THE
ECONOMIC
JOURNAL

The Economic Journal, **127** (*May*), 653–687. Doi: 10.1111/ecoj.12489 © 2017 Royal Economic Society. Published by John Wiley & Sons, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

INCOME TAXES, SORTING AND THE COSTS OF HOUSING: EVIDENCE FROM MUNICIPAL BOUNDARIES IN SWITZERLAND*

Christoph Basten, Maximilian von Ehrlich and Andrea Lassmann

We provide novel evidence on the role of income taxes for housing rents and spatial sorting. Drawing on comprehensive micro-level data, we estimate the responsiveness of households to tax differentials across municipal boundaries. Correcting for unobservable location characteristics and isolating the residential sorting component, we identify an income tax elasticity of rents of about −0.27 to −0.35. In line with non-homothetic preferences, we find that the marginal willingness to pay for lower taxes increases with income. Counterfactual calculations show how an homogenisation of taxes across jurisdictions and an increase in variation of taxes affect rents and income stratification across space.

The question whether households 'vote with their feet' in response to tax incentives is of great importance for all levels of government. On the one hand, household mobility induces competition between jurisdictions and encourages governments to minimise the tax burden for a given bundle of public goods. On the other hand, it limits governments' leeway for redistributive policies as well as for offering public goods with positive spillovers for other jurisdictions (Tiebout, 1956; Gordon, 1983). One way to quantify household responsiveness is through the extent to which lower taxes capitalise in higher house prices (Oates, 1969; Hilber, 2011).

This article's contribution is twofold. First, we provide novel estimates of income tax elasticity using a boundary discontinuity design (BDD; Black, 1999) that is incorporated in a discrete choice model where households sort into local jurisdictions according to income and heterogenous tastes over housing, taxes and local neighbourhoods (Bayer *et al.*, 2007). Exploiting variation in income taxes, housing rents and individual incomes offers an advantage compared to previous literature on property taxes and property values because the latter need to make assumptions about discount rates regarding all future expected taxes. Second, we perform two counterfactual tax policy experiments to identify the effects of:

* Corresponding author: Andrea Lassmann, Department of Management, Technology and Economics, KOF Swiss Economic Institute, ETH Zurich, Leonhardstrasse 21, G118, 8092 Zurich, Switzerland. Email: lassmann@kof.ethz.ch.

Ehrlich and Lassmann are indebted to an anonymous reviewer and the editor in charge (Rachel Griffith), for numerous invaluable comments on an earlier version of the manuscript. We thank comparis.ch, the Federal Statistical Office (BFS), the Tax Authority of Bern, the Swiss Federal Price Monitor, and Swisstopo for generously providing data. We further thank the cantonal tax offices and the cantonal education authorities, Gema Ricart (Swiss Federal Tax Administration) and Raphaël Parchet for kindly providing data on local taxes and school boundaries. Ehrlich and Lassmann are grateful for valuable comments from an anonymous referee, Patrick Bayer, Florian Chatagny, Dirk Drechsel, Christian Hilber, Henry Overman, Benedikt Rydzek, Kurt Schmidheiny and Malte Zoubek and benefited from numerous comments of seminar participants at ETH Zurich, the IIPF conference in Lugano, LSE SERC, the Royal Economic Society in Brighton, the Swiss Society of Economics and Statistics in Bern, University of Basel, University of Hamburg and ZEW Mannheim.

(*i*) a homogenisation of taxes and
(*ii*) an increase in the variation of income taxes across jurisdictions on equilibrium rents and neighbourhood stratification according to income.

Since high-income households have a relatively stronger preference for places offering low taxes than low-income households, tax differentials induce a sorting of households across jurisdictions (Schmidheiny, 2006). More precisely, low-tax jurisdictions are typically characterised by higher incomes and the resulting differences in the tax base may simultaneously affect tax policies that accordingly capitalise in the costs of housing. Apart from this, unobservable factors such as geographical amenities, local public goods and services and local neighbourhood attributes influence the attractiveness of jurisdictions and lead to differences in the costs of housing and average income across places (Ioannides, 2004; Ahlfeldt and Holman, 2015). The inherent endogeneity challenges the empirical analysis of the degree of tax capitalisation in house prices and rents.

In order to track these factors, we develop a formal framework which accounts for non-homothetic preferences and provides a starting point for an empirical analysis that draws on unique micro-geographic data sets. These contain detailed information on the universe of residences advertised for rent over the period 2005–12 and across Swiss municipalities, and on individual sociodemographic characteristics for the canton Bern. 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. The precise geo-referencing of the data allows us to analyse the variation in rents as well as in sociodemographic and residence characteristics at a very fine spatial scale. Specifically, the BDD compares residences located close together that share essentially the same amenities, neighbourhood attributes and public goods and services but face different municipal income taxes. Under the assumption that the distribution of potentially confounding variables changes continuously, we can account for the mean indirect utility provided by residences and local neighbourhood compositions and use the jump in the income tax burden across municipal boundaries to identify the mean marginal willingness to pay in terms of rents, for living in a low-tax community.

The estimated average tax elasticity amounts to about -0.27 to -0.35. By contrast, conventional hedonic regressions would suggest a tax elasticity of almost -0.9. Consistent with the assumption of non-homothetic preferences, we find that high-income households have a higher marginal willingness to pay for lower taxes compared to households with lower income. This also holds regarding the marginal willingness to pay for residing in wealthier neighbourhoods and for the moving distance that individuals are willing to incur in order to avoid higher income taxes. The estimated effects of taxes on rents cannot be explained by other factors changing discontinuously at the boundary so that the assumption of continuity of all relevant confounding factors is warranted. The counterfactual homogenisation (increased variation) of income taxes generates an adjustment in equilibrium rents that results in new spatial equilibria which feature an increase (decrease) in residential mixing according to household income.

The remainder of the article is organised as follows: Section 1 reviews the literature on taxes and costs of housing before we introduce the institutional background in Section 2. Section 3 presents the data and in Section 4, we derive an empirical model for inference about the effect of local income taxes on rents. We describe our key results and perform counterfactual analysis in Section 5, followed by numerous sensitivity checks in Section 6. Section 7 concludes with a summary of the key findings.

1. Local Taxes and the Costs of Housing

Following Oates (1969), there has been ample empirical research on the effect of taxes on property values. The dominance of centrally set income taxes has restricted most authors to studying the effect of local property taxes on house prices (Dachis *et al.*, 2012; see Ross and Yinger, 1999 and Hilber, 2011 for excellent overviews) despite the fact that the degree of decentralisation of personal income tax revenues is on average higher than the one of property taxes.[1] By contrast, empirical evidence on the capitalisation of income taxes in house prices and rents is rare (see Stull and Stull, 1991; Feld and Kirchgässner, 1997; Morger, 2013, for exceptions). Most of these studies analyse the valuation of housing characteristics and local taxes in a standard hedonic regression framework, neglecting the endogeneity bias, resulting from unobservable location characteristics. Estimating the response of house prices and rents to tax changes is complicated by the fact that capitalisation 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 sorting may arise for various reasons other than taxes. For instance, households similar in terms of sociodemographic characteristics such as income, education, or cultural background tend to cluster because of environmental and neighbourhood amenities as well as social interactions (Ioannides, 2004; Bayer *et al.*, 2007). Likewise, distance to the central business district matters because different types of households value commuting costs differently. Clustering of wealthy households in localities with more desirable characteristics is certainly relevant in Switzerland (Feld and Kirchgässner, 2001; Schmidheiny, 2006; Schaltegger *et al.*, 2011). Similarly, fixed migration costs imply that high-income individuals are more responsive to tax differences and locate in the low-tax jurisdiction which further drives up rents. As shown by Glazer *et al.* (2008), the empirical relevance of this indirect effect of taxes on the costs of housing has crucial implications for tax competition: when land is scarce and heterogeneous, local utilitarian governments may have less incentives to reduce taxes as the immigration of the rich boosts demand for desirable locations and thus causes the incumbent population to consume inferior locations. Accordingly, sociodemographic variables are intrinsically correlated with unobservable location attributes as well as local taxes. Our identification strategy accounts for this by exploiting the fact that the costs of social interactions are a smooth function of distance such that any discontinuities

[1] Brülhart *et al.* (2014) show in Table A5 that local personal income taxes are levied at least to some extent in Austria, Belgium, Estonia, Finland, Germany, Iceland, Slovenia, Spain, Switzerland, the US (OECD), Bosnia-Herzegovina, Brazil, Bulgaria, Latvia, Lithuania and Russia (non-OECD). The local tax revenue as a share of general government tax revenue from personal income taxes is on average 4.2, compared to 3.1 for property taxes. It equals up to 28.1% (Finland).

must be due to fiscal variables. It furthermore sheds light on the role of income sorting.

2. Institutional Background: Taxes, Public Goods and Housing Rents in Switzerland

By the Constitution of the Swiss Confederation, both the federal state and the 26 cantons (which are similar to US States and German *Bundesländer*) are fiscal jurisdictions. 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 determined periodically by referenda. While the federal government levies the highest share of total tax revenue (45%), it relies mostly on indirect taxes; income taxes amount to only 15% of its revenues. 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.[2] In particular, the Swiss Constitution allows cantons to decide on their own income tax schemes, including the degree of tax progression. As a consequence, heterogeneity across cantons is large. Notably, municipalities 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 rents that are endogenously determined through differences in income taxes across jurisdictions which can be traced back to heterogenous preferences for neighbourhood characteristics and income sorting. The reasons are as follows: first, inter-jurisdictional competition is limited through systems of federal and cantonal fiscal equalisation schemes in practice.[3] Second, 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. That schooling and other public goods and services that may differ at municipality boundaries are negligible in size as we show later on. Third, the large degree of regional autonomy gives rise to yardstick competition as tax rates are decided by referendum, ruling out large-scale inefficiencies as a source of differences in income taxes. Finally, the Swiss tenancy law is flexible and rents can usually be adjusted at least once a year.[4]

[2] 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 for fiscal transfers and social security account for the remaining 4% of total expenditures.

[3] For the purpose of our study only within-canton schemes are relevant which do by no means eliminate all differences in fiscal capacity. Rühli (2013) shows that in the average canton inter-municipality transfers amount to less than 20% of municipal tax revenues.

[4] In most cases, rents can be adjusted on 1 April and 1 October. The law allows an adjustment to rents customary at a place (i.e. local differences within municipalities apply) regarding existing and new tenants. Still the stock of rental contracts is likely to be less responsive to tax changes than the data on rental offer prices. Accordingly, our estimates may correspond to an upper bound of the tax elasticity of all properties.

3. Data

We use three main data sources that are combined with geographic information about Switzerland. First, the data on rents comprise about 3.5 million postings of residences offered for rent during the period 2005–12. These data stem from comparis.ch which is the most widely-used price comparison service in Switzerland. For their real estate platform comparis.ch collects all offers posted on the 17 most popular flat search engines in the Swiss market. The median residence in our data has 3.5 rooms and a living space of about 80 metre^2. 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 (corresponding to 18% of the total Swiss population). 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%.[5] The overall distribution of residences in our sample is shown in Figure 1.

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 and transaction prices, we focus on rents rather than

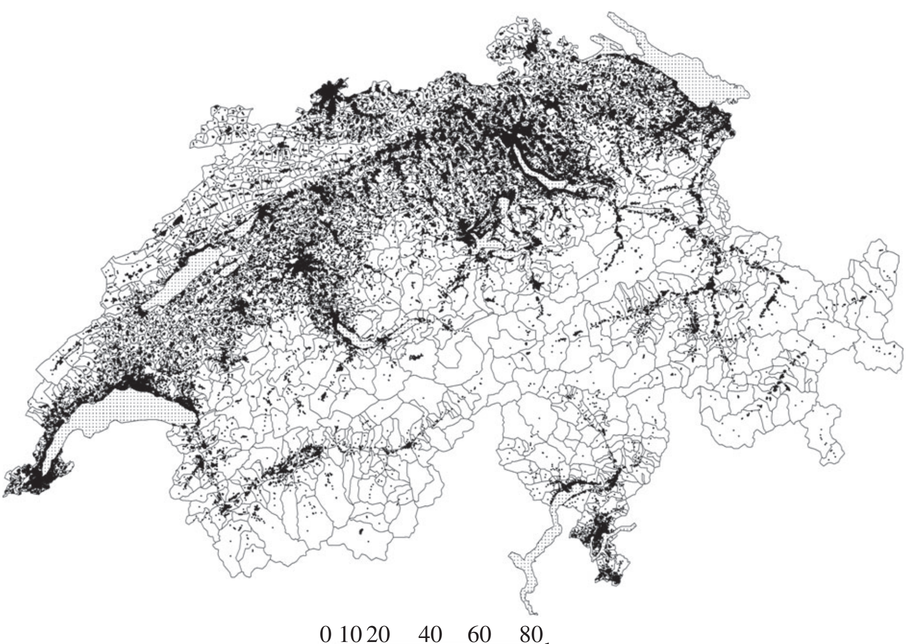

Fig. 1. *Distribution of Residences*
Note. Each dot refers to one residence for which we observe a posting containing information on the rent per metre^2 and on all covariates listed in Table 1.

[5] Population shares stem from the Swiss Federal Statistical Office (BFS). According to official statistics, the number of inhabited flats was 3,534,508 in 2011.

residential property prices.[6] Factors such as immigration and scarcity of constructible land have led to increasing demand for housing in recent years. At the same time, the supply-side response to the increase in demand was low due to factors such as spatial planning rigidities (see the 2015 OECD Economic Survey of Switzerland for an overview; see also Hilber and Vermeulen, 2016). This shows in a vacancy rate of only 0.95% in 2012 in Switzerland. As a consequence, rents are almost never negotiated but taken as given by the tenant. Importantly, the rental share is high in Switzerland, amounting to 59% of total inhabited accommodation in 2011.[7] Due to ownership-neutral taxation, renting is not restricted to those with lower income or wealth but is attractive also for many highly qualified and mobile households. Focusing on rents rather than on property prices also has a further advantage: prices must incorporate all future expected values of taxes, so that their use requires both an assumption about buyers' discount rates and one about the future trajectory of taxes. For example, a municipality with low current taxes may still be unattractive for homeowners if one believes that the municipality's spending will require higher tax rates in the future, which is not an issue for renters. Indeed, 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 and tax declarations. It covers information on education, employment status etc. The corresponding data set was provided by the Swiss Federal Statistical Office (BFS) and can be geolinked by precise addresses.[8] In addition, we employ individual-level geo-referenced information about personal net income, taxes paid, retirement status, number of children and housing expenditure from the tax authorities of the canton of Bern between 2008 and 2012. This allows us to make use of effective incomes whereas we can only distinguish between household types (low, medium, high income) for the rest of Switzerland. We merge the information on real estate, income and sociodemographics with information about the geographic location of municipal boundaries that stem from BFS spanning the period 2005–12. This determines 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 needed 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 2005–12. The information on tax rates and tax base stem from the Swiss Federal Tax Administration (ESTV) and the cantonal tax/statistical offices. More

[6] As the data set used for analysis in this article is based on offered rents, it is important to compare the data to official data sources about rents. Table A1 in online Appendix A reveals that for almost all cantons the figures are highly comparable, especially for the largest canton, Zurich, and the canton we use for counterfactual simulations, Bern.

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

[8] A detailed description is provided in online Appendix E.

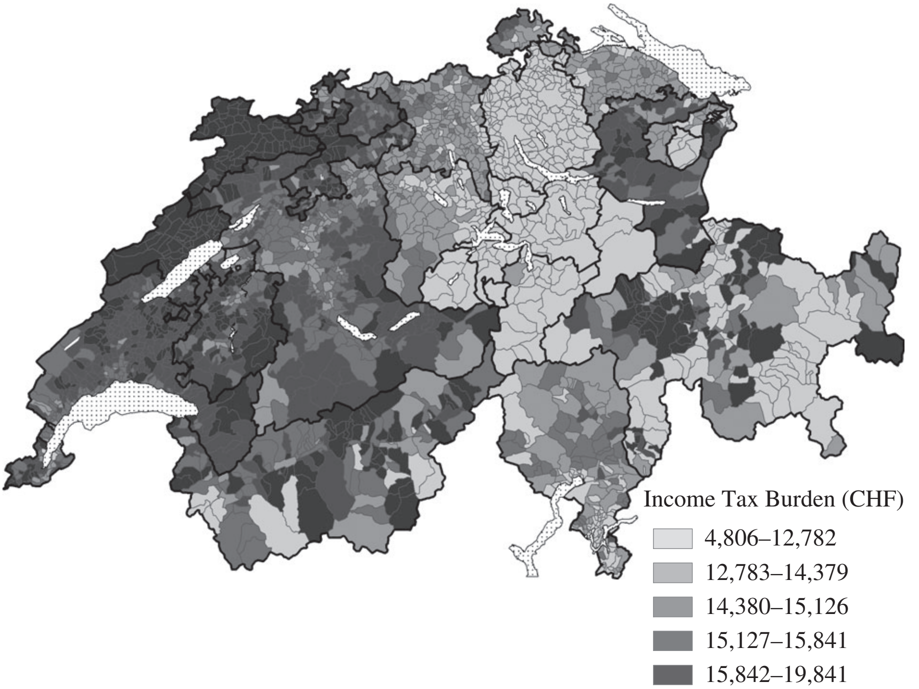

Fig. 2. *Income Tax Burden Across Swiss Municipalities*
Notes. The shading refers to the quintiles of the distribution of income tax burden. Lighter shading corresponds to a lower tax burden, darker shading to higher tax burden. The tax burden is calculated for a single household with an annual gross income of 100,000 CHF in the year 2012.

detailed information on the computation of the municipal tax burden is provided in online Appendix B.

In Figure 2, we display the total tax burden across all Swiss municipalities using different shading for the quintiles of the distribution. We plot this map for a single person with an annual income of 100,000 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 considerable 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 CHF. Interestingly, we observe municipalities in the lowest quantile of tax burden (marked in light grey) right next to municipalities in the upper quantile of the distribution (marked in dark grey). Many of these cases are municipality pairs belonging to different cantons but we find considerable variation within cantons, too.

How does the distribution of taxes relate to the distribution of rents? The map in Figure 3 plots municipal averages of rents per metre^2 separately for Switzerland. Compared to the tax burden, we observe less of a clear-cut regional pattern in the distribution of rents, as the top quintile (marked in dark grey) is scattered across western, central and eastern Switzerland. Rents are highest in the agglomerations of

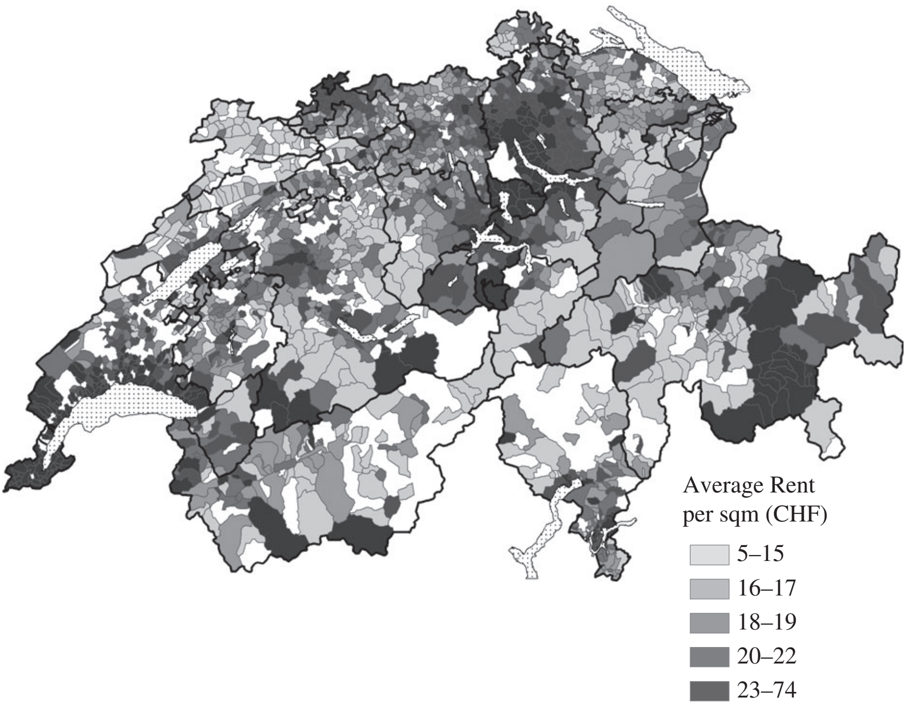

Fig. 3. *Rents Across Swiss Municipalities*
Notes. The shading refers to the quintiles of the distribution of rents per metre^2 in 2012. Lighter shading corresponds to lower rents, darker shading to higher rents.

Zurich and Geneva but we find other municipalities in the upper quintile in less agglomerated regions such as Zug, Schwyz or Graubünden (where taxes tend to be low). Of course, we cannot infer the effect of taxes on housing rents from such unconditional correlations because of many confounding factors that become evident from the maps.

4. An Empirical Strategy for Estimating the Effect of Income Taxes on Rents

This article's empirical approach identifies the impact of income taxes on housing rents based on a BDD (Black, 1999; Fack and Grenet, 2010; Gibbons *et al.*, 2013) that is nested in a residential sorting framework (Bayer *et al.*, 2007). Previous studies focused on expenditure-side effects, or more precisely, on the valuation of schooling. This analysis exploits rich information about local income taxes and provides a novel empirical test of Tiebout competition by relating rents to public revenues. With geographic location commonly being acknowledged as the key determinant of house prices and rents it seems advisable to pursue a counterfactual analysis based on observations that share the same neighbourhood characteristics. In urban economic theory and consistent with empirical evidence, the rent gradient typically decreases continuously with distance from the economic centre, publicly provided amenities, or other attractive location fundamentals (Agrawal and Hoyt, 2014). In contrast, the tax burden changes discontinuously at the municipal boundary in the multi-jurisdictional tax system of Switzerland. The analysis of narrow windows around spatial boundaries together with boundary fixed effects can thus be exploited to obtain an estimate of capitalisation of income taxes on rents. Tax-induced income sorting biases these estimates because sociodemographics jump discontinuously at boundaries. Consequently we combine this approach with an empirical strategy that takes endogenous sorting by way of location decisions that accommodate heterogenous household valuations over housing and local neighbourhoods into account and yields tractable estimates of household responsiveness to tax differences across local jurisdictions. Conveniently, this provides a starting point for counterfactual analysis that examines the effect of changes in local tax policy on the residential stratification of neighbourhoods according to income.

4.1. *A Stylised Model for Location Choice*

In order to guide our empirical analysis, we develop a simple model incorporating non-homothetic household preferences in the spirit of Deaton and Muellbauer (1980). Households derive utility from housing h and a numéraire consumption good b:

$$U(h, b), \quad \text{subject to} \quad ph + b \leq y[1 - \tau(y)],$$

where p denotes the housing rent, y refers to income, and $\tau(y)$ to the income tax. Note that the income tax schedule may be progressive and we assume that the average tax rate $\tau(y)$ as well as the marginal tax rate $\tau(y) + \tau'(y)y \in [0, 1)$. We model indirect utility as:

$$U(h^*, b^*; \alpha, \omega) = V(\tau, p; y, \eta, \alpha, \omega) = \frac{1}{\alpha} y^{\alpha}[1 - \tau(y)]^{\alpha} - \frac{\eta}{\omega} p^{\omega}, \tag{1}$$

where h^* and b^* denote optimal expenditure, $\alpha, \omega \in [0, 1]$ are taste parameters for the two goods, and $\eta > 0$ a housing parameter that captures local amenities which make a residence more valuable. Note that α measures the degree of non-homotheticity. With $\alpha = \omega = 0$ the model collapses to Cobb-Douglas demand functions.

The marginal rate of substitution between rents and taxes, $\text{MRS}_{p,\tau}$ is:

$$\frac{\mathrm{d}p}{\mathrm{d}\tau} = -\frac{\partial V/\partial \tau}{\partial V/\partial p} = -\frac{y^{\alpha}[1 - \tau(y)]^{\alpha - 1}}{\eta p^{(\omega - 1)}} < 0, \tag{2}$$

hence an increase in taxes is compensated by lower rents. Using Roy's identity, it is straightforward to show that the expenditure share of housing, $\phi = ph^*/y = \eta p^{\omega}/y^{\alpha}[1 - \tau(y)]^{(\alpha - 1)}$ is decreasing in income such that the marginal rate of substitution is unambiguously decreasing in gross income unless α equals zero:

$$\frac{\partial \text{MRS}_{p,\tau}}{\partial y} = -\frac{\alpha y^{(\alpha-1)}[1 - \tau(y)]^{(\alpha-1)}}{\eta p^{(\omega-1)}} \left\{ 1 + \frac{(1 - \alpha) y \tau'(y)}{\alpha[1 - \tau(y)]} \right\} \leq 0. \tag{3}$$

Thus, high-income households have a higher marginal willingness to pay for lower taxes relative to low-income households. Note that the second term captures the role of

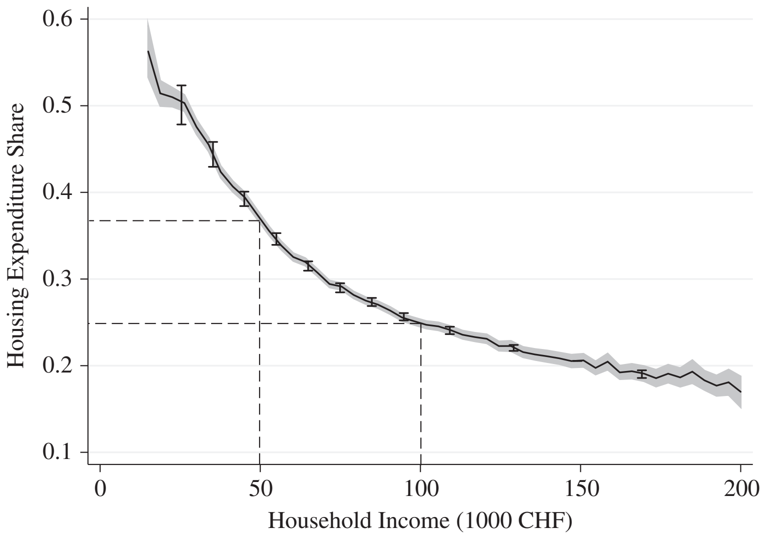

Fig. 4. *Individual Income and Housing Expenditure Share*
Notes. Data underlying this Figure stem from census data covering the canton of Bern. These data cover the year 2012 and include information about individual income and housing expenditure (measured by the rents and imputed rents in the case of owner-occupied properties; the latter are reported on the tax declarations). The local linear regression employs an Epanechnikov kernel based on a rule-of-thumb (ROT) bandwidth that minimises the conditional weighted mean integrated squared error. At 50 and 100,000 CHF the point estimates are 0.36 and 0.25 respectively. Dividing the income distribution in three equally sized groups the point estimates suggest expenditure shares of 0.41, 0.30, and 0.23 for the low (38,100 CHF), middle (69,695 CHF) and high-income (119,015 CHF) groups. For the construction of the confidence intervals, standard errors are computed as the square root of the estimate of the conditional variance of the local polynomial estimator at each grid point. This requires estimates of the residual variance at each grid point. As a pilot bandwidth the value of 1.5× ROT is used. Estimation of the variance assumes constant residual variance. The red bars are obtained from a regression of the expenditure shares on dummy variables corresponding to 12 income bins (at distances of 10k (20k) CHF). The bars correspond to the coefficients' 95% confidence bound.

progressivity as a high value of $\tau'(y)$ intensifies the high-income households' willingness to pay for lower taxes.

Figure 4 bases on individual income and housing expenditure data in Switzerland and illustrates a negative correlation between the expenditure shares for housing and gross income which supports the assumption of non-homothetic preferences.[9]

Indirect utility in municipality m is a function of local taxes, rents, individual income, amenities and individual taste $V_m(y) \equiv V(\tau_m, p_m; y, \eta, \alpha, \omega)$. For locations at the border of two municipalities (m, m'), local amenities converge to the same level such that we can compute for given income $\hat{y}$ and tastes the marginal rate of substitution

[9] Note that the density of observations is low at the lower and upper tails of the distribution (in particular at incomes below 21,735 CHF and above 167,847 CHF – the 2.5 and 97.5 percentiles of the distribution) such that the tails have to be interpreted carefully. The Figure is unconditional on relevant observable and unobservable factors accounted for in the following Sections.

between housing rents and taxes that satisfies $V_m(\hat{y}) - V_{m'}(\hat{y}) = 0$. This motivates the empirical strategy described in subsequent subsections. In a first step, we estimate the $\mathrm{MRS}_{p,\tau}$ along the lines of subsection 4.2. In a second step, we are interested in income sorting across municipalities and estimate whether $\mathrm{MRS}_{p,\tau}$ is declining in income (3). The according strategy is outlined in subsection 4.3.

4.2. *Estimating the* $\mathit{MRS}_{p,\tau}$

Let us use the following notation. We observe postings of individual residences denoted by h in year t which belong to a municipality $m \in M$. Each municipality levies an annual income tax from its residents that can vary across household type and income level. Using data on administrative boundaries, we assign to each municipality all its neighbouring municipalities. Restricting our data set to posts with non-missing and plausible values for the residence characteristics reported in Table 1, we obtain $M = 2{,}421$ municipalities.[10] We denote by $p = \{m, m'\}$ a pair of neighbouring municipalities where 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 100,000 CHF and to a single person household in our benchmark specifications. This approximately corresponds to 1.25 times the median income in Switzerland. In order to ensure 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 group because municipalities set a flat multiple on the cantonal rate such that the progression of the tax code is constant within cantons.[11]

Some municipalities, especially the ones close to city centres or with desirable amenities such as mountain or lake view, might exhibit a substantial number of second homes whose inhabitants should not consider the income tax in their location choice. Some of these locations are *ex ante* excluded because of restrictions, we place on the data (for instance, we exclude location pairs that differ in altitude by 400 metres as we describe subsequently). While for the individual flat, we cannot distinguish whether it is used as a primary or secondary home, we are able to include the share of second homes (holiday and commuter homes) at the level of municipalities in the regressions.[12]

We use this information together with data on the log rent per metre^2 at t of residence h located in m. Note that each h is uniquely assigned to a municipality and to a border point in a given year such that we can drop the subscripts m and denote the rent of h by p_{ht}. We estimate the $\mathrm{MRS}_{p,\tau}$ from a conventional hedonic regression model

[10] This figure refers to municipalities over the full time period. We assign the respective neighbouring municipalities for each year separately according to the corresponding classification and digital maps of municipalities. As some municipal boundaries were modified by mergers over the period under consideration, we use separate digital maps for each year.

[11] 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 was 80,600 CHF in Switzerland in 2012 (BFS).

[12] The share of second homes stems from the special statistical evaluation ‘Gebäude- und Wohnungsstatistik, Sonderauswertung GWS 2012’ by BFS. See Table 1 for details.

Table 1
Summary Statistics

	Full data set				Boundary sub-sample			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Main variables								
Rent/m^2 R_h	21.082	6.286	5	52	20.943	5.847	5	52
m_h^2	87.464	38.986	15	230	88.691	38.275	15	230
Tax burden T_m^{S100}	13,758.314	2,487.117	4,805	20,096	13,724.432	2,481.787	4,805	19,766
ΔT_m^{S100}	1,004.678	1,039.926	1	8,938	905.039	880.421	1	8,938
Distance (meter) D_{hb}	1,203.933	903.028	4.156	16,978	586.338	237.303	4.156	999.993
Sociodemographic covariates								
Income$_i$	51,618	126,731	0	*	61,948	84,354	0	*
Children$_i$	0.182	0.386	0	1	0.213	0.409	0	1
No. children$_i$	0.307	0.737	0	18	0.367	0.800	0	18
Retirement status$_i$	0.306	0.461	0	1	0.346	0.476	0	1
Residence covariates								
Rooms$_h$	3.462	1.376	1	10	3.537	1.353	1	10
Parquet$_h$	0.380	0.485	0	1	0.384	0.486	0	1
Lake view$_h$	0.013	0.112	0	1	0.016	0.125	0	1
Ceramicstove$_h$	0.203	0.402	0	1	0.210	0.408	0	1
Fireplace$_h$	0.049	0.217	0	1	0.053	0.224	0	1
Garden$_h$	0.097	0.296	0	1	0.106	0.307	0	1
Balcony$_h$	0.651	0.477	0	1	0.668	0.471	0	1
Terrace$_h$	0.077	0.266	0	1	0.077	0.267	0	1
Carport$_h$	0.107	0.309	0	1	0.109	0.312	0	1
Age$_h$	32.657	35.069	0	412	30.002	30.014	0	412
Floor$_h$	1.936	1.707	−2	20	1.942	1.779	−2	20

Table 1
(*Continued*)

	Full data set				Boundary sub-sample			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Municipality covariates								
Population$_m$	59,855.503	99,031.656	87	376,990	39,992.610	74,776.896	97	376,990
Area$_m$	24.131	25.513	0.310	430.110	15.216	17.827	0.310	139.150
Second homes$_m$	0.117	0.057	0.016	0.808	0.110	0.047	0.021	0.770
Property tax rate$_m$	0.236	0.463	0.000	1.500	0.207	0.425	0.000	1.500
Municipalities M	1,964				1,113			
Border points B	6,084				3,030			
Residences I	2,428,389				1,107,648			
Households (Bern)	449,117				172,114			

Notes. The summary statistics correspond to data pooled over the period 2005–12. We dropped observations with a rent of less than 5 CHF or more than 52 CHF per metre^2 (these values equal approximately the 1st and 99th percentiles), and observations with missing information about the residence covariates. Furthermore, information about the share of second homes is not available for all municipalities and years. This diminishes the sample from about 3.5 million observations to 2,428,389 residences in the full data set and from 2,421 to 1,964 municipalities. The boundary subsample consists of all residences that belong to a municipality-year dyad which features observations within a distance of 300 metres from the boundary. Accordingly, the municipality composition remains stable when limiting distance windows. All border points which correspond to geographic barriers (e.g. rivers), linguistic barriers, or canton borders have been dropped. Tax burden and rents are measured in CHF and for a single household with an annual income of 100,000 CHF. Area measured in kilometre^2 and annual population figures stem from BFS, property tax rate is measured in per mille. The second home share is measured as 1 – the share of residences registered as main residences in the number of total inhabited residences by municipality. This refers to residences rented as second homes and owned as second home and to the person who is registered to live in the property (and not an investor). For a detailed definition of the tax burden see onine Appendix B. * refers to confidential data.

of the following form as a baseline for comparison:

$$p_{ht} = \kappa_0 + \kappa_\tau \tau_{mt} + \boldsymbol{\kappa}_\mu \boldsymbol{\mu}_{mt} + \boldsymbol{\kappa}_x \mathbf{x}_{ht} + \rho_{ct} + v_{ht}, \tag{4}$$

where $\boldsymbol{\mu}_{mt}$ is a vector of municipality characteristics, including the share of second homes, log population, and log area, $\mathbf{x}_{ht}$ is a vector of detailed residence characteristics, ρ_{ct} is a set of canton-year dummies and v_{ht} is an error term that is multiway clustered at the municipality and year levels (Cameron *et al.*, 2011) to take a general form of correlation of observations within municipalities and over time into account. Because of the specification of functional form for rent and tax, κ_τ (or more generally, the coefficients estimated throughout this article) represents a tax elasticity. We absorb canton-year-specific information in order to make specification (4) comparable to the BDD approach because the latter is estimated from within-canton variation. $\mathbf{x}_{ht}$ includes the number of rooms, floor, residence age and binary indicators that equal 1 if the flat features a parquet floor, lakeview, ceramic stove, fireplace, garden, balcony, terrace and carport respectively.

To overcome the bias in κ_τ resulting 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 $b = \{1, \ldots, B\}$ using geographic information system (GIS) data. The number of border points depends on size and 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. This procedure maximises the number of residences h located in different municipalities within a close neighbourhood. With the border points at hand, we compute for each residence h the Euclidean distances $D_{ht}(b)$ 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 sufficient density of observations at the threshold. Each residence is uniquely assigned to one of the three remaining border points using the minimum distance $D_{ht} = \min\{D_{ht}(b)\}$. Finally, outcome p_{ht} is measured for each residence h 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:

$$p_{ht} = \beta_0 + \beta_\tau \tau_{mt} + \boldsymbol{\beta}_\mu \boldsymbol{\mu}_{mt} + \boldsymbol{\beta}_x \mathbf{x}_{ht} + \theta_{bt} + \varepsilon_{ht}, \tag{5}$$

where β_τ measures the tax elasticity, $\boldsymbol{\mu}_{mt}$ is a vector of municipality characteristics, $\mathbf{x}_{ht}$ is a vector of residence characteristics, θ_{bt} is the border-point and year-specific fixed effect that absorbs all variation specific to the neighbourhood of the border point in both the cross-sectional and the time dimension and ε_{ht} is an error term which we multiway-cluster at the municipality and year level to account for serial correlation along boundaries that may be due to both spatial correlation and correlation over time.[13] By

[13] Note that a residence h that has n neighbours appears in n pairs and accordingly the error terms of all these n terms are correlated because the residence-specific error term enters the error of each pair. For this reason each residence h is used only once. As the border points are measured on a very fine scale, we abstain from a spatial correction of the error term; see also Gibbons and Overman (2012). We have also estimated the regressions on data which included the residences in n pairs if applicable. To correct the error term, we used multiway clustering of the standard errors at the municipality, year and, in this case, also at the residence level; alternatively, we corrected the standard errors according to Duranton *et al.* (2011). In any case, the results were robust to the ones presented in this article(see Table D1 in online Appendix D).

restricting the sample to residences within a close proximity to the border points $D_{ht} \leq 1$ kilometre, ≤ 600 metres, ≤ 300 metres, we ensure that θ_{bt} accurately captures location-specific unobservables. Note that β_τ measures a local elasticity, which may not be representative of the overall elasticity. To tackle this potential concern, we weigh the observations by the probability that the residence unit of observation is located within the respective boundary subsample, conditional on residence and municipality characteristics (Bayer *et al.*, 2004).[14]

4.2.1. *Discussion of the identifying assumptions*

Under the assumption that unobservable determinants of rents vary continuously at the boundary, the parameter β_τ identifies the causal effect of interest for observations in the vicinity of the boundary. We choose different radii from the municipality boundary to improve on the continuity of unobserved amenities that we aim to hold constant. Hereby we face the trade-off that the number of pairs is reduced, leading to a potential loss in precision. By way of the chosen design we condition on the proximity between individual residences as well as on the proximity to a municipal boundary in terms of their distance to a common border point. Hence, this approach allows us to capture the local conditions of a small well defined area by holding both commuting costs and neighbourhood characteristics constant.

The resulting distribution of observations in space around individual border points is best described in a map as the one in Figure 5. The map exemplifies, for two neighbouring 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 residence (black) is located at the lakeside, the second type (light grey) is located close to the centres and along main traffic routes, and the third type (dark grey) is located outside the centre on the hill in both municipalities. Each of the three border points is assigned a unique fixed effect and the sample is restricted to observations within the bandwidth. This ensures that residences which are close to each other are compared and that the continuity assumption is satisfied.

A discontinuity in the density of observations on both sides of the tax border would point to a systematic difference in the supply of housing which is correlated with local taxes. This would violate the identifying assumptions. 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 6. Naturally, the frequency of postings decreases in the close neighbourhood of a municipality boundary which is due to residences being concentrated in the centre of a municipality. We observe an almost symmetric shape of the histograms on both sides of the boundary, hence there is no indication of a bias toward more postings on either side of the threshold. As an alternative test we exploit land-use data and compute the share of area covered by

[14] The weight is obtained from a probit regression of an indicator measuring whether the observation is located within the chosen radius on flat and municipality characteristics. We checked the sensitivity of this approach to the use of inverse distance weights to the CBD and to controlling for the log distance to the centre agglomeration of the municipality as well as to the neighbouring one in the regressions. The results remain unchanged.

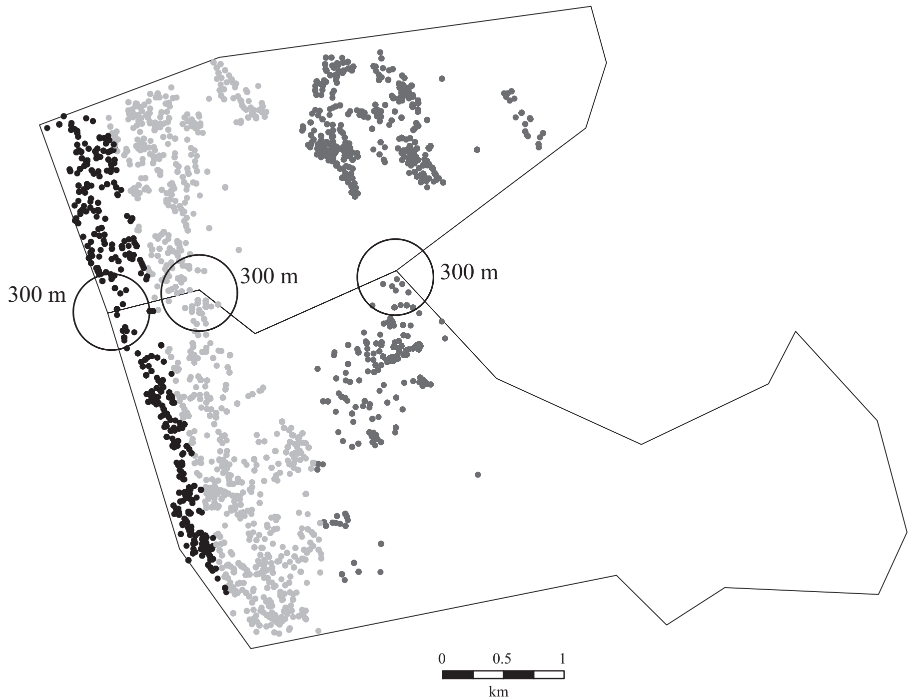

Fig. 5. *Residences and Border Points – Example: Zollikon/Küsnacht*
Notes. Each dot refers to one residence for which we observe a posting containing information on the rent per metre^2, and on all covariates listed in Table 1. The shading of the dots indicates the border point that residences were assigned to on the basis of the minimum distance. Note that Zollikon and Küsnacht are two municipalities in the canton of Zurich which are situated on the lake of Zurich. Residences marked in black (border point 1) are very close to the lake shore, ones marked in light grey are located in the CBD, while residences marked in dark grey (border point 3) are on a hill.

buildings within a certain distance from the border point. We do not find evidence for a discontinuity in the share of covered area at the border between low and high-tax municipalities which confirms the lack of systematic supply differences in the vicinity of the border.[15]

Finally, the identification strategy relies on the assumption that estimates of β_τ can be unambiguously attributed to the jump in 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 not coincide with cantonal borders.

[15] The corresponding map is shown in Figure D1 in online Appendix D. Note also that in general, the price elasticity of housing supply is the lowest in international comparison in Switzerland (see the 2015 OECD Economic Survey of Switzerland).

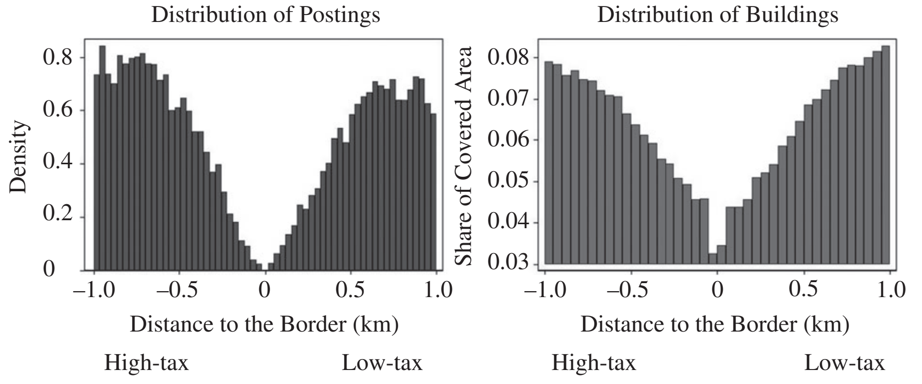

Fig. 6. *Taxes, Rents and the Municipal Border*
Notes. Treated units (lower taxes) are assigned positive distances while control units (higher taxes) are assigned negative distances. We use pooled data for 2005–12. The histograms refer to the distribution of postings in our data set depicted against the distance from the closest border point and to the area share covered by buildings within bins of 50 metre from all border points used in our analysis. The latter uses data derived from aerial images collected between 2004 and 2009 (BFS; see also online Appendix Figure D1).

We address further concerns about of the continuity assumption in the following way.

(*i*) In order to remove geographic barriers that separate municipalities at b, we drop pairs that are separated by rivers and highways or feature a difference in altitude of more than 400 metres, and ones split by language borders.[16]

(*ii*) We account for a potential positive correlation between taxes and quality of residences and the possibility that these differ systematically in m and m' by including housing characteristics in (5).

(*iii*) In order to rule out central business districts (CBD) specific effects, for instance, due to individual preferences to live at a prestigious address, we estimate the models for subsamples that exclude agglomerations (i.e. Basel, Bern, Geneva, Lausanne and Zurich) separately. This should also account for the higher second home shares in cities that we observe in the data.

(*iv*) We may be concerned about potential asymmetries in level and quality of excludable public goods between municipalities, because elementary schools are typically financed on the municipal level. This is addressed by exploiting tax variation within school districts in Section 6.

Furthermore, we analyse the potential role of the supply of other excludable public goods and services provided at the municipality level. Information from ESTV has revealed that relevant municipality-specific goods and services (besides elementary schooling) are waste, water and sewage services. We combine per unit prices from the Swiss Federal Price Monitor with information from the same source on typical annual consumption amounts. This allows us to compute moments for the absolute

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

municipality differentials in the annual costs of these three types of services. On average, the annual cumulative cost-differential is 63 CHF as displayed in Table B1 in online Appendix B. The Figures are minor compared to the annual rent and tax differentials shown in the next Section. Moreover, there is no significant correlation between the differential in taxes and the one in costs of public services (correlation coefficient of 0.04).

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 which is reflected by the rent gradient within municipalities. We test the robustness of our results to these concerns in Section 6 by adding municipality fixed effects, and by allowing for asymmetric gradient control functions.

4.3. *The Role of Sociodemographic Sorting*

In the following, we develop a strategy which accommodates heterogenous preferences about housing and taxes in a residential sorting model and accounts for the correlation of sociodemographic characteristics with unobserved local neighbourhood attributes. To shed light on the extent of sorting across municipality borders that exhibit positive tax differentials, Figure 7 visualises the discontinuity in the share of skilled *versus* unskilled and high-income *versus* low-income individuals. The Figures indicate a clear and systematic difference for each of these variables across high-tax and low-tax municipalities. These patterns suggest that addressing residential sorting is warranted.

To outline the model, let us rewrite (1) in a given year as follows:

$$V_h^i = \boldsymbol{\delta}_h + \lambda_h^i + u_h^i, \tag{6}$$

where:

$$\boldsymbol{\delta}_h = \boldsymbol{\gamma}_{0x}\mathbf{x}_h - \gamma_{0\tau}\tau_m - \gamma_{0p}p_h + \theta_b + \xi_h, \tag{7}$$

$$\lambda_h^i = \left(\sum_{j=1}^{J}\gamma_{jx}z_j^i\right)\mathbf{x}_h - \left(\sum_{j=1}^{J}\gamma_{j\tau}z_j^i\right)\tau_m - \left(\sum_{j=1}^{J}\gamma_{jp}z_j^i\right)p_h. \tag{8}$$

Vector $\boldsymbol{\delta}_h$ measures the utility provided by housing choice h (this corresponds to the mean indirect utility of h, if z_j^i is transformed to have mean zero), ξ_h represents unobserved preferences for h correlated across households, λ_h^i are preference parameters which are heterogenous across households i with j characteristics, and u_h^i are unobserved preferences not captured by ξ_h. We account for j = 3 dimensions of sociodemographic characteristics: income, number of children, and retirement status (a binary variable). The variables subsumed in row vector $\mathbf{x}$ are age of the building, floor, flat size, log distance to the CBD, an indicator for whether the building has a single or multiple units and the number of flats in the building. Equation (8) shows that the heterogeneity of preferences modelled in subsection 4.1 is accounted for in a flexible way by a characterisation of distinct individual

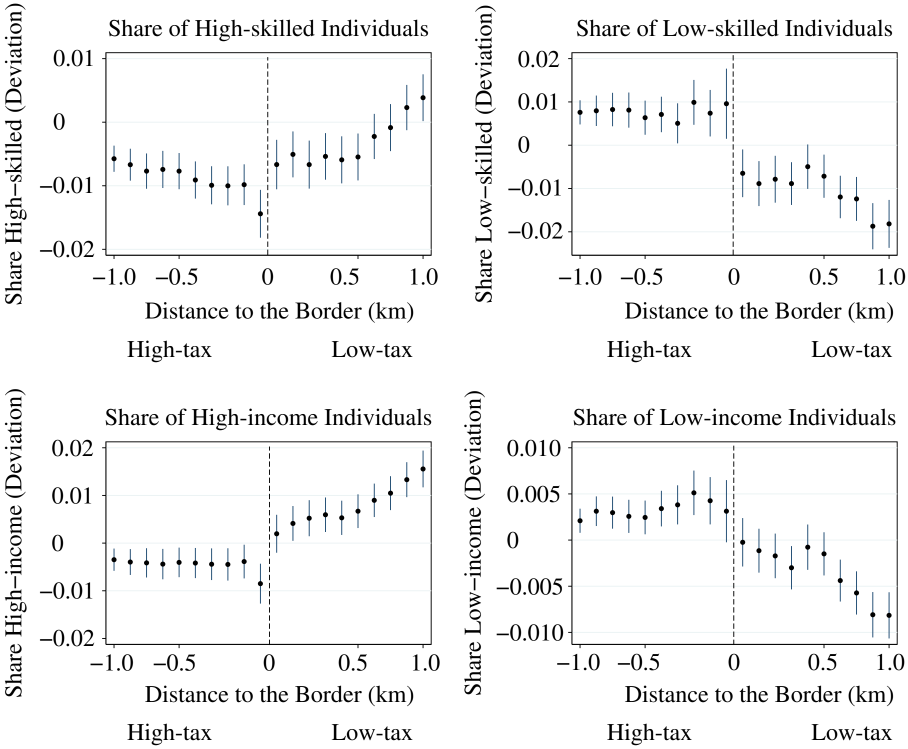

Fig. 7. *Taxes, Sociodemographic Variables, and the Municipal Border*

Notes. 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 kilometre^2 grids where we require a minimum population of 50. Each residence h 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 online Appendix E. The coefficients and 95% confidence bands in the Figures are obtained by regressions of the variable in question on border-point fixed effects and on dummy variables corresponding to bins of 100 metres. Colour Figure can be viewed at wileyonlinelibrary.com.

tastes for housing and location characteristics. The model also captures urban features related to the housing market by allowing households to value the distance to the CBD differently. In the context of the model specified in subsection 4.1, δ_h may be described as to correspond to the parameter η, while λ_h^i captures the preference parameters ω, α.

Assuming that u_h^i is drawn from the extreme value distribution (McFadden, 1973, 1978), we model location decisions in a discrete choice framework. The logit probability for the choice of h can be written as:

$$P_h^i = \frac{\exp(\delta_h + \lambda_h^i)}{\sum_{j=1}^{J} \exp(\delta_h + \lambda_h^i)}.$$

The independence of irrelevant alternatives property allows us to sample alternatives in the choice sets of individuals randomly. In order to approximate the true choice set available to an individual deciding between flats within a certain radius, we constrain the choice set by a distance of 30 kilometres within cantons.[17] With the likelihood function $\ell = \sum_i \sum_h I_h^i \ln(P_h^i)$, where I_h^i is a choice indicator function, we use contraction mapping in the spirit of Berry *et al.* (1995) to back out the vector $\boldsymbol{\delta}_h$ that solves $\sum_i (P_h^i) = S_h$ for all h, where S_h denotes supply of a house.[18] This permits revealed individual preferences to vary over the choice variables and maximises the likelihood function at the vector $\boldsymbol{\delta}_h$ where supply equals demand (the aggregate predicted choice probability for h) for each h, i.e. markets clear.

In a second step, the vector of mean indirect utilities, $\boldsymbol{\delta}_h$, is regressed on τ_m, p_h, and $\mathbf{x}_h$ using (7), as well as the border-point fixed effects subsumed in θ_b. The procedure accounts for the aforementioned unobserved location factors and leads to estimates of the mean tax elasticity that represent revealed mean preferences for taxes. Therein, we use instrumental variables for the rent to address potential correlation between housing rents and unobserved neighbourhood amenities: the number of housing units and the share of commercial land-use in the total area located at least 5 kilometres away from h but within the same canton. As is common in the literature, it is assumed that attributes of locations with a sufficient distance from a chosen unit h affect equilibrium rents but not utility at h directly such that these variables can be employed as instruments for rents.[19] As the geography in close distance to housing choice h is expected to influence utility at h in a direct way, we include the number of flats within 1.5 kilometres in the regressions. Given the estimates from both steps, the empirical model allows us to obtain estimates of the distribution of preferences for lower taxes of household by income groups.

5. Results

5.1. *Descriptive Statistics*

In total, we can assign more than 2.4 million postings containing non-missing information on metre^2 rents and many other residence characteristics to the municipality pairs we consider. These include information about detailed characteristics of the residence as listed in Table 1. Rents are measured in CHF and the annual tax burden (in CHF) refers to a single household earning 100,000 CHF per year. For a subsample limited to the canton of Bern, we employ sociodemographic information on the households' taxable income (in CHF) as well as on the number of children, and the retirement status.

[17] This is consistent if the number of houses goes to infinity and the number of individuals grows fast enough relative to the number of houses. Alternatively, other strategies that do not depend on the IIA assumption may involve a nested logit model combined with a control function approach that accounts for endogeneity of rents. The choices can be observed by the customer as all postings are publicly available free of charge.

[18] Solving $\partial\ell/\partial\delta_h$ yields $S_h - \sum_i (P_h^i) = 0$. The contraction mapping is then:

$$\delta_h^{d+1} = \delta_h^d + \ln(S_h) - \ln(\textstyle\sum_i \widehat{P}_h^i),$$

where d represents the iteration of the contraction mapping.

[19] The analogue argument in the IO literature is that product prices are influenced by the availability of close substitutes but the utility derived from consuming a certain product is not directly influenced by the attributes of the non-chosen alternatives.

The identification strategy requires us to restrict the sample to observations in the close neighbourhood of a municipal boundary which we refer to as the BDD subsample. It consists of residences that belong to a municipality-year dyad which features observations within a distance of 300 metres from the boundary. This implies that the number of municipalities remains stable when restricting the radii to alternative maximum distances of 1 kilometre and 600 metres. In total, the BDD sample comprises 1,107,648 postings in 1,113 municipalities and it focuses on 3,030 border points with sufficiently close residences on both sides of the boundary. The average rent per metre^2 amounts to 21 CHF in both the full and the BDD sample. Furthermore, the average annual tax burden is about 13,760 CHF in the full and about 13,720 CHF in the BDD data. The average tax differentials between the high and low-tax municipalities amount to about 1,000 CHF and 900 CHF respectively.

When comparing the two samples, 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. This indicates that the residences in the BDD sample represent the ones in the full data set well. In general, our data set has an excellent 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.

5.2. *BDD Results*

Table 2 reports the coefficients and standard errors based on (4) in columns (1), (2) and (5) in columns (3)–(8). In contrast to panel (*a*), panel (*b*) includes flat characteristics as explanatory variables. Panel (*c*) weights the observations by the inverse probability that an observation is located within the respective BDD sample. Uneven columns refer to samples that include all municipalities in the chosen distance window while even columns exclude the five largest cities and their neighbouring municipalities; the distance band around the border points is limited to 1 kilometre in columns (3)–(4); to 600 metres in columns (5)–(6); and to 300 metres in columns (7)–(8). The results may be summarised as follows: first, the elasticities carry a negative sign as expected and are statistically significant. Second, the inclusion of boundary fixed effects is able to considerably reduce the bias from the conventional hedonic approach and contributes substantially in terms of the models' explanatory power. Notably, columns (1) and (2) include canton-year fixed effects, and thus the marked increase in explanatory power can be attributed exclusively to spatial variation at fine scale. Third, limiting the sample to close distance bands around boundaries increases the coefficients. Fourth, the inclusion of flat characteristics lowers the $\mathrm{MRS}_{p,\tau}$, pointing to a potential correlation with taxes. The coefficients on the share of second homes and log population carry a positive sign while the one on log area is negative but neither of them affects the tax elasticity. The corresponding coefficients are small and not reported. Quantitatively, we find that an increase in the tax burden by 1% lowers rents by approximately 0.9% using the conventional OLS approach with and without agglomerations respectively. The $\mathrm{MRS}_{p,\tau}$ increases to -0.323, -0.286, and -0.241 when accounting for border point fixed effects in the 1 kilometre, 600 metre and 300 metre windows respectively. For each window and for both samples, it holds true that

Table 2

Boundary discontinuity design (BDD) for Municipal Taxes and Rents

	OLS		BDD 1 km		BDD 0.6 km		BDD 0.3 km	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel (*a*): without residence covariates								
Tax elasticity	−0.856***	−0.865***	−0.275***	−0.289***	−0.234***	−0.258***	−0.218**	−0.247**
	(0.059)	(0.060)	(0.026)	(0.024)	(0.050)	(0.055)	(0.072)	(0.081)
Adj. R^2	0.340	0.296	0.468	0.460	0.495	0.493	0.526	0.534
Panel (*b*): accounting for residence covariates								
Tax elasticity	−0.860***	−0.866***	−0.323***	−0.320***	−0.286***	−0.305***	−0.241***	−0.272***
	(0.058)	(0.059)	(0.028)	(0.025)	(0.049)	(0.056)	(0.063)	(0.072)
Adj. R^2	0.401	0.370	0.521	0.522	0.547	0.551	0.577	0.589
Panel (*c*): accounting for residence covariates and location probability within the BDD sample								
Tax elasticity	−0.900***	−0.840***	−0.333***	−0.313***	−0.292***	−0.302***	−0.277***	−0.303***
	(0.093)	(0.069)	(0.044)	(0.030)	(0.057)	(0.065)	(0.077)	(0.081)
Adj. R^2	0.407	0.351	0.486	0.498	0.487	0.519	0.510	0.547
Observations	2,428,389	1,886,430	1,107,648	899,790	559,383	470,951	207,834	178,502
No. b,t	–	–	11,885	11,258	9,133	8,592	6,275	5,851
No. M,t	14,094	14,044	5,362	5,312	5,277	5,224	4,887	4,809

Notes. ***, **, * denote statistical significance at the 1%, 5% and 10% level respectively. Standard errors are multiway clustered at the level of year and municipality. The first two columns are estimated by OLS and include canton-year fixed effects. Columns (3)–(8) refer to BDD estimates and include border-point-year fixed effects. Each specification covers the years 2005–12 and is reported for a sample of all observations in odd columns and one excluding the five largest agglomerations in even columns. All regressions include log area, log population and the share of second homes by municipality. Regressions in panels (*b*) and (*c*) control for residence covariates summarised in Table 1. Regressions in panel (*c*) are weighted by the inverse probability that an observation is located in the respective boundary subsample conditional on residence and municipality characteristics. No. M,t and b,t refer to municipality-year and border-point-year dyads.

the 95% confidence intervals of the OLS coefficients do not overlap with those of the BDD coefficients, pointing to the importance of confounding location-specific factors that are absorbed by the boundary point fixed effects. The difference between the coefficients in the samples with and without agglomerations – estimates for the latter range from -0.272 to -0.320 – tend to be negligible. The quantitative robustness suggests that the estimated tax elasticity is not confounded by differences of population density or other agglomeration effects. Finally, the difference between the estimates in panels (*b*) and (*c*) is minor, indicating that the estimates are representative of the full sample.

Overall, accounting for unobservable neighbourhood characteristics yields a $\text{MRS}_{p,\tau}$ which is only two thirds of the one obtained from conventional hedonic regressions. Moreover, decomposing the explained variance of log rents, we find that location (border point and municipality) fixed effects explain about 80%. This is much higher than the variation explained by residence characteristics, amounting to roughly 2.5%, and time fixed effects, which explain about 2%. Our results illustrate that neighbourhood characteristics represent the decisive role in explaining rents.

5.3. *Sorting Results*

We report estimates from the model described in subsection 4.3 in Table 3 for the canton Bern. These data offer information about individual income and other individual-level characteristics for the universe of residents in the canton and it can be precisely geo-referenced. The cost of reducing the sample size is relatively mild as the canton Bern features the highest number of municipalities across Swiss cantons. The following analysis bases on 386 municipalities and about 170,000 unique households.

Estimates of the tax elasticity now represent preferences for taxes corresponding to (2) which are reported across households that differ in income. More precisely, the Table indicates the $\text{MRS}_{p,\tau}$ for the mean household and for households whose income is at the 25th and the 75th percentile of the income distribution respectively. Model I shown in columns (1)–(2) allows individual characteristics to interact with flat characteristics, housing rents and taxes. Model II reported in columns (3)–(4) additionally allows for heterogeneous preferences about sociodemographic neighbourhood characteristics. Model III shown in columns (5)–(6) of the Table accounts for mobility, or migration costs by controlling for the distance to the place of residence applicable in $t-2$ as well as interactions of the moving distance with income, the number of children and the retirement status. For this, we construct a panel of movers (which represent about 6% of the individuals in the sample during the two years considered) *versus* non-movers.

The findings can be summarised as follows. First, it is noteworthy that the estimated mean elasticities are quantitatively similar to the ones estimated in the BDD model. This suggests that accounting for unobserved location characteristics eliminates most of the bias inherent in conventional house price regressions. As shown in panel (*a*), the estimated mean $\text{MRS}_{p,\tau}$ is -0.309 within a 600 metre distance band (column (1)) and amounts to -0.348 within a 1 kilometre distance radius (column (2)). Put differently, the results imply that the mean willingness to pay for a decrease in income taxes by 1% is about 0.3%. Consistent with the model presumption that preferences

Table 3

Sociodemographic Sorting

	Model I		Model II		Model III	
	0.6 km (1)	1 km (2)	0.6 km (3)	1 km (4)	0.6 km (5)	1 km (6)
Panel (*a*): tax elasticity of housing prices ($MRS_{p,\tau}$)						
25th percentile of income	−0.292*** (0.014)	−0.339*** (0.046)	−0.260*** (0.011)	−0.260*** (0.012)	−0.293*** (0.020)	−0.333*** (0.021)
Mean income	−0.309*** (0.015)	−0.348*** (0.047)	−0.266*** (0.011)	−0.267*** (0.012)	−0.299*** (0.020)	−0.337*** (0.021)
75th percentile of income	−0.338*** (0.016)	−0.364*** (0.048)	−0.276*** (0.011)	−0.276*** (0.012)	−0.308*** (0.021)	−0.348*** (0.022)
Panel (*b*): marginal willingness to pay for 1% increase of neighbourhood income						
25th percentile of income			0.001*** (0.002)	0.040*** (0.003)		
Mean income			0.021*** (0.002)	0.056*** (0.003)		
75th percentile of income			0.056*** (0.003)	0.086*** (0.004)		
Panel (*c*): marginal willingness to pay for 1% reduction in migration distance						
25th percentile of income					0.017*** (0.001)	0.018*** (0.001)
Mean income					0.017*** (0.001)	0.018*** (0.001)
75th percentile of income					0.017*** (0.001)	0.017*** (0.001)
No. *i*	65,004	172,114	65,004	172,114	57,124	136,106
No. *b*	290	441	290	441	290	441

Notes. ***, **, * denote statistical significance at the 1%, 5%, and 10% level respectively. Standard errors in parentheses. The first line in panel (*a*) reports the tax elasticity of housing rents (marginal willingness to pay for 1% tax increase) at the 25th percentile of income, the second and third lines report the corresponding effects for the mean income and the 75th percentile of income respectively. Model I includes income, number of children and retirement status. Location and flat-specific characteristics include size, floor, building age, distance to the CBD, single or multi-unit building, number of flats in building as well as all interactions of individual characteristics with flat characteristics, the rental rate and the log tax burden. Model II adds local sociodemographic characteristics measured by average income, share of children and share or retirees as well as the interactions of these location-specific characteristics with individual characteristics. Panel (*b*) reports the corresponding marginal willingness to pay for an increase in neighbourhood income for households types along the distribution of income. Model III includes the distance to the place of residence applicable in $t-2$ (moving distance) as well as interactions of the moving distance with income, the number of children and the retirement status. Panel (*c*) reports moving costs for different income levels. About 6% of the individuals moved to another residence during the two years considered. The sample covers the canton of Bern in the year 2012. No. *i* and *b* refer to individual households and border points. The first stage results corresponding to the three models are reported in online Appendix D.

are non-homothetic and higher-income individuals have a relatively stronger preference to sort into low-tax jurisdictions the $MRS_{p,\tau}$ is decreasing in income as predicted by (3). The estimates within a 600 metre radius reveal that the $MRS_{p,\tau}$ is −0.292 for households at the 25th percentile and −0.338 at the 75th percentile. These income-induced differences in the willingness to pay for lower taxes seem relatively mild, yet the results imply that a one standard deviation increase in income reduces the $MRS_{p,\tau}$ by about 7 percentage points. Note also that the changes in the $MRS_{p,\tau}$ are more pronounced at lower incomes. The $MRS_{p,\tau}$ is −0.225 for households at the 10th percentile and −0.250 at the 15th percentile. This pronounced change in the MRS at relatively low incomes is consistent with Figure 4 where the drop in expenditure shares is strongest for incomes below the 25th percentile (about 50,000 CHF). Second, accounting for sociodemographic neighbourhood characteristics eliminates some of the remaining bias in the BDD (columns (3)–(4)). However, the increase in the mean $\mathrm{MRS}_{p,\tau}$ to approximately −0.266 is minor. The same qualitative pattern compared to columns (1)–(2) holds regarding the estimates of the $\mathrm{MRS}_{p,\tau}$ over the distribution of incomes. Third, the marginal willingness to pay for residing in a high-income neighbourhood is increasing in income, thus the heterogeneity estimates provide evidence for a sorting of households based on preferences for locating in neighbourhoods with similar characteristics. Richer households are willing to pay 0.056% higher rents for a 1% increase of neighbourhood income. This figure falls to 0.021 for the mean household, and to 0.001 for lower-income households (25th percentile) as shown in column (3) in panel (*b*). Finally, panel (*c*) reports heterogeneity estimates for a 1% reduction in migration distance for different income levels. Remarkably, these are homogeneous across individuals with different incomes. Due to $\mathrm{MRS}_{p,\tau}$ decreasing in income, this nevertheless implies that high income households are willing to bear higher moving distances than low income households in order to avoid a tax increase. Generally, figures within a 1 kilometre radius exhibit the same qualitative pattern but are slightly more pronounced in quantitative terms.

The first-stage results associated with the estimates of the sorting model are reported in Table D2 in online Appendix D. Based on these estimates, further heterogeneity with regard to preferences can be observed. For instance, the interaction of the number of children with rents enters negatively while the interaction with taxes enters positively. This implies that households with children are willing to pay higher taxes for a reduction in rent per metre^2 compared to households without children. Another quantitatively important parameter pertains to the interaction of the number of children and the share of children in the neighbourhood, suggesting that spatial sorting is pronounced for families.

Finally, we would like to know how the estimates relate to the degree of tax capitalisation in rents. To summarise, we find an average $\mathrm{MRS}_{p,\tau}$ of −0.27 to −0.35. Using sample averages for monthly rent per metre^2 (20.943 CHF), residence size (88.691 metre^2), and the annual tax burden (13,724.432 CHF) according to Table 1, this corresponds to an average annual capitalisation rate between 44% and 57%, indicating that the capitalisation of taxes in rents is not complete.

5.4. *Counterfactual Experiments*

In this subsection, we perform two counterfactual experiments that illustrate the relative importance of income taxes for residential segregation in terms of income. In the first counterfactual scenario, we set all income taxes to a homogenous cantonal level. The second counterfactual experiment considers the effect of increased variation in income tax burden. The counterfactuals draw on the model for residential sorting and are compared to what we refer to as a benchmark (pre-experiment) equilibrium. We follow Bayer and McMillan (2012) and construct the benchmark equilibrium by predicting neighbourhood compositions based on revealed individual preferences over taxes, flat characteristics, local neighbourhoods, and optimal location choices according to the parameter estimates of Model I described above. For now, we assume perfect mobility in the counterfactual simulations and discuss below how the experiments may be affected by migration costs as included in Model III.

5.4.1. *Benchmark equilibrium*

Since the data do not necessarily correspond to an equilibrium, we start with the computation of a benchmark equilibrium which ensures that a spatial equilibrium situation is reached. With this at hand, we can adequately assess the location responses in the two counterfactual tax constellations. We use the results obtained in subsection 5.3 and solve for the rental price vector using (7), building upon Berry (1994) who shows that such a unique vector of market-clearing prices exists. To be precise, the rent of residence h in the benchmark equilibrium is obtained by employing the first-stage estimate of $\boldsymbol{\delta}_h$ and the estimated coefficients θ_b, γ_{0x}, γ_{0p}, $\gamma_{0\tau}$. Given the set of market-clearing rents denoted by $\bar{p}_h$, we calculate the choice probabilities $\bar{P}_h^i$ using the estimates in λ_h^i jointly with the estimated $\boldsymbol{\delta}_h$, whereby the error components u_h^i and ξ_h are set to zero.

The model displays a remarkable goodness-of-fit in predicting the location choices as we assign almost 90% of the households their correct housing choice. Accordingly, we may think of the benchmark equilibrium as an appropriate representation of the equilibrium distribution of households across space in the data. As we aim to assess the degree of income sorting under different tax regimes, it appears instructive to measure the exposure of individual households to different income groups (or other sociodemographic groups). The exposure rates in the benchmark equilibrium reveal the degree of income sorting under the current tax regime and can be compared to those in the counterfactual experiments. For this, the predicted probabilities (predicted demand) for each neighbourhood are aggregated by households' income quartiles. This income-specific probability mass is divided by the total probability mass over all income categories assigned to the respective neighbourhood. For each household, we then compute the exposure to each income category using the predicted location probabilities and the predicted income compositions of these locations. For instance, take a low-income household i which is assigned a range of different neighbourhoods. For each of these neighbourhoods, we compute the predicted income composition. Hence, i's exposure rate to high-income neighbourhoods is the sum of probability-weighted shares of high-income earners in i's predicted neighbourhood choices.[20] Table 4 describes the income exposure rates along the quartiles of household income. As is evident from the Table, income stratification is prevalent. For instance, the exposure of an average household in the first quartile to households in the same quartile is 28.7%, whereas the one to households in the fourth quartile of the income distribution is only 21.4%. This implies that the chance for a low-income household to live next to another low-income household is 34% higher than the probability that it lives next to a household in the fourth quartile. Similarly, high-income households are disproportionately exposed to wealthier neighbourhoods

[20] Note that we net out the household's own probability mass when computing the neighbourhood's income composition.

Table 4

Pre-experiment Exposure Rates

	Household income level			
Exposure rates (%)	1. Quartile	2. Quartile	3. Quartile	4. Quartile
Neighbourhood income 1. Quartile	28.672	26.163	23.448	21.230
Neighbourhood income 2. Quartile	26.276	25.682	24.517	23.572
Neighbourhood income 3. Quartile	23.637	24.536	25.607	26.346
Neighbourhood income 4. Quartile	21.416	23.619	26.427	28.852

Note. The Table reports average exposure rates (in %) of household income categories (column dimension) to neighbourhood income categories (row dimension) as described in subsection 5.4 for the year 2012 in the canton of Bern.

as the typical neighbourhood of a household in the upper quartile consists to 28.8% of households in the same quartile and only to 21.2% of ones in the first.

For every counterfactual experiment, we can now compute a new set of housing rents and choice probabilities that are used to update the equilibrium neighbourhood composition and the according exposure rates.

5.4.2. *Homogeneous taxes*

We set τ_m equal to the cantonal average and employ (7) to obtain the corresponding vector of rents which is used to compute the probabilities for each household to reside in each neighbourhood. This allows us to:

(*i*) shed light on the changes in these probabilities for neighbourhoods that increase or decrease the tax burden according to the experiment; and to

(*ii*) analyse the effect on the local income composition by calculating changes in the exposure rates compared to the benchmark equilibrium described above.

A homogenisation of local income taxes to the cantonal average implies that about half of the municipalities face a tax increase and the other half a decrease. On average, this amounts to a 0.4% (0.3%) increase (decrease) of the tax burden in those municipalities that are characterised by a below (above) average tax burden. These modifications of local tax burden directly translate into new market-clearing rents. Denote by $\tilde{p}_h$, $\tilde{P}^i_h$ the housing rents and choice probabilities in the counterfactual simulation. Then the correlation between the change in housing rents and the change in the probability to opt for the corresponding housing choice is $\mathrm{Corr}[(\tilde{p}_h - \bar{p}_h)/\bar{p}_h, (\tilde{P}^i_h - \bar{P}^i_h)/\bar{P}^i_h]$. This equals 0.828 for households in the fourth quartile of the income distribution whereas the analogous correlation for households in the lowest quartile of the distribution is -0.729. Consistent with this observation, the correlation between the change in probability mass and the change in taxes, $\mathrm{Corr}[(\tilde{\tau}_m - \bar{\tau}_m)/\bar{\tau}_m, (\tilde{P}^i_h - \bar{P}^i_h)/\bar{P}^i_h]$, equals -0.826 in the fourth quartile of household income and 0.734 in the first quartile. The figures indicate that homogenising local income taxes induces

Table 5

Change of Exposure Rates in Counterfactual Exercises

	Household income level			
	1. Quartile	2. Quartile	3. Quartile	4. Quartile
Panel (*a*): change of exposure rates with homogeneous income taxes (%)				
Neighbourhood income 1. Quartile	−2.897	−0.670	1.050	2.424
Neighbourhood income 2. Quartile	−0.631	−0.319	0.248	0.609
Neighbourhood income 3. Quartile	1.024	0.272	−0.466	−0.834
Neighbourhood income 4. Quartile	2.504	0.717	−0.833	−2.199
Panel (*b*): change of exposure rates with increased variation in income taxes (%)				
Neighbourhood income 1. Quartile	4.314	0.994	−1.656	−3.581
Neighbourhood income 2. Quartile	0.971	0.201	−0.312	−0.750
Neighbourhood income 3. Quartile	−1.613	−0.336	0.579	1.378
Neighbourhood income 4. Quartile	−3.671	−0.860	1.389	2.952

Notes. In panel (*a*) we set the income tax burden in all municipalities to the cantonal average. In panel (*b*) we raise the tax burden by one standard deviation in those municipalities with above-average tax burden and reduce it by one standard deviation in those with below-average tax burden. The Table reports changes in average exposure rates (in percentage points) of household income categories (column dimension) to neighbourhood income categories (row dimension) relative to the pre-experiment exposure rates displayed in Table 4. For details see subsection 5.4.

high-income households to move to locations that face a reduction in tax burden and a corresponding increase in rents whereas low-income households find it attractive to move to locations that are characterised by a tax increase.

Panel (*a*) of Table 5 illustrates the effect of homogenising taxes on income sorting across neighbourhoods by reporting the changes of income exposure rates between the benchmark equilibrium and the equilibrium with homogeneous taxes. According to the model, low-income households experience an increase of about 2.5 percentage points in their exposure to high-income households. At the same time, the neighbourhood an average high-income household is exposed to consist of about 23.6% low-income households when income taxes do not vary by municipalities while the model predicts this share to be 21.2% in the existing tax constellation. Hence, a homogenisation of income taxes across space would yield a moderate reduction in income sorting even though stratification along the income distribution would not fully vanish. The remaining differences can be traced back to differences in the existing housing stock which is valued differently depending on income and location characteristics (such as the distance to the CBD).

5.4.3. *Increased variation in taxes*

The second counterfactual experiment considers the effect of increased variation in the income tax burden. This is achieved by raising the tax burden by one standard deviation in those municipalities with above-average tax burden and reducing it by one standard deviation in those with below-average tax burden. Compared to the first experiment, this experiment generates somewhat more pronounced changes in the tax burden of almost 5%. Accordingly, we observe more pronounced location responses. Panel (*b*) of Table 5 reports the differences in exposure rates compared to the benchmark equilibrium. We observe that income stratification increases considerably due to increased variance in taxes. For instance, mixing of lower with high-income households drops by 3.7 percentage points while the probability of a high-income household to reside next to one in the same income category increases by almost 3 percentage points.

So far, we have assumed perfect mobility, yet we can use Model III to consider the role of migration costs. The residence characteristics in (7) and (8) contain the log distance to the respective alternative. Let us characterise the distance π the average household is willing to migrate in order to avoid a 1% tax increase while holding the rental rate constant by $\mathrm{d}V_h^i = -\gamma_{0\pi}\mathrm{d}\pi - \gamma_{0\tau}\mathrm{d}\tau - \gamma_{0p}\mathrm{d}p = 0$, where $\mathrm{d}\tau = 1$ and $\mathrm{d}p = 0$. Since the tax burden as well as π are measured in logarithmic terms, we obtain the average marginal willingness to pay for a 1% reduction in migration distance from Table 3 as $\gamma_{0\pi}/\gamma_{0p} = 0.299/0.017 \cong 17.6$. This elasticity seems high but, given that only 6% of the individuals in our data moved and the average moving distance is only 15, 8 kilometres, a 1% increase in tax burden corresponds to a moving distance of 2.8 kilometres. These back-of-the envelope calculations should be interpreted cautiously as our setting does not account for fixed costs of moving. Nevertheless, the results suggest that moving costs are sizable.

6. Sensitivity

In this Section, we further test the sensitivity of the results along the following lines.

6.1. *Inclusion of Property Taxes and Supply*

We cannot rule out that other factors confound the estimated coefficients on the income tax elasticity. A candidate that may bias the coefficient on income taxes is the property tax as municipalities might compensate low-income tax rates by levying higher property taxes. We account for this potential concern by collecting information about property taxes (measured in per mille) and including them in the analysis. In addition, we relax the assumption of fixed supply. We include a variable measuring log supply of newly built residences at the grid-level in the regressions. The respective results are reported in Table 6.[21] Overall, the estimates are very well in line with those reported in Table 2. As the inclusion of property tax rates leaves our results unchanged, we cautiously conclude that municipalities do not seem to compensate for

[21] It is noticeable that property taxes are levied only in a fraction of cantons (for a detailed description of the data, see online Appendix B). We furthermore plotted property against income taxes for the canton of Bern and found a slight positive correlation between the two. Log supply is computed based on information from the official land registers in the cantons of Zurich and Bern. The results are robust to different measures, as well as to lags and leads of supply. Note that the sample size is considerably reduced by the inclusion of supply as the information is limited to two cantons. In any case the assumption of supply not differing across borders is valid as shown in Figure 6, rendering the BDD approach valid.

Table 6

Sensitivity to Property Taxes and Supply

	BDD 1 km		BDD 0.6 km		BDD 0.3 km	
	(1)	(2)	(3)	(4)	(5)	(6)
Tax elasticity	−0.354***	−0.425***	−0.266***	−0.305***	−0.171***	−0.256*
	(0.037)	(0.099)	(0.049)	(0.100)	(0.056)	(0.152)
Property tax	0.009	−0.022	0.031**	0.007	0.037	−0.015
	(0.017)	(0.022)	(0.015)	(0.023)	(0.023)	(0.040)
Supply new buildings	no	yes	no	yes	no	yes
Adj. R^3	0.533	0.475	0.560	0.506	0.606	0.555
Observations	695,014	304,221	305,109	127,919	99,860	40,658
No. b,t	7,409	2,773	5,415	1,940	3,541	1,261

Notes. ***, **, * denote statistical significance at the 1%, 5% and 10% level respectively. Tax elasticity refers to income tax elasticity. Standard errors are multiway clustered at the level of municipality and year. Each specification refers to BDD estimates, covers the years 2005–12, and includes border-point-year fixed effects, log area, log population and the share of second homes by municipality, as well as the residence covariates summarised in Table 1. Even columns include the property tax at the municipality level as an additional explanatory variable. Uneven columns include the property tax and log supply of new buildings at the level of 1×1 kilometre^2 grids.

lower income taxes by increasing property taxes and that income sorting is the remaining channel for taxes to differ. The omission of supply seems to somewhat bias the $MRS_{p,\tau}$, however the corresponding decrease is due to a selection effect as estimating the regressions on the sample of municipalities with non-missing information about supply leads to slightly lower elasticities as well. We refrain from reporting and interpreting the coefficient on supply as it is simultaneously determined with demand. The sign and significance of the coefficient on the property tax rate is not stable, and the coefficient is small in magnitude.

6.2. *Matching Residences Across Boundaries*

As an alternative approach, we follow Fack and Grenet (2010) and Gibbons *et al.* (2013) and match each posting h with a counterfactual h' separately for each year. The counterfactual rental offer $p_{h't}$ is calculated as the inverse distance-weighted mean of the rents of all residences k that are located within a certain radius and are posted in the same year as offer h but belong to municipality m' rather than to m. The methodology is presented in detail in online Appendix C and the corresponding results are reported in Table C1. Generally, the elasticities are very similar to the ones in Table 2 where in terms of geographic precision the 1 kilometre (600 metres) window in the matching residences across boundaries (MBDD) comes closest to the 600 metres (300 metres) window in the BDD.[22] Hence, this approach allows us to exploit variation in rents for units that are situated just a stone's throw away from each

[22] This is because the BDD exploits two residences on both sides of the border and 600 metres (300 metres) distance from the border which feature an approximate distance between each other of 1.2 kilometres (600 metres) as used in the MBDD.

other, with an average distance between reference and counterfactual unit of only 240 metres in the smallest window.

6.3. *The Role of Schooling*

At the municipality level in Switzerland, the only relevant public good that appears excludable is elementary schooling even if in practice municipal borders are rather permeable regarding school choice. Secondary and tertiary education are entirely determined at the cantonal (state) level. Note also that schooling is primarily public in Switzerland, and the average quality of public schools is perceived superior to private schools.[23] As mentioned in Section 2, a large degree of homogeneity across municipalities is guaranteed in elementary schooling as well such that the main reason for differences in schooling should be driven by the composition of pupils which is captured by the design in the previous subsection. Yet, to address further concerns about the possibility of quality differences in elementary schools, we limit our analysis to residences along boundaries that are located within the same school district yet face different levels of income taxes.[24] Since the number of municipality pairs drops considerably when using only boundaries within the same school district, we focus on observations within the 1 kilometre and 600 metres windows. The corresponding results are summarised in Table C2 in online Appendix C. Limiting the sample to units within the same school district, the benchmark MBDD estimates decrease somewhat but confidence intervals overlap. These negligible differences are striking in particular when considering the substantial change in sample composition which affected among others the moments of the tax differentials. The estimate of β_τ is well in line with our previous findings and suggests that unobservable school differences are not confounding our estimates.

6.4. *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 at 500 metres and 150 metres from the true boundary point for either municipality within a (border-point-specific) pair. Then, we reassign the newly treated observations the low tax burden and the newly non-treated observations the high tax burden. Table D3 in online Appendix D reports the corresponding results for the BDD model and for a shift of the true boundary towards the low or high-tax municipalities. Hence, in the former case, we contaminate the treated and in the latter case we contaminate the control units. None of the estimates is significant and the coefficients’ magnitudes are far from our benchmark.

[23] According to BFS, only 2.7% of elementary school pupils were registered in a private school in 2012/3.

[24] To do so, we digitalise maps on school districts for cantons where applicable and link this information to our data. This is done for the Cantons of Zurich, Bern, Aargau, Fribourg and Vaud where the boundaries of school attendance zones do not always coincide with municipal boundaries.

6.5. *Functional Form Misspecification*

In line with most regression discontinuity approaches, we account for spatial trends that lead to average rent differences across boundaries by including: (*i*) a cubic polynomial function of distance to *b*; and (*ii*) the geographic coordinates of locations. This accounts for asymmetric rent gradients in low and high-tax municipalities. Thus, it allows for the possibility that rents are decreasing in distance from the high-tax centre where public goods are provided while the distance from a low-tax centre is irrelevant or enters even with an opposite sign. Results presented in columns (1) and (2) of Table D4 in online Appendix D are nearly identical and confirm the robustness of the estimated tax elasticities to the inclusion of flexible forms of distance to the boundary.

Furthermore, non-linearities of taxation may apply. 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 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. We approach the potential nonlinear relationship between rents and local income taxes: (*i*) by including a quadratic term of the tax differential to the MBDD specification; and (*ii*) by focusing on comparisons of observations h and their counterfactual i', where the tax differential exceeds the median tax differential in the sample, amounting to 731 CHF. Columns (3) and (4) of Table D4 suggest that both approaches leave previous results qualitatively unaffected.

6.6. *Endogeneity of Municipality Boundaries and Taxes*

Due to municipality mergers, the data are geo-referenced by year. If municipality mergers were more or less likely for neighbours 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 D4.

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 is unlikely that the population composition featured discontinuities in the absence of municipal boundaries and natural irregularities as the costs of social interactions are a smooth function of distance. Yet, our measures of house characteristics as well as sociodemographics may be incomplete. As an additional test, we include municipality fixed effects that absorb all time-invariant differences across municipalities in the specifications and identify the tax elasticity exclusively from variation in the level of taxes over time. This should moderate the issue as households need time to move as well as to adjust housing characteristics to the new tax differential (Baum-Snow and Ferreira, 2014). The corresponding results are reported in column (6) of Table D4. The coefficients remain quantitatively very similar to our benchmark results. Note that municipality fixed effects capture average sociodemographics which determine the preferences of the electorate but they do not reflect the distribution of households within municipalities, i.e. at the boundaries as addressed in the sorting approach above.

6.7. *Heterogeneity in Households and Residences*

As a final sensitivity check, we study the heterogeneity of tax elasticity with regard to different residence types. We observe heterogenous responses to income taxes for different residence types when splitting the sample into quartiles based on residence size and estimating the tax elasticity separately for each of the four quartiles. Table D5 in online Appendix D reports the corresponding results and shows throughout all specifications an increased responsiveness of smaller residences which becomes more accentuated the smaller the distance band we choose. This points to an increased mobility of smaller households compared to larger ones and is consistent with household size determining migration cost which turns out significant in the first stage of the sorting model (see Table D2 in online Appendix D).[25]

7. Concluding Remarks

Households exhibit heterogenous preferences over housing, taxes, the quality of local public goods and services, local amenities and the sociodemographic characteristics of their neighbours. This article sheds light on households' responsiveness to income tax differentials across municipalities in Switzerland and the resulting income composition of neighbourhoods. The degree of income tax capitalisation 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. This article has corrected for unobservable location characteristics and has accommodated the estimation of tax elasticity by way of a BDD in a discrete choice framework of household sorting using comprehensive household-level data on rents and sociodemographic characteristics.

We estimate an average income tax elasticity of about −0.27 to −0.35, corresponding to capitalisation rates of 44–57%. These estimates amount to about two thirds of the estimate from conventional hedonic regressions, pointing to the important role of unobservable location and neighbourhood characteristics. Our results show that high-income individuals have a higher marginal willingness to pay for lower taxes and thus systematically sort into low tax jurisdictions. This is consistent with the location choice model developed in the article that incorporated non-homothetic preferences over housing. The framework is instrumental to examining the consequences of tax reforms for housing rents and for the spatial equilibrium distributions of heterogeneous households. A homogenisation of local income taxes is shown to yield a moderate reduction in income sorting across

[25] This split also accounts for the potential concern that rents for larger flats (for which demand is generally lower and for which purchasing is an attractive substitute) may be negotiated to some extent.

jurisdictions. Following such an experiment, high-income households would move to neighbourhoods previously characterised by high taxes and inhabited mostly by low-income households. The reverse holds true for the location response of low-income households.

FINMA
University of Bern
Swiss Federal Institute of Technology ETH Zurich

Additional Supporting Information may be found in the online version of this article:

Appendix A. Data on Housing Rents.
Appendix B. Data on Municipal Taxes and Fees.
Appendix C. MBDD Methodology.
Appendix D. Additional Results.
Appendix E. Sociodemographic Census Data.
Data S1.

References

Agrawal, D.R. and Hoyt, W.H. (2014). 'Commuting and taxes: theory, empirics, and welfare implications', Working Paper No. 4852, CESifo.
Ahlfeldt, G.M. and Holman, N. (2015). 'Distinctively different: a new approach to valuing architectural amenities', SERC Discussion Paper No. 0171, Spatial Economics Research Centre, LSE.
Banzhaf, H.S. and Farooque, O. (2013). 'Interjurisdictional housing prices and spatial amenities: which measures of housing prices reflect local public goods?', *Regional Science and Urban Economics*, vol. 43(4), pp. 635–48.
Baum-Snow, N. and Ferreira, F. (2014). 'Causal inference in urban and regional economics', Working Paper No. 20535, NBER.
Bayer, P. and McMillan, R. (2012). 'Tiebout sorting and neighborhood stratification', *Journal of Public Economics*, vol. 96(11), pp. 1129–43.
Bayer, P., McMillan, R. and Rueben, K. (2004). 'An equilibrium model of sorting in an urban housing market', Working Paper No. 10865, NBER.
Bayer, P., Ferreira, F. and McMillan, R. (2007). 'A unified framework for measuring preferences for schools and neighborhoods', *Journal of Political Economy*, vol. 115(4), pp. 588–638.
Berry, S.T. (1994). 'Estimating discrete-choice models of product differentiation', *The RAND Journal of Economics*, vol. 25(2), pp. 242–62.
Berry, S., Levinsohn, J. and Pakes, A. (1995). 'Automobile prices in market equilibrium', *Econometrica*, vol. 63 (4), pp. 841–90.
Black, S.E. (1999). 'Do better schools matter? Parental valuation of elementary education', *Quarterly Journal of Economics*, vol. 114(2), pp. 577–99.
Brülhart, M., Bucovetsky, S. and Schmidheiny, K. (2014). 'Taxes in cities', Cahiers de Recherches Economiques du Département d'Econométrie et d'Economie politique (DEEP) No. 14.04, Université de Lausanne, Faculté des HEC.
Cameron, C.A., Gelbach, J.B. and Miller, D.L. (2011). 'Robust inference with multiway clustering', *Journal of Business & Economic Statistics*, vol. 29(2), pp. 238–49.
Dachis, B., Duranton, G. and Turner, M.A. (2012). 'The effects of land transfer taxes on real estate markets: evidence from a natural experiment in Toronto', *Journal of Economic Geography*, vol. 12(2), pp. 327–54.
Deaton, A. and Muellbauer, J. (1980). 'An almost ideal demand system', *American Economic Review*, vol. 70(3), pp. 312–26.
Duranton, G., Gobillon, L. and Overman, H.G. (2011). 'Assessing the effects of local taxation using microgeographic data', Economic Journal, vol. 121(555), pp. 1017–46.

Fack, G. and Grenet, J. (2010). 'When do better schools raise housing prices? Evidence from Paris public and private schools', *Journal of Public Economics*, vol. 94 (1–2), pp. 59–77.
Feld, L.P. and Kirchgässner, G. (1997). 'Die Kapitalisierung von Steuern und öentlichen Leistungen in den Mietzinsen: Eine empirische Überprüfung der Tiebout-Hypothese für die Schweiz', in (H. Schmid and T. Slembeck, eds.), *Finanz- und Wirtschaftspolitik in Theorie und Praxis*, pp. 64–92, Bern: Haupt.
Feld, L.P. and Kirchgässner, G. (2001). 'Income tax competition at the state and local level in Switzerland', *Regional Science and Urban Economics*, vol. 31(2–3), pp. 181–213.
Gibbons, S., Machin, S. and Silva, O. (2013). 'Valuing school quality using boundary discontinuities', *Journal of Urban Economics*, vol. 75(C), pp. 15–28.
Gibbons, S. and Overman, H.G. (2012). 'Mostly pointless spatial econometrics?', *Journal of Regional Science*, vol. 52(2), pp. 172–91.
Glazer, A., Kanniainen, V. and Poutvaara, P. (2008). 'Income taxes, property values, and migration', *Journal of Public Economics*, vol. 92(3-4), pp. 915–23.
Gordon, R.H. (1983). 'An optimal taxation approach to fiscal federalism', *Quarterly Journal of Economics*, vol. 98(4), pp. 567–86.
Hilber, C.A. (2011). 'The economic implications of house price capitalization: a survey of an emerging literature', SERC Discussion Paper No. 0091, Spatial Economics Research Centre, LSE.
Hilber, C.A. and Vermeulen, W. (2016). 'The impact of supply constraints on house prices in England', Economic Journal, vol. 126(591), pp. 358–405.
Ioannides, Y.M. (2004). 'Neighborhood income distributions', *Journal of Urban Economics*, vol. 56(3), pp. 435–57.
McCrary, J. (2008). 'Manipulation of the running variable in the regression discontinuity design: a density test', *Journal of Econometrics*, vol. 142(2), pp. 698–714.
McFadden, D. (1973). 'Conditional logit analysis of qualitative choice behavior', in (P. Zarembka, ed.), *Frontiers in Econometrics*, pp. 105–42, New York: Academic Press.
McFadden, D. (1978). 'Modelling the choice of residential location', in (A. Karlqvist, F. Snickars and J. Weibull, eds.), *Spatial Interaction Theory and Planning Models*, pp. 75–96, New York: North Holland.
Morger, M. (2013). 'Heterogeneity in income tax capitalization and its effects on segregation within Switzerland', Working Paper, Swiss Federal Tax Administration.
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*, vol. 77(6), pp. 957–71.
Ross, S. and Yinger, J. (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*, pp. 2001–60, vol. 3, Amsterdam: Elsevier.
Rühli, L. (2013). 'Irrgarten Finanzausgleich. Wege zu mehr Effzienz bei der interkommunalen Solidarität', Avenir Suisse.
Schaltegger, C.A., Somogyi, F. and Sturm, J.E. (2011). 'Tax competition and income sorting: evidence from the Zurich metropolitan area', *European Journal of Political Economy*, vol. 27(3), pp. 455–70.
Schmidheiny, K. (2006). 'Income segregation and local progressive taxation: empirical evidence from Switzerland', *Journal of Public Economics*, vol. 90(3), pp. 429–58.
Stull, W.J. and Stull, J.C. (1991). 'Capitalization of local income taxes', *Journal of Urban Economics*, vol. 29(2), pp. 182–90.
Tiebout, C.M. (1956). 'A pure theory of local expenditures', *Journal of Political Economy*, vol. 64(5), pp. 416–24.