitivity checks in Section 7. The last section concludes with a summary of the key findings.

2   Local taxes and housing prices

Following Oates (1969) there has been ample empirical research on the effect of taxes on property values, but as a matter of the design of tax institutions in the majority of countries, most authors have studied the effect of local property taxes on housing prices (see Ross and Yinger, 1999 and Hilber, 2011 for excellent overviews) while empirical evidence regarding local differences in income taxes and their link to housing prices can hardly be obtained (exceptions are Morger, 2013 and Stull and Stull, 1991).[1] Most of these studies follow the standard hedonic approach to estimate the valuation of housing characteristics and local amenities which however is prone to endogeneity bias resulting from unobservables. Accordingly, in order to address these issues the distinction of possible sources for differences in taxation is essential. These include the variation in household preferences, the efficiency of public goods provision, and the size of the tax base.

Different tax rates between municipalites may lead to rent differentials via the following channels. First, there is a direct effect as disposable income increases with lower taxes. Second, tax savings increase with income and accordingly high-income individuals have a higher willingness to pay for

---

[1] Further empirical tests of Tiebout's sorting mechanism with regard to changes in environmental quality and schooling are manifold (e.g., Banzhaf and Walsh, 2008; Rothstein, 2006). For a discussion of the estimation of housing prices and demand see Sheppard (1999).

housing in the low-tax municipality. With inelastic supply high-income individuals outbid lower income individuals which results in income sorting across municipalities. Third, fixed migration costs imply that high-income individuals are more responsive to tax differences. Suppose there was an exogenous shock to the tax difference between two municipalities. In accordance with the channels described above, this will boost the demand for housing in the low-tax municipality. Yet, there are potential feedback mechanisms on local taxes because a favorable selection of high-income citizens allows for the provision of a given level of public goods at lower tax rates. Similarly, the new selection of citizens may change the set of municipal preferences and strive for a lower level of government expenditures which will further reduce taxes.

The sorting process outlined above is likely to be incomplete once heterogeneous preferences are taken into account (Schmidheiny, 2006b). Conversely, complete income stratification would occur if household preferences varied only over income (see e.g., Epple and Platt, 1998; Epple and Sieg, 1999). For instance, households may self-segregate in the sense that individuals prefer to locate close to other households with similar characteristics. This can explain the clustering of household types within municipalities and across municipal boundaries (Ioannides, 2004). Moreover, individuals may freeride on externalities that local neighborhoods generate across boundaries or exhibit heterogeneous preferences over local public goods. An example for the latter is the relationship between housing prices and schooling quality that has been subject to previous empirical work using similar identification strategies (see, e.g. Bayer, Ferreira, and McMillan, 2007; Black, 1999; Fack and Grenet, 2010; Gibbons, Machin, and Silva, 2013). Similar to these studies showing that the inclusion of neighborhood effects reduces the coefficient of school quality in hedonic regressions, we find that disregarding local unobservables produces an upward bias in the elasticity of housing prices with regard to income taxes. Moreover, we show that income sorting explains a considerable part of the effect. Empirical evidence in favor of incomplete income sorting in Switzerland is found in Hodler and Schmidheiny (2006), Schmidheiny (2006a), and Schaltegger, Somogyi, and Sturm (2011). With incomplete sorting and moving costs income taxes should not fully capitalize in housing prices.

In general, these sources of tax variation raise severe endogeneity problems due to unobserved heterogeneity. Previous work has not been able to account for both the variation in neighborhood characteristics that may reflect individual preferences about the level and quality of public goods and the local variation in income taxes. Utilizing an institutional background

5

that leads to local differences in income taxes together with detailed georeferenced data on individual rental prices, apartment types and sociodemographic characteristics we are able to assess the effects of income taxes on housing prices taking spatial sorting into account.

3 Institutional background: taxes, public goods and housing prices in Switzerland

By the Constitution of the Swiss Confederation, both the federal state and the 26 cantons (which are similar to U.S. States and German *Bundesländer*) are fiscal jurisdictions. It is noticeable that cantons have the main tax authority and the federal state's right to levy taxes is limited both in duration and extent. In practice, cantonal law delegates the authority to levy taxes to the municipalities such that all three state levels – the federal state, the cantons, and the municipalities – are tax jurisdictions, i.e., they set their own tax rates and levy their own income taxes. Elements of direct democracy are pronounced and thus tax rates are set periodically by referenda.[2] While the federal government levies the highest share of total tax revenue (45%), it relies mostly on indirect taxes and income taxes amount to only 15% of its revenue. On the other hand, cantons and municipalities, accounting for 33% and 23% of the total tax revenue respectively, mainly impose direct taxes such as income taxes, which account for 60% of the cantonal and for 70% of the municipal revenue. The shares in total expenditures are 33%, 41%, and 22%, for the federal, cantonal, and municipal level, respectively.[3] In particular, the Swiss Constitution allows cantons to set their own income tax schemes, including the degree of tax progression. As a consequence, the heterogeneity across cantons is large. Notably, municipalities have to set their tax rates as a flat multiple of the cantonal rate such that progression is homogeneous within cantons.

The unique combination of institutional characteristics leads to housing prices that are endogenously determined through differences in income

---

[2]The local (legislative) tax-setting authorities are usually the cantonal parliaments and the municipality councils but exceptions (including compulsory referenda) exist. Tax rates are determined annually in most cantons.

[3]All figures refer to the year 2011 and stem from the Swiss Federal Finance Administration (EFV). Fiscal revenue shares are based on figures for the three levels (federal, cantonal, municipal). Expenditures include fiscal transfers and social security accounts for the remaining 4% of total expenditures.

taxes across jurisdictions that can be traced back exclusively to heterogenous preferences and income sorting. The reasons are as follows: first, interjurisdictional competition is limited through systems of federal and cantonal fiscal equalization schemes in practice. Second, as pointed out by Hodler and Schmidheiny (2006), a minimum level and quality of public goods provision is regulated by (mostly) cantonal law. For instance, teacher salaries and school class size are determined by the cantons and tutoring is limited on the basis of equality considerations, hence not only the level but also the quality of schooling is sufficiently homogeneous. Note also that secondary education (starting at age 13) and tertiary education are completely determined at the cantonal level. Third, the large degree of regional autonomy gives rise to yardstick competition as tax rates are set by referendum, ruling out large-scale inefficiencies as a source of differences in income taxes. Finally, the Swiss tenancy law is flexible and accordingly, rents can usually be adjusted at least once a year.[4] In the following section we illustrate the spatial differences that prevail both regarding the tax rates and rents across cantons and municipalities.

4 Data

We use three data sources that are combined with geographic information about Switzerland. First, the data on housing prices comprise about 3.5 million postings of residences offered for rent during the period 2005-2012. These data stem from *comparis.ch* which is the most widely-used price comparison service in Switzerland offering among others information on real estate, insurance, and phone contracts. For their real estate platform *comparis.ch* collects all offers posted on the 17 most popular apartment search engines in the Swiss market. According to official statistical sources, the total number of (owned and rented) apartments was 4,131,342, and the number of inhabited apartments was 3,534,508 in 2011, thus the number of vacancies was about 600,000.[5] The number of unique observations in the same year covered by our sample is about the same. These figures indicate that the data coverage of our sample is excellent. Our data also reflect the distribution of residences across different categories very well. The median apartment in our

---

[4]In most cases, rents can be adjusted on first April and first October in any year. The law allows an adjustment to rents customary in a place (allowing for local differences within municipalities) regarding existing and new tenants.

[5]Source: Swiss Federal Statistical Office (BFS).

data has 3.5 rooms and a living space of about 80 square meters. Regarding the geographic distribution of postings across cantons we observe a strong concentration of observations (about 20%) in the Canton of Zurich which is approximately consistent with the population share. In 2011, about 18 % of the Swiss population lived in the Canton of Zurich. In addition, the data provide a good coverage of more rural areas such as the cantons Aargau and Schwyz with 51,259 and 11,223 observations in 2011 (corresponding to 8.2% and 1.8% of the total) which fits approximately their population shares of 7.7% and 1.9%.[6] The overall distribution of residences in our sample is shown in Figure 1.

– Figure 1 here –

All prices in the data at our disposal reflect offer prices rather than transaction prices. To account for the potential measurement error that might arise from the difference between offer prices and transaction prices, we focus on rents rather than residential property prices.[7] Immigration has been an important factor that led to increasing demand for housing in recent years. As a consequence, rents are almost never negotiated but taken as given by the tenant. Moreover, real estate brokers offer residences online, in which case the advertisement appears in the data at our disposal. In general, brokers are very rare in this market because they are regulated to earn a flatrate of 500 Swiss Francs (CHF) for matching a landlord and a customer. Importantly, the rental share is relatively high in Switzerland, amounting to 59% in total inhabited accommodations in 2011.[8] Migration costs are likely to be lower for tenants than for property owners and thus capitalization should be more pronounced. Finally, a recent study by Banzhaf and Farooque (2013) for the Los Angeles area has shown that rents capture spatial differences in contemporaneous amenities and income better than property prices.

Second, information on sociodemographic characteristics of individual households is obtained from census data. It covers information on education, employment status, occupation, age, legal status, native language, etc. The

---

[6]Population shares stem from the Swiss Federal Statistical Office (BFS).

[7]Our dataset includes information on residential properties. Yet, we disregard these data as we believe that the effect can be identified more precisely from rental prices.

[8]The residential property share in Switzerland is particularly low in comparison to the neighboring countries (56.4% in Austria, 58.1% in France, 45% in Germany, 77.1% in Italy). Sources: Swiss Federal Statistical Office (BFS), Euroconstruct 2013. The low property share is mostly due to the fact that with the exception of *Valais*, individual apartments could not be purchased until 1965.

corresponding dataset was provided by the Swiss Federal Statistical Office (BFS) and can be geolinked by precise addresses.[9] We match the information on real estate and sociodemographics to information about the geographic location of municipal boundaries that stem from BFS, and to data on the municipal tax burden, both spanning the period 2005–2012. The former allows us to determine the municipality in which individual residences are located as well as the distance from each residence to every municipal boundary in every year. Annual maps are necessary because due to several reforms, the number of municipalities diminished from 2,777 in 2005 to 2,495 in 2012.

Third, we calculate the local income tax burden depending on household types and covering the years 2005-2012. Differences in tax legislation imply that the calculation of the tax base varies across cantons (deductions from gross income are set at the cantonal level). Applying the cantonal tax rate which relies on a continuous and progressive tax schedule to the tax base yields the so-called *Einfache Steuer*. This sets the basis for calculation of the municipal and church taxes, with the corresponding flat tax rate usually being defined in %. More specifically, the overall tax burden is calculated as follows. We collected data on cantonal, municipal and church tax rates, and on poll and school taxes if applicable for every municipality in all years as well as the majority religious shares.[10] We further use data from the Swiss Federal Tax Administration (ESTV) on the total tax burden at the municipal level. These data cover municipalities with at least 2,000 inhabitants until 2010 and all municipalities in later years. Using these data together with the cantonal tax rates we infer the cantonal tax burden i.e., the *Einfache Steuer*. With the cantonal tax burden and the municipal tax multipliers at hand we obtain the municipal tax burden which is added to the cantonal tax burden for all years. This computation is done by household status and income, where we distinguish between single households and married households with 2 children and income= $\{80; 100; 150; 200; 400\}$ tsd. CHF.[11]

---

[9]We use the year 2000 as this is the last census which built on a questionnaire that the entire population of Switzerland had to answer. Accordingly, it shows the demographic, social, and economic developments at the finest possible spatial scale. Later censuses are based on samples and data are available at the municipality level only for a few variables.

[10]The municipal tax multipliers are typically made available on the cantonal or municipal webpages, and where not available we contacted the respective authorities for information. The majority religious shares at the municipality level stem from the 2000 census. All results are robust to using only the observations covered by ESTV data.

[11]More detailed information on the computation of the municipal tax burden is provided in Appendix A.

For reasons of consistency, we apply this approach to all municipalities in all years and test the precision of our data by comparing the tax burden stated by ESTV for municipalities with more than 2,000 inhabitants with our own calculations. The fuzziness resulting from our calculations is minor. Moreover, the ranking of municipalities in terms of their tax level is perfectly matched by our computations and the negligible deviations in the absolute tax burden are due to the tax authority using slightly different representative characteristics for the computation of the tax base.[12]

– Figure 2 here –

In Figure 2 we display the total tax burden across all Swiss municipalities using different colors for the quintiles of the distribution. We plot this map for a single person with an annual income of 100 tsd. CHF in 2012. The western part of Switzerland generally exhibits the highest taxes while the central part levies relatively low taxes. Overall, we observe a large variation. The minimum tax burden applies in the municipality of Wollerau (Schwyz) and corresponds to 4,806 CHF while the highest tax burden applies in Les Planchettes (Neuchâtel) and corresponds to 19,841. Interestingly, we observe municipalities in the lowest quantile of tax burden (marked in light yellow) right next to municipalities in the upper quantile of the distribution (marked in dark blue). Many of these cases are municipality pairs belonging to different cantons but we find considerable variation within cantons, too. The Canton of Graubünden displays the largest variance in tax burden across municipalities for our representative single person household with a range of 4,997 (Rongellen) to 17,547 (Schmitten). We have plotted the distribution of municipal taxes in the Canton of Zurich separately in Panel B. Two interesting observations are immediately evident from the map. First, the most attractive locations at the lake side display very low taxes. Second, the highest taxes are levied at the border of the canton relatively distant from the agglomerations' core.

– Figure 3 here –

How does the distribution of taxes relate to the distribution of rents? The maps in Figure 3 plot the average rent per square meter in a municipality

---

[12]Contrasting the ESTV data about the tax burden for municipalities with more than 2,000 inhabitants with our own calculations yields a correlation coefficient of 0.98 and a mean difference in the annual tax burden of only 0.6% of annual income across all cantons and years. The results are robust to using data covered by ESTV only.

for Switzerland and for the Canton of Zurich, respectively. Compared to the tax burden, we observe less of a clear regional pattern in the distribution of rents as the top quintile (marked in dark red) is scattered across western, central and eastern Switzerland. Rents are highest in the agglomerations of Zurich and Geneva but we find other municipalities in the upper quintile in less agglomerated regions such as around Zug, Schwyz or in Graubünden (where taxes tend to be low). Focusing on the Canton of Zurich we observe the highest rents in the lake-side municipalities and the lowest rents at the cantonal borders which corresponds to the inverse of the distribution of taxes as illustrated in Figure 2. Of course we cannot infer the effect of taxes on the prices of housing from such unconditional correlations because of many confounding factors that become evident from the maps. For instance, numerous of the high-rent and low-tax municipalities in the Canton of Zurich are located in the most attractive environment close to the lake and in commuting distance to the city center while the low-rent and high-tax regions tend to be in the hinterland and offer less convenient public transport links to the city. The municipalities with high rents and low taxes in Graubünden are located in some of the most popular tourist regions while municipalities featuring low rents and relatively (yet not always) higher taxes tend to be located in densely populated mountain areas.

5   Estimating the effect of income taxes on housing prices

This paper's empirical approach identifies the impact of income taxes on rental prices based on a boundary discontinuity design as developed in Black (1999) and Bayer, Ferreira, and McMillan (2007) and refined by Fack and Grenet (2010) and Gibbons, Machin, and Silva (2013). In contrast to these studies focusing on expenditure-side effects – or more precisely, on the valuation of schooling –, our analysis relates rental prices to public revenues. Unobservable local conditions determine the level of taxes and rents at the same time which implies that conventional hedonic housing prices regressions yield biased estimates of the tax elasticity. With geographic location commonly being acknowledged as the key determinant of housing prices it seems advisable to pursue a counterfactual analysis based on observations that share the same neighborhood characteristics. For instance, Albouy and Ehrlich (Albouy and Ehrlich) conclude that space is a much greater deter-

minant of land values than time even when studying data over the housing boom and bust cycle. In urban economic theory and consistent with empirical evidence, the rent gradient typically decreases continuously with distance from the economic center, publicly provided amenities, or other attractive location fundamentals. In contrast, the tax burden changes discontinuously at the municipal boundary in the multijurisdictional tax system of Switzerland. The analysis of narrow windows around spatial boundaries together with boundary fixed effects can thus be exploited to obtain a causal effect of income taxes on rents.[13]

5.1 Boundary discontinuity design (BDD)

Let us use the following notation. We observe postings of individual residences denoted by i in year t which belong to a municipality $m \in M$. Each municipality levies an annual income tax Tax_{mt} – which can vary across household type and income level – from its residents. Using data on administrative boundaries we can assign each municipality all its neighboring municipalities. Restricting our dataset to posts with non-missing and plausible values for the residence characteristics reported in Table 1, we obtain $M = 1,989$ municipalities.[14] We denote by $p = \{m, m'\}$ a pair of neighboring municipalities where each combination of municipalities is considered only once and each residence is uniquely assigned to one municipality pair according to the minimum distance to the municipal boundary.

   The log tax burden is denoted by τ_{mt} which refers to a gross income of 100 tsd. CHF and to a single person household in our benchmark specifications. This corresponds approximately to 1.5 times the median income of an employee in Switzerland. In order to insure a sufficient degree of institutional homogeneity we restrict the analysis to municipalities belonging to the same canton. This implies that the results are not affected by the choice of income

   [13]Note that this approach is similar but not equivalent to a spatial regression discontinuity design, where identification is based on the continuity of distance as a running (or forcing) variable; e.g., Lalive (2008), Dell (2010); see also Eugster and Parchet (2013), Basten and Betz (2013), Egger and Lassmann (2013) for such designs for Switzerland. In contrast to most spatial RDD, the distance from the boundary has no uniform economic interpretation and we match individual residences on both sides of the boundary.

   [14]This refers to the full time period. We assign the respective neighboring municipalities for each year separately according to the corresponding classification and digital maps of municipalities. As some municipal boundaries were modified over the period under consideration we use separate digital maps for each year which are available from BFS.

group because the progression of the tax code is constant within cantons.[15] Likewise, we neglected the federal tax in the calculation of the tax burden for the same reason. Yet, the responsiveness of households may depend on household characteristics and accordingly we analyze the robustness of our results for a representative family with two children.[16]

We use this information together with data on the log rent per square meter at t of residence i located in m, to estimate the tax elasticity from a conventional hedonic regression model of the following form as a baseline for comparison:

$$r_{imt} = \eta + \omega \tau_{mt} + \rho_t + u_{imt}, \tag{1}$$

where ρ_t is a set of year dummies and u_{imt} is an error term that may be clustered at the municipality-year level to take a general form of correlation of observations within municipalities and over time into account.

To overcome the bias in ω that results from omitting relevant factors that are correlated with income taxes but affect the outcome independently, we determine for each municipality pair p the latitude and longitude for up to $B = 24$ different border points $b = \{1, ..., B\}$ using geographic information system (GIS) data. The number of border points depends on the size and the shape of the common boundary. Municipality pairs sharing a long common boundary and those with a very wiggly boundary line are assigned more border points.[17] With the border points at hand, we compute for each residence i the Euclidean distances $D_{imt}(b)$ from all common border points $b \in B$ of the municipality pairs it belongs to. In the next step, we determine for each pair the three border points featuring the greatest density of residences surrounding them. All other border points are dropped because our approach requires a high density of observations at the threshold. Each residence is uniquely assigned to one of the three remaining border points using the minimum distance $D_{imbt} = \min\{D_{imt}(b)\}$. Finally, outcome r_{imbt}

---

[15]For two municipalities with the same degree of progression the ranking of tax burdens between m and m' remains the same over all income groups. This implies that the notation can neglect indices referring to household types and income levels. The median gross income of an employee was 67,500 CHF in Switzerland in 2012 (see BFS, 2012).

[16]We provide the results for annual incomes of 80, 150, 200 and 400 tsd. CHF for a single household as well as for a representative family with two children in an online appendix. All results are stable across different income categories.

[17]This procedure maximizes the number of residences i located in different municipalities within a close neighborhood.

is measured for each residence i located in municipality m and assigned to border point b at year t.

This procedure allows us to estimate a BDD model which can be stated as follows:

$$r_{imbt} = \alpha + \beta \tau_{mt} + \theta_{bt} + \varepsilon_{imbt}, \tag{2}$$

where β measures the tax elasticity, θ_{bt} is the border-point- and year-specific fixed effect that absorbs all variation specific to the neighborhood of the border-point in both the cross-sectional and the time dimension, and ε_{imbt} is an error term which we cluster at the boundary-year level to account for serial correlation along boundaries that may be due to both spatial correlation and correlation over time.[18] By limiting the sample to residences within a small neighborhood of the border-points we ensure that θ_{bt} accurately captures location-specific unobservables.

5.2   Matching residences across boundaries (MBDD)

As an alternative approach, we follow Fack and Grenet (2010) and Gibbons, Machin, and Silva (2013) and match each posting i on one side of b in m with a counterfactual i' on the other side of b in m'. The counterfactual transaction $r_{i'm'bt}$ is calculated as the distance-weighted geographic mean of the prices of all transactions j that took place within a certain radius and in the same year as transaction i but were located in municipality m' rather than in m:

$$r_{i'm'bt} = \sum_{j=1}^{J} \frac{\frac{1}{D_{ij}}}{\sum_{j=1}^{J} \frac{1}{D_{ij}}} r_{jm'bt}. \tag{3}$$

Note that $\frac{1}{D_{ij}}$ refers to the inverse distance between a residence j and reference sale i (both posted in the same year).[19] Assuming that the reference transaction and the counterfactual are sufficiently close, they share

---

[18]Note that in general, a residence i that has n neighbors will be in n pairs and accordingly the error terms of all these n terms will be correlated. This correlation occurs because the residence-specific error term enters the error of each pair. For this reason each residence i is used only once. We assign each i a unique municipality pair according to the minimum distance to the boundary.

[19]For the sake of computational ease, we use the inverse distance between postings i and j via the border point. With many border points it can be shown that this yields virtually the same values as computing direct distances.

the same unobservable time-varying neighborhood effect and the regression model (referred to as MBDD) becomes:

$$r_{imbt} - r_{i'm'bt} = \gamma(\tau_{mt} - \tau_{m't}) + (\epsilon_{imbt} - \epsilon_{i'm'bt}). \tag{4}$$

Parameter γ reflects the tax elasticity of rents and can be estimated by regressing the price differential between reference and counterfactual rent on the corresponding tax differential. Since the identification assumptions are more likely to be valid for observations with a small distance, we estimate an inverse-distance weighted model. We apply a weight $\sum_{j=1}^{J} \frac{1}{D_{ij}}$ to each pair $\{i, i'\}$ such that observations in a close neighborhood receive more weight. Standard errors are clustered at the boundary-year level to account for the serial correlation in the error term induced by differencing. Note that this approach differs from the model in equation (2) by eliminating common spatial trends through spatial differencing and by taking the distance between individual residences at opposing sides of the boundary into account. Since we pool the data over the years 2005-2012, we additionally allow for a common time trend λ_t by including year dummies. While the tax burden differential varies across years, all municipalities within each pair in our boundary sample generally offer the higher/lower tax rate over the full time period.

5.3   Discussion of the identifying assumptions

Under the assumption that unobservable determinants of rents vary continuously at the boundary, the parameters β and γ identify the causal effect of interest for observations in the vicinity of the boundary in the BDD and MBDD approaches, respectively. We choose the radius from the municipality boundary such that it guarantees both the consistent estimation of the coefficients of interest as well as sufficient precision. Hence, all estimates are reported for three alternative bandwidths: we use samples of residences located within 1 km, 600 meters, and 300 meters from the closest border point. Note that municipality-pair-specific distances instead of border-point-specific distances would be adequate to condition on the commuting costs to the neighbor municipality but would not capture the local conditions of a small well defined area: a residence i in m may be as close to the p-boundary as a residence j located in neighboring municipality m' but still i and j might be distant from each other and located in environments of very different attractiveness. In contrast, by way of the chosen design we condition on the

15

proximity of residences i and j as well as on the proximity to the municipal boundary in terms of their proximity to a common border point. Hence, this approach allows us to hold neighborhood characteristics constant.

The resulting distribution of observations in space around individual border points is best described in a map as the one in Figure 4. The map exemplifies for two neighboring municipalities, Küsnacht and Zollikon (Canton of Zurich), how individual observations are allocated to and compared across municipal border-points on a narrow spatial scale: the first type of residences (green) is located at the lakeside, the second type (blue) is located close to the centers and along main traffic routes, and the third type (red) is located outside the center on the hill in both municipalities. Each of the three border points will be assigned a unique fixed effect and the sample will be restricted to observations within the bandwidth. Hence, using the 300 meter radius the red border point may not be used in the regressions due to the low density of residences in the northern municipality. In order to avoid differences in the sample composition of border points we restrict all specifications to border points covered by the 300 meter boundary sample.[20] The MBDD framework further improves on the continuity assumption as it attributes a greater weight to those residence pairs that are closer than others within each circle drawn in Figure 4. Specifically, we construct for each dot in the southern part of the circle a counterfactual residence which is a weighted average of the dots on the northern side of the boundary and within the circle according to equation (3).

The identifying assumptions would be violated if we observed a bunching of residences and postings on the low (or high) tax side of the border. A discontinuity in the density of the observations would point to a systematic difference in the supply of housing which is correlated with local taxes. As suggested by McCrary (2008), we illustrate the density of our observations in equally sized bins of the distance from the boundary in the upper left panel of Figure 5. Naturally, the frequency of postings decreases in the close neighborhood of a municipality boundary which is due to residences being concentrated in the center of a municipality. This process holds true in treated and control municipalities and we observe an almost symmetric shape of the histograms on both sides of the boundary. Even the immedi-

---

[20]The distance between municipalities may have implications for strategic tax setting. Accordingly, border points/municipality pairs that enter the sample only for larger distances may be systematically different from others and we restrict our analysis to municipality pairs and border points that display a sufficiently high density of residences with the smallest bandwidth.

ate bins at both sides of zero distance have approximately the same density. Hence, there is no indication of a bias toward more postings on either side of the threshold.

– Figure 4 here –

Finally, our identification strategies rely on the assumption that estimates of β and γ can be unambiguously attributed to the jump in the income tax. If instead other factors varied discontinuously at b, our estimates could not isolate the effect of taxes. It is not plausible to disentangle the effect of income tax differences at cantonal borders from other political, economic and institutional differences. For this reason we generally restrict our sample to municipal boundaries which do no coincide with cantonal borders. Further violations of the continuity assumption may include: (i) geographic barriers that separate municipalities at b, (ii) apartment characteristics being systematically different in m and m', (iii) asymmetric level and quality of *excludable* public goods between municipalities, and (iv) individual preferences to live at a particular address in central business districts.

We address these concerns in the following way: (i), we drop pairs that are separated by rivers and highways or feature a difference in altitude of more than 400 meters. In addition we drop pairs split by language borders.[21] (ii), we account for apartment characteristics by using residuals obtained from first-stage regressions of the log rent per square meter on a comprehensive set of characteristics in all specifications.[22] (iii), we exploit tax variation within school districts in Section 7 because elementary schools are typically financed on the municipal level. Other publicly provided goods such as health services, roads, cultural events are either not municipality-specific, regulated on the cantonal and federal level, or not exclusively limited to local residents. In the latter case the usage costs become a continuous function of distance. Identifying the effects only from tax rate variation over time – e.g. adding municipality fixed effects – further supports this argument. In order to rule

---

[21]German, French, Italian, and Romansh are official languages in Switzerland. Language borders are defined by the majority of the respective language speakers within municipalities.

[22]These include the variables summarized in the next section, nonlinear and interaction terms of them as well as year dummies. Moreover, all results remain very robust to alternative specifications of the first-stage. In our preferred first-stage specification, we explain about 15% of the variation in log rents. The corresponding regression tables are displayed in an online appendix.

out CBD specific effects (iv), we estimate equations (2) and (4) for observations that exclude the agglomerations of Basel, Bern, Geneva, Lausanne, and Zurich separately.

6 Results

6.1 Descriptive statistics

In total we can assign about 2.5 million postings containing non-missing information on square meters, rents, and many other apartment characteristics to the municipalities which we consider between the years 2005 and 2012.[23] These include information about detailed characteristics of the residence as listed for the full dataset in columns (1) and (2) of Table 1. Apart from age, number of rooms, and the apartment's floor these are binary variables that capture information on the quality and type of the residence. We use the following sociodemographic covariates: the shares of high- and low-income population, of high- and low-skilled population, of foreigners, of unemployed, of commuters, as well as the average age. These data are based on individual information which is georeferenced and assigned to grids of 1×1 km. Then, we compute the respective shares and averages on the basis of the total population within each grid. Income groups are constructed according to detailed occupation categories and job functions. For details about the tax data and the sociodemographic variables see Appendices A and B.

Regarding our identification strategy we resort to observations in the close neighborhood of a municipal boundary. The boundary sub-sample (columns (3) and (4) of Table 1) consists of residences that could be assigned to a border point b which features at least 2 observations on both sides of the boundary and within a distance of 300 meters. Accordingly, the municipality composition remains stable when restricting the boundary sub-sample to alternative maximum distances of 1 km and 600 meters. All border points which correspond to geographic or linguistic barriers have been dropped. Moreover, we drop municipality pairs with a tax differential of less than 100 CHF and we keep only residences that lie in a 1×1 km grid which has a population of at least 50 individuals. The latter should ensure that the

---

[23]Note that we loose about one million of postings due the following reasons: First, we drop observations with missing or non-credible information regarding rents, rooms, and square meters. Second, we drop postings offering a rent of less than 5 CHF or more than 52 CHF per square meter. These values equal approximately the 1st and 99th percentiles.

sociodemographic information is representative of the neighborhood. Tax burden and rents are measured in CHF where T_m^{S100} refers to the annual tax burden of a single household earning 100 thousand CHF per year. In total the boundary sub-sample consists of 151,011 postings in 211 municipalities. The average rent per square meter amounts to 21.1 and and 22.2 CHF in the full sample and the boundary sub-sample, respectively. Furthermore, the average annual tax burden (T_m^{S100}) is about 13,900 CHF in both samples. When comparing the boundary sub-sample – which consists of only about 10% of the total number of municipalities – to the full dataset, we observe that the moments of the data are remarkably similar not only for the main variables but also for the sociodemographic and residence covariates. The boundary sample focuses on 190 border points with sufficiently close residences on both sides of the boundary and 701 associated grid cells. Accordingly, a border point is assigned on average about 800 observations and 3.6 grid cells (1×1 km).

– Table 1 here –

Table 1 provides further information about the data. The average apartment has about 3.5 rooms, is about 40 years old, and is located on the second floor. More than 70% of the residences either have a balcony or a terrace and 40% have a parquet floor. The share of high-skilled is about 10% and the share of foreigners amounts to about 23%. Overall, these figures seem reasonable and representative. Limiting the sample with respect to a distance of 600 and 300 meters from the closest border point results in samples of 89,058, and 30,980 observations, respectively. Within these distance bands, the observed average rents are 21.8 CHF and 21.4 CHF. The figures for the average annual income tax burden of our benchmark household with a gross annual income of 100 thousand CHF amount to 14,113 CHF and to 14,202 CHF. In general, our dataset has a remarkable support at the border regions which allows us to condition very precisely on the location of residences. This is a feature of the data coverage, the relatively small size of municipalities, and the high population density in Switzerland.

6.2   Graphical assessment of boundary discontinuities

In Figure 5 we illustrate the symmetry of the distribution of residences across boundaries as described earlier, as well as the deviations in tax burden, rents per square meter, and residual rents from the border-point averages for observations on the low- and high-tax side of each boundary. Note that we pool

all border-points and structure the data such that high-tax municipalities are assigned negative distances whereas low-tax municipalities have positive distances from the boundary. Moreover, the figures show local averages over distance bins of 100 meters as well as flexible polynomial fits that were estimated separately on both sides of the boundary.[24] In each figure we restrict the sample to a distance band of 1 km on either side of the boundary which corresponds to the maximum bandwidth used in the estimations.

By construction, the difference in tax burden between municipalities on opposite sides of the boundary features a discontinuity at the boundary (see upper right panel of Figure 5). On average the difference between taxes (T_m^{S100} in neighboring low- and high-tax municipalities is about 750 CHF or 5% of the sample average in T_m^{S100}. The municipal distance from the boundary is measured by the average distance of observed residences within a municipality. The lower panel of Figure 5 suggests that the jump in the rental price per square meter and even more so, the one in residual rents is pronounced at the municipal boundary. We observe an average difference in log rent per square meter and residual rent of 0.02 and 0.03, respectively. In both cases, this corresponds to approximately 1% of the dependent variables' sample average. Accordingly, the figures suggest that a 5% tax differential results in a 1% rent difference or a 20% capitalization rate. The 95% confidence intervals of the polynomial fits indicate that the discontinuities are highly significant for all three variables.

Figure 6 visualizes the discontinuity in the share of skilled versus unskilled and high-income versus low-income individuals in order to shed light on the extent of sorting across municipalities which is a prominent channel to explain the rent differential. The figures are constructed in the same way as described above and indicate a clear discontinuity at the boundary. For each variable the jump at the boundary is strongly significant. At the boundary, the share of high-skilled and high-income individuals increases by about 45 and 50% in the low-tax municipality compared to the high-tax municipality. Conversely, the share of low-skilled and low income individuals jumps substantially when moving from the low-tax to the high-tax side of the boundary. In the next section we analyze the significance of these differences and quantify the elasticities as well as the relative importance of

---

[24]An alternative and qualitatively equivalent way of displaying the discontinuities would be to regress the variable in question on border-point fixed effects and on dummy variables for each bin where the coefficients on these distance dummies correspond to points on the polynomial fit displayed in our figures. We obtain very similar graphs with both approaches and accordingly chose the continuous version.

income sorting.

– Figures 5 and 6 here –

6.3   BDD results

Table 2 reports the coefficients and standard errors based on equation (1) in columns (1) and (2) and equation (2) in columns (3)-(8). The table is structured as follows. We use log rental prices per square meter as outcome in Panel A and residual rents purged of the influence of apartment characteristics in Panel B. The latter are derived as the residuals from a first stage regression of log rents on flexible functions of the apartment characteristics reported in Table 1.[25] Uneven columns refer to specifications that include all municipalities in the chosen sample while even columns exclude the five largest agglomerations. This leads to 605 and 522 border points. The distance band around the border points is limited to 1 km in columns (3) and (4). It is reduced to 600 and 300 meters in columns (5)-(6) and columns (7)-(8), respectively. The results may be summarized as follows. First, the elasticities carry a negative sign as expected and are statistically highly significant. Second, the inclusion of boundary fixed effects – columns (3)-(8) – is able to reduce the bias from the conventional hedonic approach considerably and contributes substantially in terms of the models' explanatory power. Third, limiting the sample to close distance bands around boundaries lowers the coefficients further. Fourth, taking residence characteristics into account increases the estimates of the tax elasticity slightly and becomes robust to the results in Panel A the smaller we choose the distance to the boundary. Quantitatively, we find that an increase in the tax burden by 1% lowers rental prices by approximately 0.5% using the conventional approach. This effect reduces to -0.248 and to -0.258 regarding outcomes in Panel A versus Panel B when accounting for border point fixed effects and limiting the distance of observations to 300 meters from the boundary. It amounts to -0.282 and -0.311 when excluding observations located in large agglomerations. Note also that excluding agglomerations does not systematically affect the tax elasticity: while the effect is somewhat more pronounced when including agglomerations in the 600 meter sample the opposite is true for

---

[25]We estimate numerous first-stage specifications and account for the fact that the premium attached to apartment characteristics may vary across cantons and may be different in cities compared to more rural areas. Our results remain robust to these alternative first-stages.

the 300 meter sample. In any case the difference in the coefficients is negligible which suggests that the estimated tax elasticity is not confounded by differences of population density or other agglomeration effects.

Overall, the size of the estimates drops by roughly 50% when accounting for unobservable neighborhood characteristics compared to the conventional hedonic regressions in columns (1) and (2). Moreover, border fixed effects explain about 40% of the variation in the data.

– Table 2 about here –

6.4   MBDD results

Results based on the approach in equation (4) are shown in Table 3. The structure of the table and the sample composition is exactly the same as in Table 2. We additionally report the average distance between reference residences and the counterfactual that matches observations located in the neighboring municipality. The elasticities are remarkably similar to the ones in Table 2, however, unobservable neighborhood characteristics seem to be absorbed already for relatively large bandwidths. While the BDD coefficients drop further when moving from a bandwidth of 1 km to 300 meters, the MBDD's coefficients are comparable to the ones in column (8) of Table 2 when using the 1 km window. Hence, due to contrasting individual residences and assigning more weight to nearby units, the MBDD represents the more efficient approach compared to the BDD. Moreover, the coefficients are quite robust over all three samples. Our preferred estimates correspond to the 300 meter boundary sample which display an income tax elasticity that amounts to roughly -0.253 in Panel A and to -0.264 in Panel B (column (5)). Again, excluding agglomerations does not systematically affect the estimates.

– Table 3 about here –

6.5   The role of sociodemographic sorting

Apart from the direct capitalization effect of tax differentials it is evident from Figure 6 that tax differentials lead to a sorting of households. We quantify the share of rent differences between high- and low-tax municipalities that is due to sociodemographic sorting by including the vector of census block covariates summarized in Table 1 on the right-hand side of equations (2) and (4). Table 4 reports the coefficients and standard errors from both

equations with distance bands restricted to 600 and 300 meters. Compared to the preferred specifications with 300 meters distance in Table 3 the tax elasticity of rents is reduced by 20-30% and amounts to -0.181 and to -0.199 for log rents and residual rents, respectively. The corresponding specifications disregarding local sociodemographics yield elasticities of -0.253 and -0.264 for log rents and residual rents, respectively. This suggests that low taxes attract high-income households and accordingly a larger housing budget which drives up rents in addition to the direct capitalization effect.

– Table 4 about here –

7 Sensitivity

Our results have been shown to be very stable across the two approaches. In general, matching residences to individual counterfactuals in the neighboring municipalities allows for a high degree of precision compared to the BDD even with relatively larger windows. We test the sensitivity of these results and in particular for the preferred MBDD specifications along the following lines:

The role of schooling. At the municipality level in Switzerland the only relevant public good that appears excludable is elementary schooling. As mentioned above a large degree of homogeneity across municipalities is guaranteed in elementary schooling as well. The main reason for differences in schooling should be the composition of pupils. This represents a channel for the indirect effect of taxes on housing prices via sorting which is captured by the chosen design. Yet, in principle the quality of elementary schools may change discontinuously at the municipal boundary and accordingly raises concerns about the consistency of our estimates. In order to address this concern, we digitalized maps on school districts for cantons where applicable: these are the Cantons of Zurich, Bern, Aargau, Fribourg, and Vaud. In these cantons the boundaries of school attendance zones do not always coincide with municipal boundaries. Accordingly we focus on observations along boundaries that are located within the same school district yet face different levels of income taxes. The corresponding results are summarized in Table 5.

Since the number of municipality pairs drops considerably when using only boundaries within the same school district we focus on the estimations

within the 600 meter window. As has been shown in Tables 2 and 3 this bandwidth is appropriate to capture neighborhood specific unobservables. Panels A (log rent per square meter) and B (residual rents) report the coefficients for the BDD and MBDD approaches where sociodemographic sorting is accounted for in columns (2) and (4). While the coefficients' magnitude tends to increase somewhat when accounting for schools and using log rent per square meter as the dependent variable, the opposite is true for residual rents. Limiting the sample to units within the same school district, the benchmark MBDD estimate increases from -0.305 to -0.420 and decreases from -0.395 to -0.323 in the log rent and residual rent specifications, respectively. However, in none of the specifications we can reject that the coefficients are identical for the two samples at conventional levels of significance. Note that restricting the sample to residences located in the same school district reduces the number of observations substantially and accordingly raises the standard errors but most coefficients remain remarkably significant. Moreover, the estimates reported in Table 5 are consistent with the finding that about two thirds of the tax elasticity are due to direct capitalization effects and approximately one third is driven by household sorting. Comparing columns (2) and (1) and columns (4) and (3), we observe a reduction of the estimated tax elasticity between 13 and 29% when including sociodemographic covariates.

– Table 5 about here –

Placebo discontinuities. We address the possibility that rent differentials are erroneously attributed to tax differentials by shifting municipal boundaries artificially within both m and m'. For this we set new 'fake' boundaries which are shifted 300 meters from the true boundary point for either municipality within a (border-point-specific) pair. Then, we reassign the newly treated observations the low tax rate and the newly non-treated observations the high tax rate. Finally, we estimate the BDD model using the new tax burden variable. Table 6 reports the corresponding results for log rents per square meter as well as residual rents and for a shift of the true boundary towards the low- or high-tax municipalities. Hence, in the former case we contaminate the control units and in the latter case we contaminate the treated. None of the estimates in Table 6 is significant and the coefficients' magnitudes are far from our benchmark. This falsification exercise adds further confidence in our results as it shows that spatial trends in unobservables can be effectively eliminated by spatial differencing.

– Table 6 about here –

24

Functional form misspecification. In line with most regression discontinuity approaches, we account for spatial trends that lead to average rent differences across boundaries by adding flexible forms of distance or coordinates to equation (4):

$$r_{imbt} - r_{i'm'bt} = \delta(\tau_{mt} - \tau_{m't}) + f(l_{imbt}) - f(l_{i'm'bt}) + (\nu_{imbt} - \nu_{i'm'bt}), \quad (5)$$

where $f(l_{imbt})$ and $f(l_{i'm'bt})$ include (i) a cubic polynomial function of distance to b, and (ii) the geographic coordinates of locations. Results are presented in columns (1) and (2) of Table 7 and are directly comparable to the ones shown in column (5) of Table 3. The coefficients are nearly identical which confirms the robustness of the estimated tax elasticities to the inclusion of flexible forms of distance to the boundary. Note that spatial trends across boundaries have been taken into account already by weighting the regressions using the inverse distance.

Furthermore, nonlinearities of tax rates may apply. The MBDD approach identifies the coefficient on $(\tau_{mt} - \tau_{m't})$ which reflects the tax benefit from living in municipality m compared to living in municipality m'. For some municipality pairs the tax differential is only minor and may not lead to an aggregate price response, at least as long as the differential is below individual fixed migration costs. Similarly, it may be the case that for sufficiently high levels of taxes all mobile households leave a municipality and only the elderly or other tax-inelastic groups remain. In this scenario, the marginal effect of a tax increase will be negligible. Accordingly, it seems plausible to expect a nonlinear relationship between rents and local income taxes. We approach the potential nonlinear response by (i) including a quadratic term of the tax differential in equation (4); and by (ii) focusing on comparisons of observations i and their counterfactual i' where the tax differential exceeds the median tax differential in the sample, amounting to 731 CHF. Columns (3) and (4) suggest that both approaches leave previous results qualitatively unaffected. At the same time column (3) provides some indication that the elasticity diminishes with high tax differentials.

Endogeneity of municipality boundaries and taxes. Due to municipality mergers, the data are georeferenced by year. However, if municipality mergers are more or less likely for neighbors with large tax differentials this would induce a selection bias. We address the possibility that the tax differential is simultaneously determined with the tax base by dropping unstable municipal boundaries as a robustness check. Hence, we focus on the subset of

municipalities that maintained their boundaries unchanged over the entire time horizon in column (5) of Table 7.

A further concern relates to pre-existing spatial differences in the composition of the population which led to the establishment of municipal boundaries along the current lines and renders taxes endogenous. It seems very unlikely that the population composition featured discontinuities in the absence of municipal boundaries and natural irregularities as the costs of social interactions tend to be a smooth function of distance. Of course, the continuity assumption in the absence of municipal boundaries does not hold if natural geographic barriers such as rivers or mountains prevail. We address this concern by dropping boundary segments coinciding with natural irregularities from the sample throughout all regressions. Moreover, we include municipality fixed effects in MBDD specifications. Thereby, we absorb all time-invariant differences in the population composition of municipalities and identify the tax elasticity exclusively from variation in the level of taxes over time. The corresponding results are reported in column (6) of Table 7 and confirm our results. Exploiting exclusively variation over time reduces the efficiency of the estimates but the coefficients remain significant and almost identical in magnitude. Note that municipality fixed effects capture the average sociodemographics which determine the preferences of the electorate but they do not reflect the distribution of households within municipalities and in particular the sorting at the boundary as captured by our sociodemographic covariates. Thus, we still observe a reduction of the coefficient by about one third when adding the sociodemographic covariates to the specification with municipality fixed effects.

– Table 7 about here –

Heterogeneity in households and residences. As a final sensitivity check we study the heterogeneity of the tax elasticity with regard to different household and residence types. In column (7) of Table 7 we report the results for a representative married couple with two children and annual income of 100 tsd. CHF. We obtained similar results for four additional income choices of 80, 150, 200, and 400 tsd. CHF.[26] The point estimates remain qualitatively unchanged compared to our benchmark. This can be explained by the fact that the tax progression across income groups and household types is fairly stable within cantons. Yet, we observe heterogenous responses to income

---

[26]These results and a battery of further specifications are made available in an online appendix.

taxes for different residence types: we split the sample into quartiles based on apartment size and estimated equation (4) separately for each of the four quartiles. Table 8 reports the corresponding results and shows throughout all specifications an increased responsiveness of smaller apartments. This points to an increased mobility of smaller households compared to larger ones and is consistent with household size as a determinant of migration cost.

– Table 8 about here –

8   Concluding remarks

Local income taxes directly capitalize in housing prices and indirectly affect the latter through a spatial sorting of households according to income. The degree of capitalization and spatial sorting is of key importance for the optimal design of many policy measures as well as for the configuration of fiscal federalism in general. Previous studies have been confined to property taxes and were complicated by unobservable confounding factors such as heterogeneous household preferences, environmental amenities, and the quality of local public goods and services. This paper corrects for unobservable location characteristics and disentangles the direct capitalization effect and the role of household sorting using comprehensive apartment/household-level data on rents and sociodemographic characteristics. We identify the income tax elasticity using a boundary discontinuity design and alternatively a matching approach. In both cases, we estimate an income tax elasticity of about 0.26 which amounts to about one half of the estimate from conventional hedonic regressions and thus points to the important role of unobservable location characteristics. Moreover, our results show that about one third of the overall effect is due to household sorting whereas two thirds can be attributed to direct capitalization of fiscal variables in housing prices.

References

Albouy, D. and G. Ehrlich. The distribution of urban land values: Evidence from market transactions. Lincoln Institute Product Code: WP14XX.

Banzhaf, H. S. and O. Farooque (2013). Interjurisdictional housing prices and spatial amenities: Which measures of housing prices reflect local public goods? *Regional Science and Urban Economics 43*(4), 635–648.

Banzhaf, H. S. and R. P. Walsh (2008). Do people vote with their feet? An empirical test of Tiebout. *American Economic Review 98*(3), 843–63.

Basten, C. and F. Betz (2013). Beyond work ethic: Religion, individual, and political preferences. *American Economic Journal: Economic Policy 5*(3), 67–91.

Bayer, P., F. Ferreira, and R. McMillan (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy 115*(4), 588–638.

Bayer, P. and R. McMillan (2012). Tiebout sorting and neighborhood stratification. *Journal of Public Economics 96*(11), 1129–1143.

Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *The Quarterly Journal of Economics 114*(2), 577–599.

Dell, M. (2010). The persistent effects of peru's mining mita. *Econometrica 78*(6), 1863–1903.

Egger, P. and A. Lassmann (2013). The causal impact of common native language on international trade: Evidence from a spatial regression discontinuity design. CEPR Discussion Papers 9441, C.E.P.R. Discussion Papers.

Epple, D. and G. J. Platt (1998). Equilibrium and local redistribution in an urban economy when households differ in both preferences and incomes. *Journal of Urban Economics 43*(1), 23–51.

Epple, D. and H. Sieg (1999). Estimating equilibrium models of local jurisdictions. *Journal of Political Economy 107*(4), 645–681.

Eugster, B. and R. Parchet (2013). Culture and taxes: Towards identifying tax competition. Economics Working Paper Series 1339, University of St. Gallen, School of Economics and Political Science.

Fack, G. and J. Grenet (2010). When do better schools raise housing prices? Evidence from Paris public and private schools. *Journal of Public Economics 94*(1-2), 59–77.

Gibbons, S., S. Machin, and O. Silva (2013). Valuing school quality using boundary discontinuities. *Journal of Urban Economics 75*(C), 15–28.

Gordon, R. H. (1983). An optimal taxation approach to fiscal federalism. *The Quarterly Journal of Economics 98*(4), 567–586.

Hilber, C. A. L. (2011, October). The Economic implications of house price capitalization: A survey of an emerging literature. SERC Discussion Papers 0091, Spatial Economics Research Centre, LSE.

Hodler, R. and K. Schmidheiny (2006). How fiscal decentralization fflattens progressive taxes. *FinanzArchiv: Public Finance Analysis 62*(2), 281–304.

Ioannides, Y. M. (2004). Neighborhood income distributions. *Journal of Urban Economics 56*(3), 435–457.

Lalive, R. (2008). How do extended benefits affect unemployment duration: A regression discontinuity approach. *Journal of Econometrics 142*(2), 785–806.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

Morger, M. (2013). Heterogeneity in income tax capitalization and its effects on segregation within switzerland.

Oates, W. E. (1969). The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the tiebout hypothesis. *Journal of Political Economy 77*(6), 957–971.

Parchet, R. (2012). Are local tax rates strategic complements or substitutes? ERSA conference papers ersa12p313, European Regional Science Association.

Ross, S. and J. Yinger (1999). Sorting and voting: A review of the literature on urban public finance. In P. C. Cheshire and E. S. Mills (Eds.), *Handbook of Regional and Urban Economics*, Volume 3 of *Handbook of Regional and Urban Economics*, Chapter 47, pp. 2001–2060. Elsevier.

Rothstein, J. M. (2006). Good principals or good peers? Parental valuation of school characteristics, Tiebout equilibrium, and the incentive effects of competition among jurisdictions. *American Economic Review 96*(4), 1333–1350.

Schaltegger, C. A., F. Somogyi, and J.-E. Sturm (2011). Tax competition and income sorting: Evidence from the zurich metropolitan area. *European Journal of Political Economy 27*(3), 455–470.

Schmidheiny, K. (2006a). Income segregation and local progressive taxation: Empirical evidence from switzerland. *Journal of Public Economics 90*(3), 429–458.

Schmidheiny, K. (2006b). Income segregation from local income taxation when households differ in both preferences and incomes. *Regional Science and Urban Economics 36*(2), 270–299.

Sheppard, S. (1999). Hedonic analysis of housing markets. In P. C. Cheshire and E. S. Mills (Eds.), *Handbook of Regional and Urban Economics*, Volume 3 of *Handbook of Regional and Urban Economics*, Chapter 41, pp. 1595–1635. Elsevier.

Stull, W. J. and J. C. Stull (1991). Capitalization of local income taxes. *Journal of Urban Economics 29*(2), 182–190.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy 64*, 416.

Tables and figures

Figure 1: DISTRIBUTION OF RESIDENCES
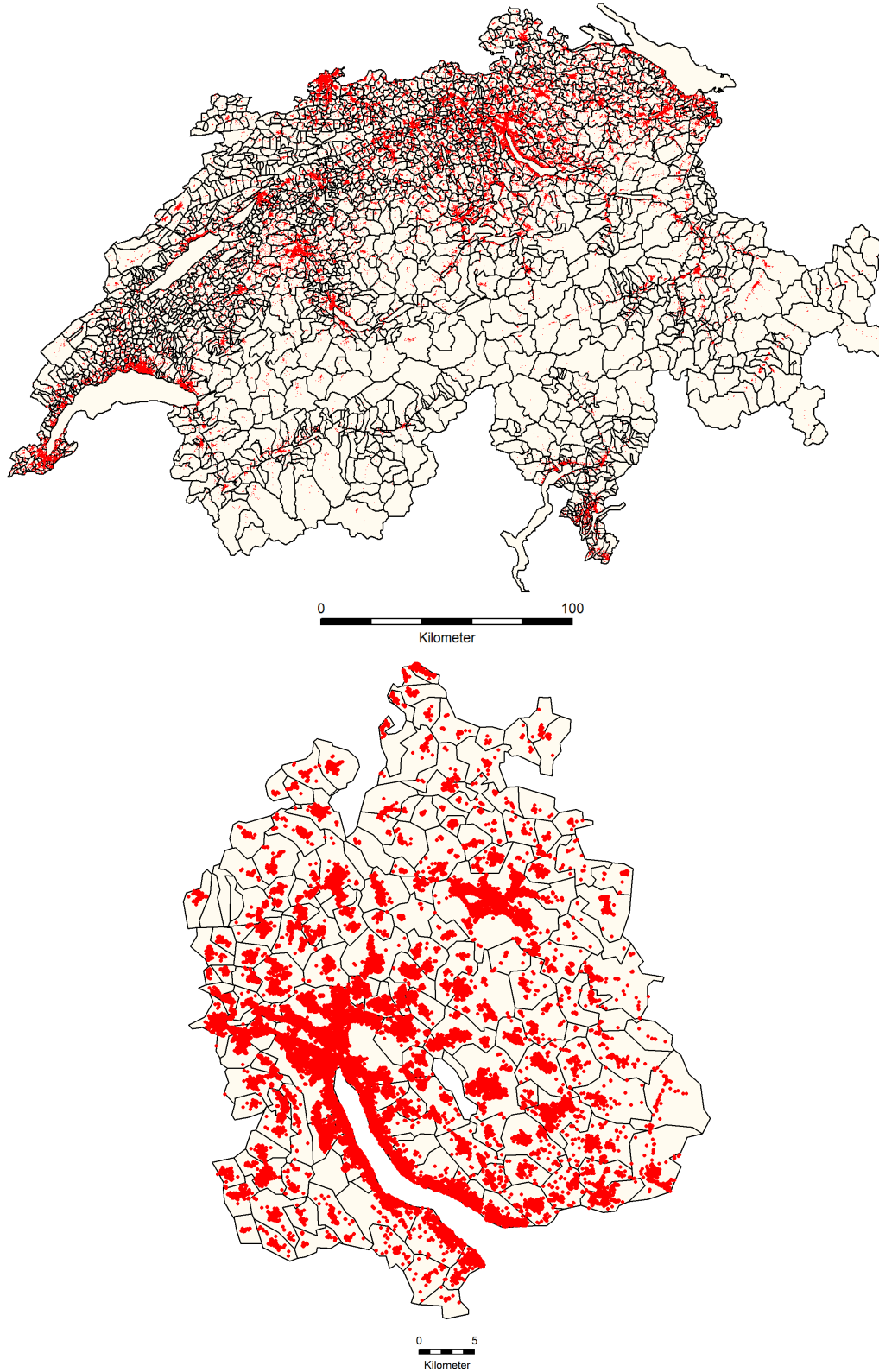
Note: Each dot refers to one residence for which we observe a posting containing information on the rent per square meter, and on all covariates listed in Table 1. The lower map focuses on the Canton of Zurich.

Figure 2: INCOME TAX BURDEN ACROSS SWISS MUNICIPALITIES
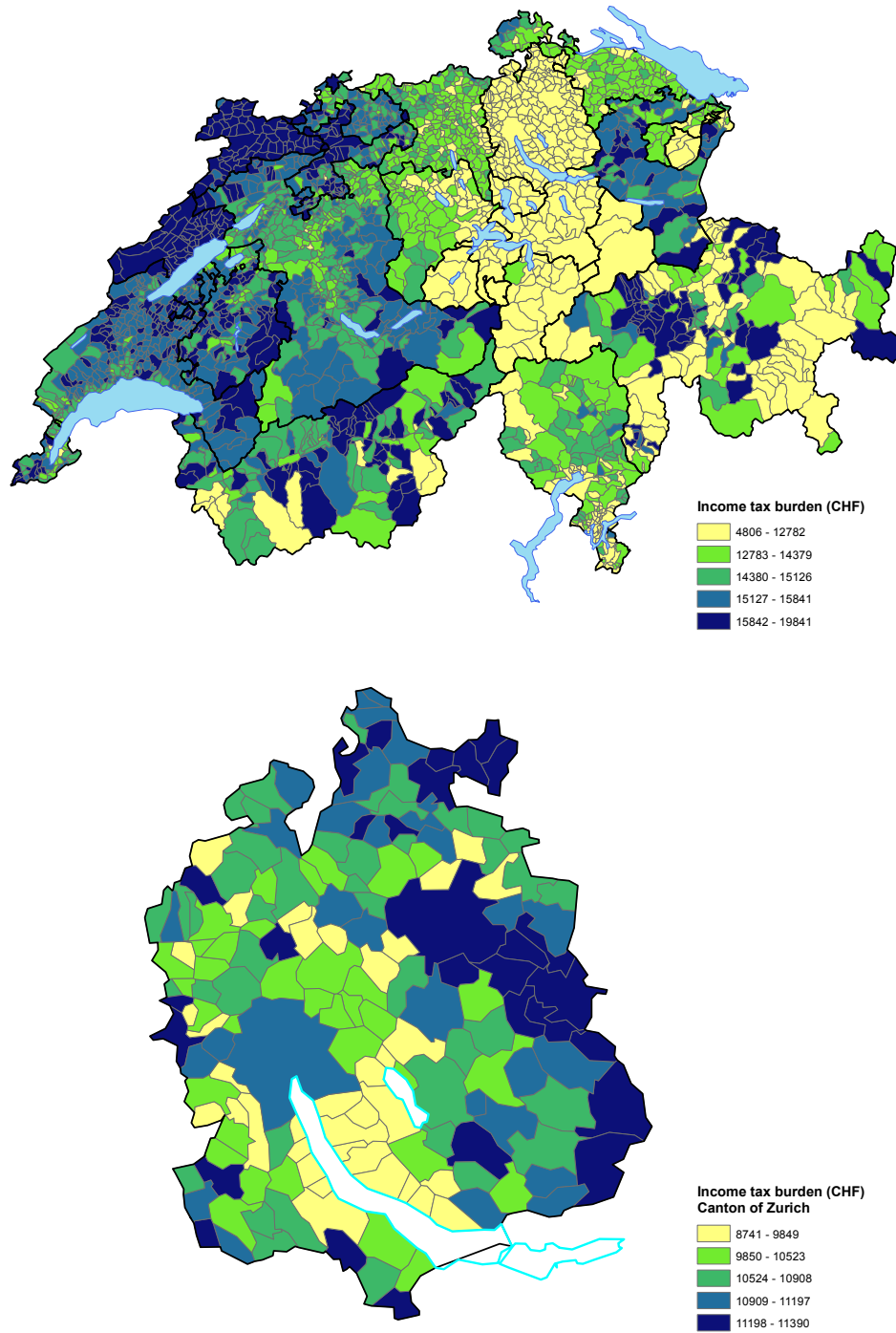
Note: The colors refer to the quintiles of the distribution of income tax burden in Switzerland and in the Canton of Zurich. Lighter colors correspond to lower tax burden, darker colors to higher tax burden. The tax burden is calculated for a single household with an annual gross income of 100,000 CHF in the year 2012.

Figure 3: Rents across Swiss Municipalities

Average rent per sqm (CHF)
- 5 - 15
- 16 - 17
- 18 - 19
- 20 - 22
- 23 - 74

**Average rent per sqm (CHF)
Canton of Zurich**
- 14 - 19
- 20 - 21
- 22 - 23
- 24 - 26
- 27 - 41

Note: The colors refer to the quintiles of the distribution of rents per square meter in 2012 where the upper map plots Switzerland and the lower map focuses on the Canton of Zurich. Lighter colors correspond to lower rents, darker colors to higher rents. We consider only border points that are assigned at least 2 residences. Municipalities that do not feature a border point with sufficient observations are dropped from the analysis and are marked white areas in the above maps.

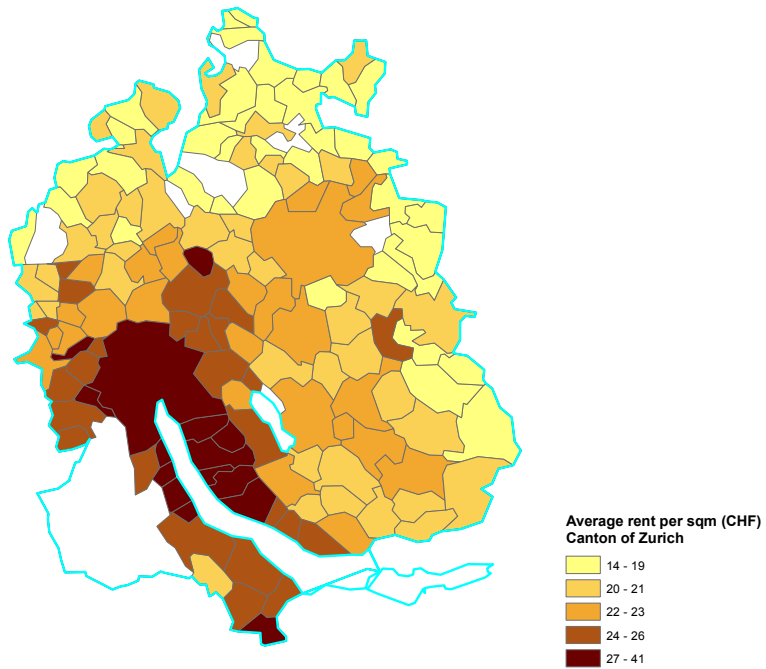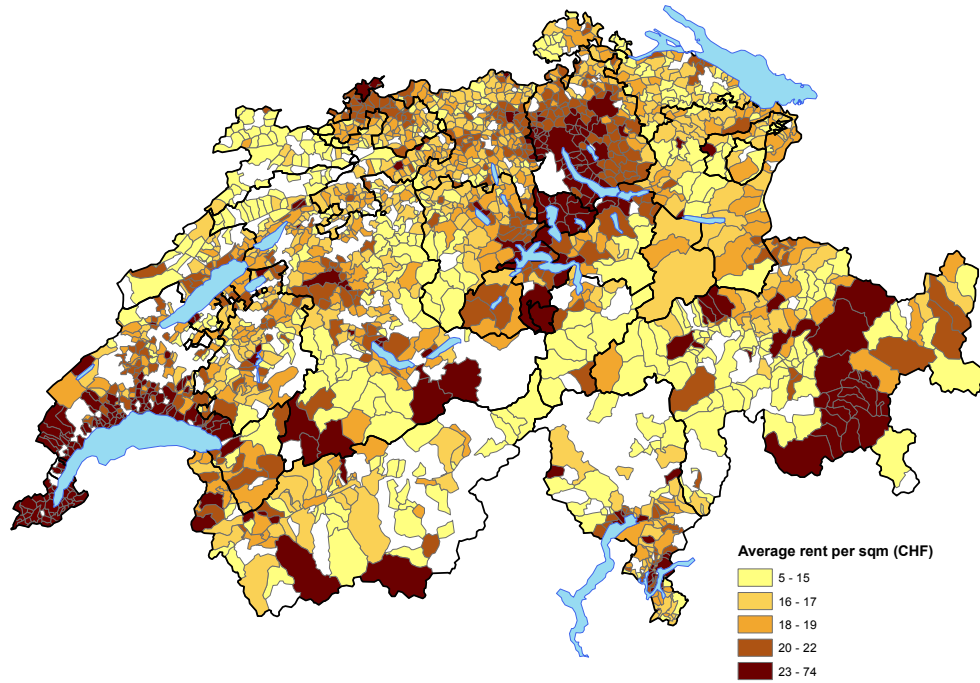Figure 4: Residences and Border Points – Example: Zollikon/Küsnacht
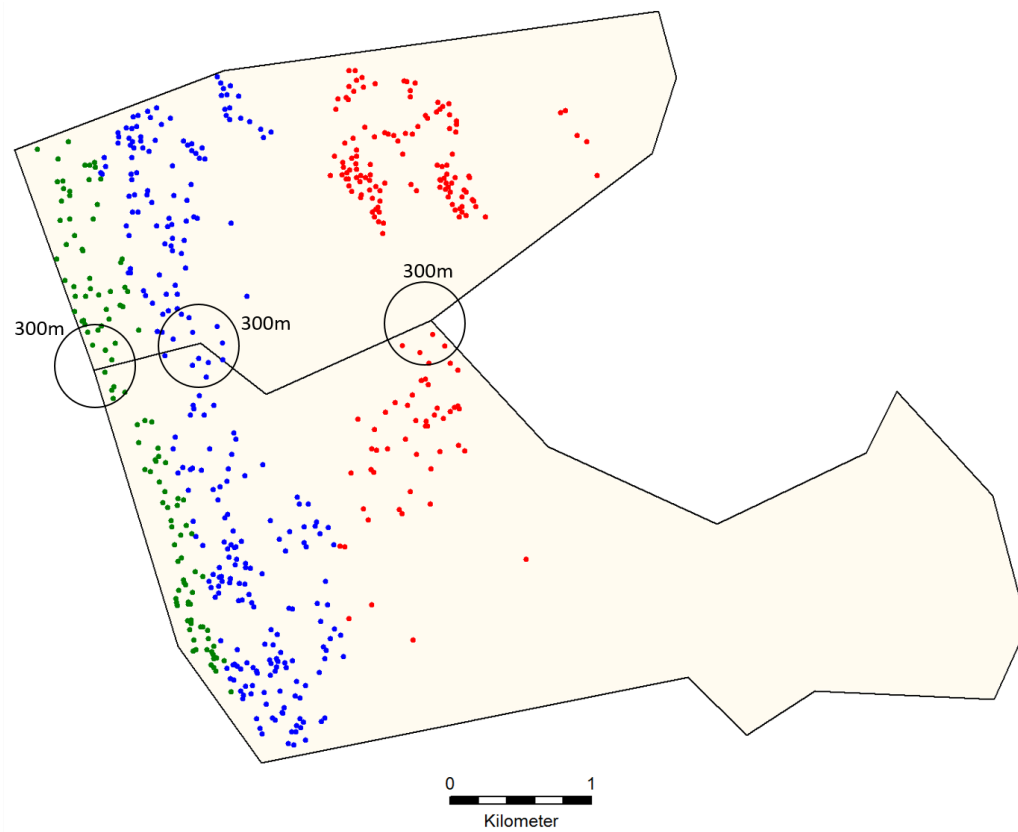
Note: Each dot refers to one residence for which we observe a posting containing information on the rent per square meter, and on all covariates listed in Table 1. The colors of the dots indicate the border point that residences were assigned to on the basis of the minimum distance. Note that Zollikon and Küsnacht are two municipalities in the Canton of Zurich which are situated at the lake of Zurich. Residences marked in green (border point 1) are very close to the lake shore while residences marked in red (border point 3) are on a hill.

Figure 5: Taxes, Rents, and the Municipal Border
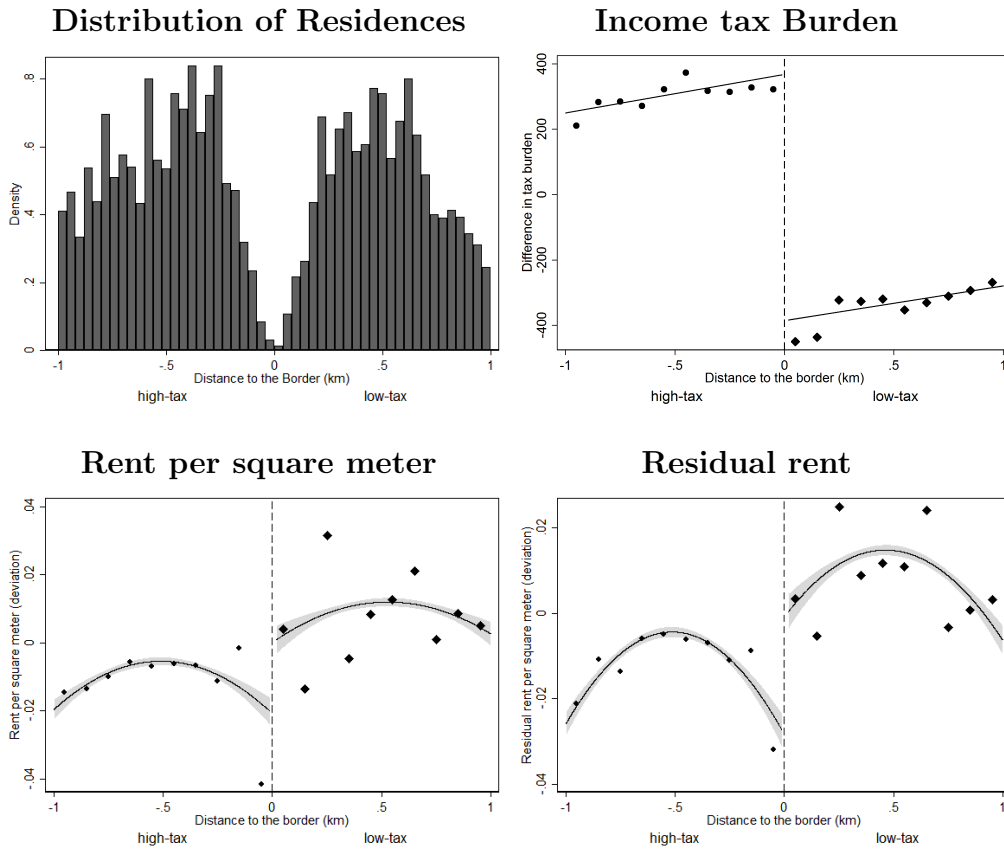
Note: Treated units (lower taxes) are assigned positive distances while control units (higher taxes) are assigned negative distances. We use pooled data for 2005-2012. The figures show local averages of variables in equally sized bins of 100 meters which are plotted against the distance from the respective boundary point within 1 km. Rents per square meter and residual rents are measured in logarithmic terms, and the tax burden in absolute terms. All variables are measured in terms of the deviation from the respective border-point-year average. We fit linear predictions for income tax burden and quadratic functions for rent per square meter and residual rent on both sides of the threshold and display the corresponding 95% confidence intervals.

Figure 6: TAXES, SOCIODEMOGRAPHIC VARIABLES, AND THE MUNICIPAL BORDER
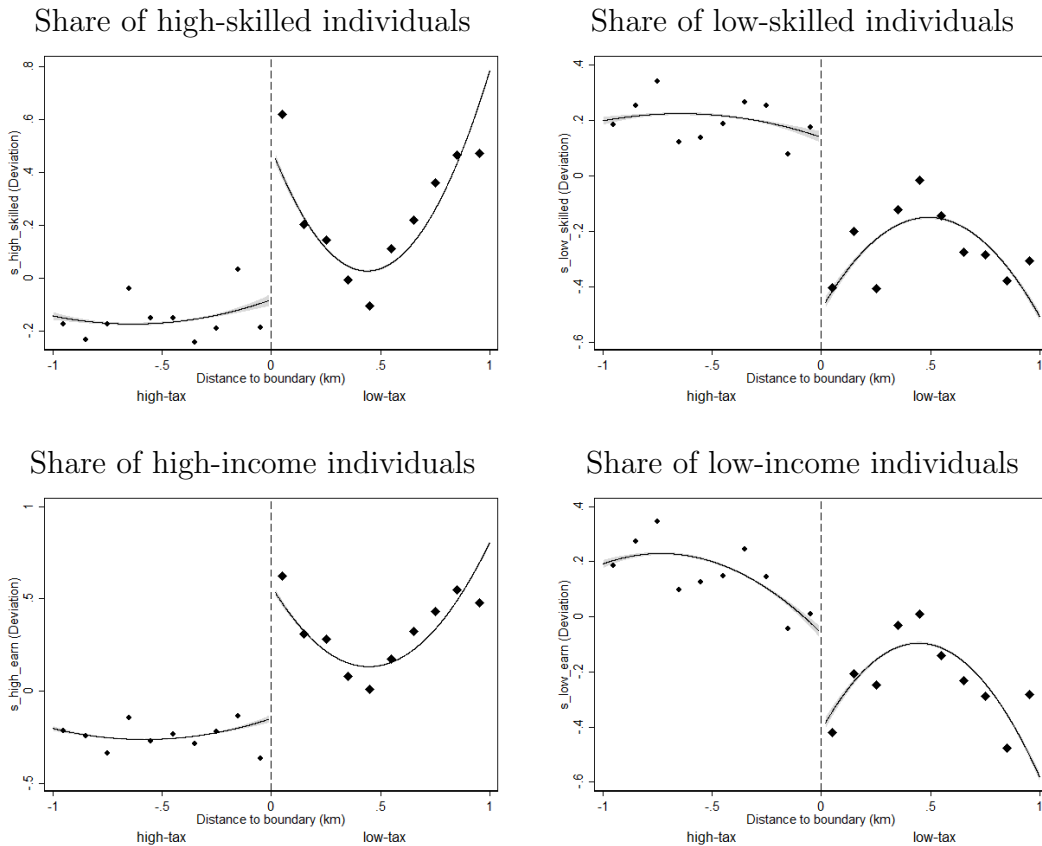
Note: Treated units (lower taxes) are assigned positive distances while control units (higher taxes) are assigned negative distances. The shares are computed on the basis of the population in 1×1 km grids where we require a minimum population of 50. Each residence i is assigned the sociodemographics of the grid g it is located in. For more details on the classification of high- (low-) income (skill) groups see Appendix B. The figures show local averages of variables in equally sized bins of 100 meters which are plotted against the distance from the respective boundary point within 1 km. Outcomes are measured in shares and as the deviation from the respective boundary-year average. We fit quadratic functions on both sides of the threshold and display the corresponding 95% confidence intervals.

Table 1: Summary statistics

	Full data set		Boundary sub-sample	
	Mean	Std. Dev.	Mean	Std. Dev.
Main Variables				
Rent/m^2 R_i	21.140	6.261	22.247	6.430
Tax burden T_m^{S100}	13,875.760	2,409.827	13,849.775	2,497.446
Distance (meter) D_{ib}	-	-	538.177	244
Sociodemographic Covariates				
Share of high-income$_g$	0.124	0.061	0.139	0.074
Share of low-income$_g$	0.111	0.040	0.112	0.046
Share of high-skilled$_g$	0.100	0.070	0.117	0.077
Share of low-skilled$_g$	0.302	0.088	0.294	0.101
Average age$_g$	39.580	3.496	40.354	3.499
Share of foreigners$_g$	0.239	0.115	0.257	0.123
Share of unemployed$_g$	0.032	0.017	0.032	0.015
Share of commuters$_g$	0.439	0.049	0.433	0.042
Population$_g$	3604.573	3589.298	2811.350	2038.478
Residence Covariates				
Rooms$_i$	3.528	1.324	3.507	1.325
Age$_i$	32.585	35.746	40.354	3.499
Floor$_i$	1.933	1.696	1.916	1.755
Parquet$_i$	0.376	0.484	0.367	0.482
Cellar$_i$	0.175	0.380	0.175	0.380
Attic$_i$	0.041	0.199	0.040	0.195
Garden$_i$	0.098	0.297	0.105	0.307
Terrace$_i$	0.078	0.268	0.077	0.266
Balcony$_i$	0.653	0.476	0.667	0.471
Lift$_i$	0.370	0.483	0.347	0.476
Renovated$_i$	0.078	0.269	0.081	0.273
Carport$_i$	0.108	0.310	0.098	0.298
Lake view$_i$	0.013	0.113	0.035	0.184
Furnished$_i$	0.042	0.201	0.065	0.247
Bathtub$_i$	0.040	0.195	0.045	0.207
Dryer$_i$	0.048	0.214	0.042	0.200
Washingm$_i$	0.058	0.234	0.048	0.214
Dishwash$_i$	0.155	0.362	0.160	0.366
Ceramicstove$_i$	0.204	0.403	0.213	0.409
Openkitchen$_i$	0.049	0.215	0.043	0.202
Floorheat$_i$	0.012	0.111	0.010	0.099
Fireplace$_i$	0.050	0.218	0.061	0.240
Municipalities M	1,989		211	
Grids G	8,399		701	
Border points B	6,193		190	
Residences I	2,516,518		151,011	

Notes: The summary statistics correspond to data pooled over the period 2005-2012. We dropped observations with a rent of less than 5 CHF or more than 52 CHF per square meter, and observations with missing information on rents, the number of rooms, and square meters. This diminishes the sample from about 3.5mn. observations to 2,516,518 observations in the full dataset. The boundary sub-sample consists of all residences that could be assigned to a border point b which features at least 2 observations on both sides of the boundary and within a distance of 300 meters. Accordingly, the municipality composition remains stable when restricting the boundary sub-sample to residences at maximum distances of 600 and 300 meters. All border points which correspond to geographic barriers (e.g. rivers) or linguistic barriers have been dropped. Moreover, we drop municipality pairs with a tax differential of less than 100 CHF and we keep only residences that lie within a 1 × 1 km grid g with a population of at least 50 individuals. The latter should ensure that the sociodemographic information is representative of the neighborhood. Tax burden and rents are measured in CHF. T_m^{S100} refers to the tax burden of a single household with an annual income of 100,000 CHF. For a detailed definition of the sociodemographic variables and the tax burden see Appendices A and B.

Table 2: BDD for municipal taxes and rental prices 2005-2012

A. Log rent per m²

	OLS		BDD 1km		BDD 0.6km		BDD 0.3km	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tax elasticity	-0.489***	-0.503***	-0.369***	-0.377***	-0.345***	-0.274***	-0.248**	-0.282**
	(0.062)	(0.035)	(0.044)	(0.063)	(0.062)	(0.082)	(0.115)	(0.142)
Obs.	260,454	165,332	151,011	120,439	89,058	75,218	30,980	26,470
No. b,t	–	–	605	522	605	522	605	522
Agglom.	Yes	No	Yes	No	Yes	No	Yes	No
adj. R^2	0.124	0.186	0.506	0.505	0.521	0.509	0.547	0.551

B. Residual rent

	OLS		BDD 1km		BDD 0.6km		BDD 0.3km	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tax elasticity	-0.540***	-0.525***	-0.463***	-0.450***	-0.456***	-0.383***	-0.258***	-0.311***
	(0.052)	(0.035)	(0.046)	(0.064)	(0.064)	(0.079)	(0.100)	(0.117)
Obs.	260,454	165,332	151,011	120,439	89,058	75,218	30,980	26,470
No. b,t	–	–	605	522	605	522	605	522
Agglom.	Yes	No	Yes	No	Yes	No	Yes	No
adj. R^2	0.143	0.185	0.465	0.481	0.484	0.487	0.522	0.541

***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively. Standard errors are clustered at the municipality-year level in columns (1) and (2) and at the border-point-year level in columns (3)-(8). The first two columns are estimated by OLS (including year fixed effects) and columns (3)-(8) refer to BDD estimates (including border-point-year fixed effects). Each specification is reported for a sample of all observations in uneven columns and one excluding the 5 largest agglomerations in even columns covering the years 2005-2012. We account for 605 and 522 boundary-year groups in the samples with and without agglomerations, respectively. Residual rents are obtained from a regression of the log rent per square meter on the residence covariates listed in Table 1 and interactions between them.

Table 3: MBDD for municipal taxes and rental prices 2005-2012

A. Log rent per m²

	1km Sample		0.6km Sample		0.3km Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Tax elasticity	-0.284***	-0.263***	-0.305***	-0.263***	-0.253***	-0.257**
	(0.052)	(0.073)	(0.067)	(0.084)	(0.097)	(0.116)
Obs.	151,011	120,439	89,058	75,218	30,980	26,470
No. b,t	605	522	605	522	605	522
Agglom.	Yes	No	Yes	No	Yes	No
adj. R²	0.015	0.011	0.016	0.012	0.009	0.008
Avg.dist	1.019	0.988	0.717	0.710	0.411	0.419

B. Residual rent

	1km Sample		0.6km Sample		0.3km Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Tax elasticity	-0.389***	-0.363***	-0.395***	-0.348***	-0.264***	-0.265**
	(0.051)	(0.068)	(0.066)	(0.080)	(0.090)	(0.103)
Obs.	151,011	120,439	89,058	75,218	30,980	26,470
No. b,t	605	522	605	522	605	522
Agglom.	Yes	No	Yes	No	Yes	No
adj. R²	0.025	0.019	0.025	0.018	0.010	0.009
Avg.dist	1.019	0.988	0.717	0.710	0.411	0.419

***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively. The sample covers the years 2005-2012. Standard errors are clustered at the border-point-year level. All regressions include year fixed effects and are estimated by WLS. The reference observations are spatially differenced with respect to a counterfactual outcome (according to equation (3)) at 605/522 border points. Weights correspond to the inverse distance to the counterfactual (the average distance between the reference unit and all residences used to construct the counterfactual residence). Each specification is reported for a sample of all observations in uneven columns and one excluding the 5 largest agglomerations in even columns. Residual rents are obtained from a regression of the log rent per square meter on the residence covariates listed in Table 1 and interactions between them.

Table 4: Sociodemographic sorting

A. Log rent per m²

| | BDD | | MBDD | |
| | 0.6km | 0.3km | 0.6km | 0.3km |
	(1)	(2)	(3)	(4)
Tax elasticity	-0.333***	-0.226*	-0.244***	-0.181*
	(0.070)	(0.126)	(0.076)	(0.108)
Obs.	89,058	30,980	89,058	30,980
No. b,t	605	605	605	605
Socio-demog.	Yes	Yes	Yes	Yes
adj. R^2	0.524	0.553	0.052	0.048
Avg.dist			.717	.411

B. Residual rent

| | BDD | | MBDD | |
| | 0.6km | 0.3km | 0.6km | 0.3km |
	(1)	(2)	(3)	(4)
Tax elasticity	-0.405***	-0.255**	-0.308***	-0.199**
	(0.068)	(0.102)	(0.069)	(0.097)
Obs.	89,058	30,980	89,058	30,980
No. b,t	605	605	605	605
Socio-demog.	Yes	Yes	Yes	Yes
adj. R^2	0.486	0.526	0.055	0.037
Avg.dist			.717	.411

***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively. The sample covers the years 2005-2012 and includes agglomerations. Standard errors are clustered at the border-point-year level. Columns (1) and (2) refer to BDD estimates whereas columns (3) and (4) correspond to the MBDD specifications and are estimated by WLS. Weights correspond to the inverse distance to the counterfactual (the average distance between the reference unit and all residences used to construct the counterfactual residence). Residual rents are obtained from a regression of the log rent per square meter on the residence covariates listed in Table 1 and interactions between them. We include linear terms of all sociodemographic covariates listed in Table 1. These variables are computed on the basis of the population of 1×1 km grids. We locate each residence i in the grid and assign the corresponding sociodemographic covariates.

Table 5: (M)BDD with schooling

A. Log rent per m^2

	BDD		MBDD	
	(1)	(2)	(3)	(4)
Tax elasticity	-0.498**	-0.433*	-0.420*	-0.306#
	(0.227)	(0.226)	(0.226)	(0.209)
Obs.	21,014	21,014	21,014	21,014
No. b,t	151	151	151	151
Socio-demog.	No	Yes	No	Yes
adj. R^2	0.434	0.444	0.028	0.103
Avg.dist			0.689	0.689

B. Residual rent

	BDD		MBDD	
	(1)	(2)	(3)	(4)
Tax elasticity	-0.404***	-0.285*	-0.323*	-0.244#
	(0.144)	(0.154)	(0.164)	(0.157)
Obs.	21,014	21,014	21,014	21,014
No. b,t	151	151	151	151
Socio-demog.	No	Yes	No	Yes
adj. R^2	0.351	0.359	0.025	0.088
Avg.dist			0.689	0.689

***, **, *, # denote statistical significance at the 1%, 5%, 10% and 15% level, respectively. The sample covers the years 2005-2012. Standard errors are clustered at the municipality-year level. All regressions are based on the sample with agglomerations and a distance limited to 600 meters from the boundary point. This yields 150 border-point-year groups of units that lie within the same school district but face different municipal income taxes. Columns (1) and (2) refer to the BDD estimates whereas columns (3) and (4) correspond to the MBDD specifications which are estimates by WLS. Weights correspond to the inverse distance to the counterfactual (the average distance between the reference unit and all residences used to construct the counterfactual residence). Residual rents are obtained from a regression of the log rent per square meter on the residence covariates listed in Table 1 and interactions between them. Each specification is reported without (in uneven columns) and with (in even columns) sociodemographic covariates. These are listed in Table 1.

Table 6: Placebo tests

	Border shifted towards treated		Border shifted towards control	
	Log rent/m^2	Residual rent	Log rent/m^2	Residual rent
Tax elasticity	0.152	-0.041	0.104	-0.022
	(0.110)	(0.086)	(0.089)	(0.089)
Obs.	54,472	57,548	54,472	57,548
No. b,t	522	522	522	522
adj. R^2	0.551	0.514	0.519	0.498

All standard errors are clustered at the border-point-year level. Regressions are based on the sample without agglomerations and a distance limited to 600 meters from the border point.

Table 7: SENSITIVITY (I)

A. Log rent per m²

	Distance	Coordinates	t-squared	t-Median	Stable	Municipality FE	Married & 2 Children
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tax elasticity	-0.262***	-0.238**	-0.262***	-0.272***	-0.239**	-0.294*	-0.254***
	(0.096)	(0.096)	(0.095)	(0.097)	(0.101)	(0.164)	(0.097)
Tax elasticity²			1.075**				
			(0.464)				
Obs.	30,980	30,980	30,980	13,481	28,129	30,980	30,980
No. b,t	605	605	605	248	594	605	605
adj. R²	0.014	0.014	0.012	0.023	0.010	0.213	0.009
Avg.dist	0.411	0.411	0.411	0.411	0.411	0.411	0.411

B. Residual rent

	Distance	Coordinates	t-squared	t-Median	Stable	Municipality FE	Married & 2 Children
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tax elasticity	-0.277***	-0.246***	-0.272***	-0.273***	-0.242***	-0.298**	-0.264***
	(0.090)	(0.086)	(0.088)	(0.091)	(0.093)	(0.150)	(0.090)
Tax elasticity²			0.994**				
			(0.444)				
Obs.	30,980	30,980	30,980	13,481	28,129	30,980	30,980
No. b,t	605	605	605	248	594	605	605
adj. R²	0.015	0.017	0.013	0.026	0.010	0.191	0.010
Avg.dist	0.411	0.411	0.411	0.411	0.411	0.411	0.411

***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively. The sample covers the years 2005-2012. All standard errors are clustered at the border-point-year level. All regressions include year fixed effects and are estimated by WLS. The reference observations are spatially differenced with respect to a counterfactual outcome (according to equation (3)) at 605/522 border points. Weights correspond to the inverse distance to the counterfactual (the average distance between the reference unit and all residences used to construct the counterfactual residence). Residual rents are obtained from a regression of the log rent per square meter on the residence covariates listed in Table 1 and interactions between them. Regressions are based on a distance limited to 300 meters from the border point. Column (1) includes a cubic function of distance from the border point, column (2) includes x- and y-coordinates, column (3) includes a squared term of the tax differential, column (4) compares observations with above median tax differential (i.e., >731 CHF); column (5) shows the results based on a sample of municipalities that are stable over time; column (6) includes municipality fixed effects in addition to border-point-year fixed effects; and column (7) focuses on married households with two children and an annual gross income of 100,000 CHF.

Table 8: SENSITIVITY (II)

A. Log rent per m^2

	1. Quartile	2. Quartile	3. Quartile	4. Quartile
	(1)	(2)	(3)	(4)
Tax elasticity	-0.357**	-0.366***	-0.274**	-0.189
	(0.158)	(0.126)	(0.117)	(0.122)
Obs.	8,055	8,239	7,806	6,880
No. b,t	520	513	547	540
adj. R^2	0.017	0.023	0.019	0.013
Avg.dist	0.411	0.411	0.411	0.411

B. Residual rent

	1. Quartile	2. Quartile	3. Quartile	4. Quartile
	(1)	(2)	(3)	(4)
Tax elasticity	-0.414***	-0.291**	-0.280***	-0.160
	(0.130)	(0.137)	(0.103)	(0.120)
Obs.	8,055	8,239	7,806	6,880
No. b,t	520	513	547	540
adj. R^2	0.029	0.013	0.023	0.011
Avg.dist	0.411	0.411	0.411	0.411

***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively. The sample covers the years 2005-2012. All standard errors are clustered at the border-point-year level. All regressions include year fixed effects and are estimated by WLS. The reference observations are spatially differenced with respect to a counterfactual outcome (according to equation (3)) at 605/522 border points. Weights correspond to the inverse distance to the counterfactual (the average distance between the reference unit and all residences used to construct the counterfactual residence). Residual rents are obtained from a regression of the log rent per square meter on the residence covariates listed in Table 1 and interactions between them. Regressions are based on a distance limited to 300 meters from the border point. Columns (1)-(4) show the results for quartiles of apartment size (1st qu.: 15-61, 2nd qu.: 61.5-82, 3rd qu.: 82.3-107, 4th qu.: 107.3-230 square meters).

Appendix A: Data on municipal taxes

There are numerous differences in the calculation of taxes across the 26 cantons. These include differences in deductions, in the cantonal tax rates (progressive versus flat rate), the character of the municipal and church tax rates (percentage, multiple), and the tax base (*Einfache Steuer*, cantonal tax, taxable income).

The total income tax from cantonal, municipal and church taxes was calculated by canton and year (2005-2012) as follows. We used figures published by *Eidgenössische Steuerverwaltung* (ESTV) for all municipalities with at least 2,000 inhabitants before 2010 and full information for all municipalities for 2010-12 to trace back the so-called canton-specific *Einfache Steuer* which sets the basis for the calculation of the local tax burden. For most cantons, the formula is as follows.

$$\text{Base} = \frac{1}{G} \sum_{g=1}^{G} \frac{\text{Tax} - \text{Poll tax}}{\text{tcant} + \text{tmun} + \text{tchurch}},$$

where *Base* is the *Einfache Steuer*, G denotes the number of municipalities, Tax is the local tax burden published by ESTV, *Poll tax* is a canton- or municipality-specific poll tax, *tcant* is the cantonal, *tmun* the municipal, and *tchurch* the church tax rate.

We used the church tax rate according to the majority confession by municipality using 2000 census data (BFS). All calculations are done by year and canton. We obtained the income tax as the sum of the three and add the poll tax if applicable for all years 2005-2012. The income tax was calculated for different levels of annual gross income (80, 100, 150, 200, 400 tsd. CHF) and for two different types of households (single households and married households with two children). The sources were as follows: cantonal tax rates (ESTV), municipal tax rates (cantonal tax or statistical offices), church tax rates (cantonal tax or statistical offices or local parishes), poll tax (cantonal tax offices). The latter were levied by the Cantons of Geneva, Lucerne, Nidwalden, Schaffhausen, Solothurn, Uri, Vaud, and Zurich. They differed by municipality in the Cantons of Vaud and Solothurn. We generally neglected federal income taxes (note that these are irrelevant anyway as we focus solely on within-canton comparisons).

Exceptions are explained in the following. Calculations were adjusted for the Cantons of Appenzell-Innerrhoden and Graubünden, where municipal taxes were expressed in % of the cantonal tax, for the Canton of Valais,

where the church tax was expressed in % of the municipal tax, and for the Canton of Jura, where the church tax was expressed in % of the cantonal tax. *Base* was then multiplied accordingly by the cantonal, municipal, and church tax rates.

In case municipalities within canton reported non-uniform and positive school and/or district tax rates (Schwyz, Zug, Appenzell-Innerrhoden, Thurgau), we added these tax rates to the municipal tax rate. Note that municipalities were always a subset of districts. Three municipalities did not levy income taxes but instead only a school tax (Einsiedeln, Küssnacht am Rigi, Gersau) from 2010 onwards. Only 38 administrative municipalities in 2 cantons (Appenzell-Innerrhoden and Thurgau) consisted of more than one (at most 5, on average 2.2) school municipalities *and* reported different school taxes within the administrative municipality (note that there is more municipalities but these do not charge varying school taxes).

In the Canton of Uri, taxes were calculated as a percentage of the taxable income until 2008. To obtain the tax burden, we first regressed the tax burden minus the poll tax on the municipal and the church tax rate and on a constant and obtained the *Einfache Steuer* from the latter in all years before 2009. This procedure is in line with Parchet (2012). We estimated these regressions by year. We then calculated the total income tax burden from cantonal, municipal and church taxes for all municipalities from the following formula:

$$\text{Tax} = \text{Poll tax} + \text{Base} \times \text{tcant} + \text{taxable income} \times (\text{tmun} + \text{tchurch}).$$

The taxable income was obtained by subtracting deductions for public and private insurance and pension contributions, business and general expenses. We assumed a typical number of deductions to proxy the taxable income and in a next step, calculated the total tax burden due. Specifically, deductions for public insurance schemes were 7.1% of the gross annual income (5%, 1.1%, and 1% thereof for public pension, unemployment, and accident insurance, respectively), resulting in net income. Deductions for private pension schemes amounted to 6,192 CHF in the years 2005 and 2006, and to 6,365 in 2007 and 2008. In addition, a general insurance deduction of 1,500 CHF was subtracted. Deductions for general expenses amounted to 2000 CHF. Finally, we deducted business expenses by utilizing the minimum of 3% of the net income and 3800 CHF.

For the Canton of Basel-Stadt, we used the income tax burden as published by ESTV in all years and for all three municipalities of the canton.

Appendix B: Sociodemographic data

Sociodemographic variables stem from the 2000 micro census and were kindly provided by the Swiss Federal Statistical Office (BFS). We obtained xy-coordinate-based individual data (anonymized) that we aggregate within spatial grids of 1×1 km. The variables are constructed as follows.

Share of high-income. The share of high-income individuals in the total laborforce (working plus nonworking population). To construct this variable, we used information on socio-professional categories as defined by BFS. High-income individuals are defined as individuals whose job function is top management, independent professions (e.g. physicians, lawyers, architects), other self-employed excluding agriculture, academic jobs, and managerial staff (Oberstes Management, Freie Berufe, Andere Selbständige ohne Landwirtschaft, Akademische Berufe und oberes Kader). We calculate the sum of high-income individuals within the grid and divide it by the total laborforce within the grid to obtain the share.

Share of low-income. The share of low-income individuals in the total laborforce. Low-income individuals are defined as individuals occupying non-skilled job functions (ungelernte Angestellte und Arbeiter). The calculation of the share is otherwise analogous to the share of high-income.

Share of high-skilled. The share of high-skilled individuals in the adult population. To construct this variable, we used information on the highest completed educational degree. High-skilled individuals are defined as individuals with a college of higher education or university degree. We calculate the sum of high-skilled individuals within the grid and divide it by all individuals (excluding missing observations, school-aged, or younger) within the grid to obtain the share.

Share of low-skilled. The share of low-skilled individuals in the adult population. To construct this variable, we used information on the highest completed educational degree. Low-skilled individuals are defined as individuals without an educational degree and completed compulsory schooling. The calculation of the share is otherwise analogous to the share of high-skilled.

Average age. The average of the age of all individuals within the grid.

Share of foreigners. The share of non-native born in the total population. We calculate the sum of foreigners within the grid and divide it by the population within the grid.

Share of unemployed. The share of unemployed in the total laborforce. We calculate the sum of unemployed within the grid and divide it by the total laborforce within the grid.

Share of commuters. The share of commuters in the total adult population. We calculate the sum of commuters within the grid and divide it by the population within the grid.

Population. The total number of individuals in the census (7,452,075).