CESifo WORKING PAPERS

9445
2021

November 2021

Fiscal Autonomy and Self-Determination

Gabriel Loumeau, Christian Stettler

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9445*

Fiscal Autonomy and Self-Determination

Abstract

This paper studies the equilibrium effects of local fiscal autonomy accounting for benefits from self-determination. It proposes a quantifiable structural equilibrium framework in which imperfectly mobile heterogeneous households sort themselves across jurisdictions under endogenous public good provision. We calibrate the framework to fit the economic and geographic characteristics of the Canton of Bern using household-level data. In particular, we exploit quasi-natural policy variation in voting rights to quantify benefits from self-determination, and employ machine learning methods to accurately represent the local political process. We find that restricting local fiscal autonomy decreases welfare for (almost) all households.

JEL-Codes: H710, H770, R510.

Keywords: fiscal autonomy, self-determination, decentralization, household, equilibrium, quasi-natural variation.

*Gabriel Loumeau**
ETH Zürich, Department of Management, Technology, and Economics
Leonhardstrasse 21
Switzerland – 8092 Zurich
loumeau@ethz.ch

Christian Stettler
ETH Zürich, Department of Management, Technology, and Economics
Leonhardstrasse 21
Switzerland – 8092 Zurich
stettler@ethz.ch

*corresponding author

November 24, 2021
We are particularly grateful to workshop participants in Lugano, Munich and Zurich. In particular, we thank (in alphabetical order) Simon Büchler, Maximilian von Ehrlich, Sergio Galletta, Marko Köthenbürger, Patrick Moran, Costanza Naguib and Kurt Schmidheiny for helpful comments and suggestions.

1 Introduction

The economic literature on fiscal decentralization and local tax autonomy generally ignores benefits from self-determination; that is to say, the benefits of letting individuals decide locally on their own. Yet, benefits from self-determination are a key motivation for decentralization in the first place. As Henkin (1987) puts it: "*[Decentralization] permits more meaningful individual participation in self-government.*" Similarly the Preamble of the European Charter of Local Self-Government (1985) states that "*[...] local authorities with real responsibilities can provide an administration which is both effective and close to the citizen.*"[1] Instead, the literature tends to trade off efficiency gains and negative externalities implied by decentralization.[2] Whereas such trade-off is evidently relevant, the welfare effects of (de)centralization are particularly unclear when considering self-determination benefits. In that context, is decentralization welfare-improving? Does accounting for benefits from self-determination leads to more stable conclusions? Are welfare gains/losses from decentralization evenly distributed?

We answer these questions by developing and applying a realistic quantifiable spatial equilibrium framework of fiscal autonomy accounting for benefits from self-determination. In the framework, heterogeneous households (in any finite number of dimensions) chose where to reside and work, as well as the local public finance outcomes at their place of residence. The choice of local public outcomes includes multiple tax instruments (i.e., tax multipliers on personal income and housing), the level of public spending, and is impacted by horizontal fiscal equalization. When choosing, households face heterogeneous migratory and commuting costs. A better/worse tax-benefit linkage is allowed to capitalize into housing prices.

To calibrate the framework, we profit from two highly precise data-sets. First, we employ unique data on the universe of about 450'000 households in the Canton of Bern for the years from 2012 to 2016. We construct this data-set by merging register-based population census data to the entire earning records from the Swiss Old-Age and Survivors' Insurance. Our final data-set contains a wide array of household characteris-

[1]By 1988, 47 countries of the Council of Europe had ratified the charter. For more information, see original text and documentation: here.

[2]Key contributions include, among others, Hamilton (1975, 1976); Davoodi and Zou (1998); Hoyt and Lee (2003); Faguet (2004); Enikolopov and Zhuravskaya (2007); Barankay and Lockwood (2007); Albouy (2009); Rodríguez-Pose and Ezcurra (2010); Calabrese et al. (2012); Fajgelbaum et al. (2019). Epple and Nechyba (2004) survey the literature.

tics, including *joint household income, average age, the number of adults and children, the municipality of residence, and the municipality of origin*. This data allows us to accurately model households' preferences for their municipality of residence, tax rates and spending levels. Second, we complement our data on the universe of households by annual survey data from the Swiss Households Panel (SHP, Voorpostel et al., 2020) covering the ten annual waves between 2000 and 2009. This data allows us to estimate precisely the utility benefits from self-determination at the individual level in a quasi-natural setting. We also use publicly available local public finance information, as well as municipality-to-municipality travel times on the road network.

Equipped with the calibrated framework, we simulate two reforms affecting local fiscal autonomy: (i) tax harmonization via upper and lower bounds imposed to local tax multipliers, (ii) minimum mandatory spending levels per capita. We find that fiscal decentralization increases aggregate welfare significantly. Fully harmonizing local taxes and imposing high mandatory spending levels lead to an aggregate welfare decline of -0.8% and -0.5%, respectively. This result holds for parameter values determining the benefits from self-determination as low as 20% of the quasi-experiment-based estimates. Almost all household-groups benefit from higher local fiscal autonomy. This is particularly true for richer households with less than two children. Households with three or more children are, *ceteris paribus*, the only group experiencing an increase in welfare when the degree of fiscal decentralization decreases (both via tax harmonization or mandates on local public spending).

Main contributions. This paper innovates in three main dimensions.

First, it proposes a spatial quantifiable framework of local public finance with realistic geography. As such, this paper contributes to bridging the gap between the vast literature on Tiebout-type equilibria following Tiebout (1956) and Oates (1969),[3] and the spatial equilibrium literature following Rosen (1979) and Roback (1982).[4] The framework is realistic in that (i) households heterogeneity may be defined in any finite number of dimensions (including heterogeneous migratory and commuting costs), (ii)

[3]Key contributions in this literature include, among others, Brueckner (1982); Epple and Platt (1998); Banzhaf and Walsh (2008); Schmidheiny (2006b); Calabrese et al. (2012). Surveys of the literature are offered in Epple and Nechyba (2004), Ross and Yinger (1999), and Rubinfeld (1987).

[4]Fajgelbaum et al. (2019) study how local taxes distort individuals' and firms' location choices in a spatial equilibrium framework. Albouy (2009) and Brülhart et al. (2021) study how mobile workers and individuals bear the burden of taxation. Redding and Rossi-Hansberg (2017) survey the quantitative spatial equilibrium literature.

it can be applied to any real world geography setting, (iii) its representation of local public equilibrium matches the complexity of observed settings with endogenous public spending levels, multiple tax instruments, and fiscal equalization, (iv) it allows for endogenous capitalization effects of tax rates and public spending into housing prices. Moreover, it includes commuting costs which have often been ignored in local public finance equilibrium frameworks. Finally, to represent crowding out due to limited land in a static context, we propose a "rich chooses first" residential allocation mechanisms. This also allows for effective zero density in a conditional logit approach to residential choice.

Second, it is the first to account structurally for utility benefits from self-determination when studying fiscal decentralization. Past research has often shy away from analyzing these mechanisms – despite acknowledging their importance – due to a lack in the availability of the necessary data. A notable applied exception is Flèche (2021) who study the effect of Swiss centralization reforms on residents' life satisfaction. In the baseline model, benefits from self-determination are modeled as a procedural utility: *utility gain from choosing for oneself* irrespective of the outcome.[5] This modelling assumption has a long tradition dating back to the birth of modern nations. The UN for instance often claims a broad "right of self-determination of all peoples" (see, among others, Lind and Tyler, 1988; Lane, 1988; Buchanan, 1992; Sen, 1999; Dahl, 2020). A wish to locally self-determinate motivates many local independence movements (e.g., Catalonia, Quebec, Scotland, etc.). The same reasoning applies to local constituents wishing to decide on their own how much tax to levy locally, and how much to spend for their local community.[6]

To retrieve the parameter governing the procedural utility benefits, we make use of an extensive household level survey in the context of a 2005 voting rights reform in the Canton of Geneva. Foreigners residing in Geneva and who have been resid-

[5]Benefits from self-determination may alternatively be defined as a local informational advantage, which we do in robustness Section B. As Oates (1999) puts it: "individual local governments are presumably much closer to the people and geography of their respective jurisdictions; they possess knowledge of both local preferences and cost conditions that a central agency is unlikely to have." To study the information advantage, we exploit an institutional particularity of the Canton of Bern: 93% of municipalities decide on local outcomes via a citizen's assembly. This includes many municipalities of more than 1,000 inhabitants and up to 11,000 inhabitants. The remaining 7% host a municipal parliament. This allows us to set a difference-in-difference estimation strategy – with population as running variable – to retrieve the parameter governing the informational advantage of local decision making.

[6]This consideration also intersects with the literature on heterogeneous groups and (national or regional) borders following the seminal paper by Alesina and Spolaore (1997). Panizza (1999), for instance, studies the consequences of two levels of government in a fiscal federation.

ing in Switzerland for 8 years or more were allowed to vote *locally* (even though they were not naturalized) from Jan. 1st, 2005. This reform gives rise to two quasi-natural settings, which we both exploit. First, in a *Regression Discontinuity Design* (RDD), we compare foreigners in Geneva just below and just above the 8 years threshold. Second, in a *Difference-in-Difference* (DiD) approach, we compare foreigners residing since more than 8 years in Geneva versus non-reformed Cantons (including Bern) before and after 2005. Both identification strategies deliver qualitatively and quantitatively very similar results. Using a voting rights reform to retrieve procedural utility parameters offers two advantages over more standard centralization reforms. First, it allows a more natural distinction between a reaction to gaining/losing voting power and a reaction to gaining/losing voting power on a given item. Second, the gain/loss of voting power is complete, as opposed to centralization reforms in which voter still have some (indirect) control on the centralized items.

Third, the last main innovation of this paper concerns the use of Machine Learning and ensemble methods to represent the local political process, and ultimately to predict local tax rates (Athey and Imbens, 2019; Athey et al., 2019; Varian, 2014). In a standard public finance setting (i.e., a large number of heterogeneous household types – defined as the cross-product of many household characteristics – and a comparatively small number of municipalities), we show that this approach substantially outperforms the standard linear model. It is also much more flexible in the modelling of preferences than the commonly used median voter based approach to local political decision making. This approach, as well as its performance, relies on the availability of precise and extensive household level data on the universe of households in the Canton of Bern for the years from 2012 to 2016. As key figure illustrating the performance of this approach, the mean absolute deviation of the out-of-sample prediction amounts to 3.75%, which is 5.94 percentage points below the measure for the ordinary least squares regression. In turn, the accuracy of the predictions of local tax multipliers allows us to predict household type local densities with great accuracy, as reveled in the model validation exercise in Section 5.

The paper proceeds as follows. We detail the institutional background in Section 2. We then present the framework in Section 3, before presenting the data used in Section 4 and calibrating the framework to fit Bern's economic and geographic characteristics in Section 5. Finally, we simulate reforms limiting fiscal decentralization and study their equilibrium effects in Section 6. Section 7 concludes.

2 Institutional background

In this section, we provide key facts about the Swiss institutional set-up that are reflected in our theoretical framework and our empirical approach. In doing so, we also highlight the reasons that make the Canton of Bern a particularly well suited setting to study fiscal decentralization.

2.1 Fiscal decentralization in Switzerland

The Swiss confederation features one of the highest degrees of fiscal decentralisation to the municipal level in the world (see Figure 1a for an international comparison). Swiss municipalities – the smallest political units – account for more than one fifth of the general governments total revenue and expenditure.[7] Consequently, municipal autonomy is also large. Municipalities independently manage a large number of public services, including schools, local planning and social welfare.[8]

As documented in Figure 1b, and in contrast to the local jurisdictions of most other federally organized countries, Swiss municipalities largely rely on residence-based income taxation (67% of total tax revenues), whereas property taxation represents only 9% of total tax revenues.[9] Income is taxed at the federal, cantonal and municipal level. Municipalities thus compete against each other and higher government levels for the same tax base.

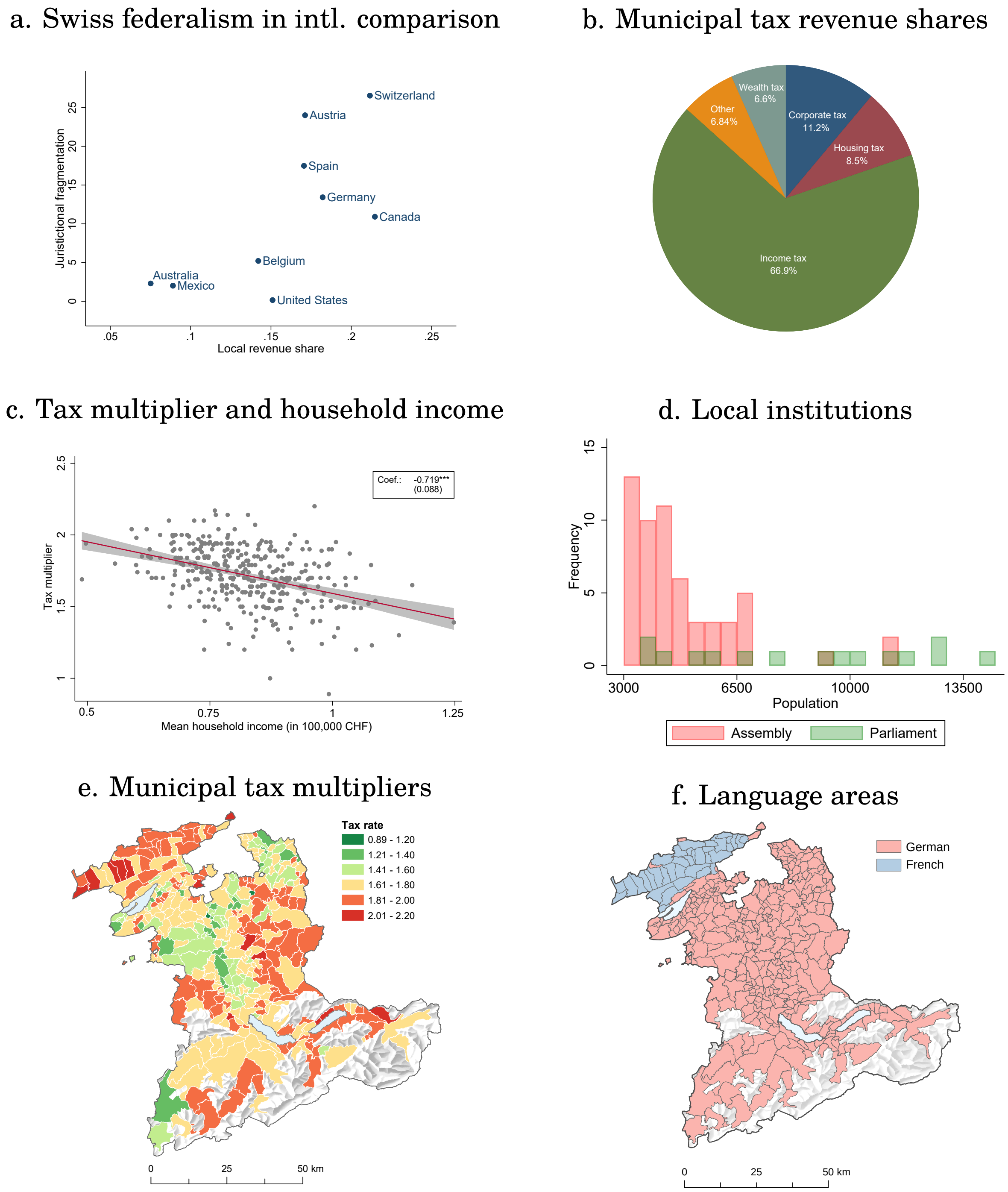
The federation and the cantons have the autonomy to design income tax schemes. This includes the degree of tax progression as well as exemptions and deductions. Municipalities are bound to the cantonal tax schemes. However, they decide on a scalar tax multiplier that applies to the cantonal tax schedule. Therefore, tax multipliers basically constitute the only instrument for municipal tax policy, which allows for perfect comparability of tax levels *within cantons*. As shown by Figure 1e, differences in municipal tax multipliers, and hence municipal tax rates, are substantial, ranging from 0.89 to 2.20. The unweighted mean tax multiplier across the canton's municipalities is 1.72. For a taxable income of CHF 100,000 this tax multiplier is equivalent

[7] Federal Fiscal statistics are taken from the Swiss Federal Finance Administration (SFFA): Switzerland's financial statistics for 2016 - Annual Report.

[8] While municipalities provide a wide range of public services, their services are often regulated through cantonal or federal regulations. Municipalities' real spending autonomy is therefore more limited than suggested by expenditure data. See Rühli (2012) for a comprehensive report on municipal autonomy in Switzerland.

[9] The remaining tax revenue stems from corporate (11%), wealth (7%) and other taxation.

Figure 1: STYLIZED FACTS



Notes on Figure 1a: The figure presents jurisdictional fragmentation (average number of municipalities per 100,000 inhabitants, 2016) and local revenue decentralization (local revenue as a share of general government revenue, 2016) for the 9 countries that are denoted by the OECD as federal or quasi-federal countries. *Source: Own calculations based on data from the OECD (Regional Statistics) and IMF (GFS). Notes on Figure 1b:* The figure shows municipal revenue from different taxes as a percentage of total municipal tax income for the 352 municipalities in the Canton of Bern and for 2016. *Source: Financial Statistics (FINSTA) of the Canton of Bern. Notes on Figure 1c:* The figure presents predicted values and 95% confidence intervals from a linear regression of municipal tax multiplier on mean household income (in hundred thousand CHF) for the 352 municipalities of Bern in 2016. *Notes on Figure 1d:* The figure provides a distribution of the municipalities in the canton of Bern across population size (for municipalities with a population between 2000 and 14500 inhabitants) by the type of legislative institution. Figure A1 shows the distribution across the full population level. *Notes on Figure 1e:* The figure shows the level of municipal tax multiplier for the 352 municipalities in the Canton of Bern and in 2016. Shaded gray areas are mountains. *Notes on Figure 1f:* The figure shows the main language for each of the 352 municipalities in the Canton of Bern and in 2016. Shaded gray areas are mountains.

to a municipal income tax rate of 7.8% (unmarried) and 6.6% (married).

Due to Switzerland's high degree of municipal tax decentralization, spatial income sorting and local tax competition are very active. Schmidheiny (2006a), Schmidheiny (2006b), Schaltegger et al. (2011) and more recently Basten et al. (2017) and Schmidheiny and Slotwinski (2018) show that high-income households sort into municipalities with low income tax rates and high housing prices. Figure 1c provides descriptive evidence specifically for the Canton of Bern. The figure plots the mean income of households of the 352 municipalities against the municipal tax multiplier. The red line represents the fitted value of an OLS regression of the tax multiplier on the mean household income. The slope coefficient is -0.72 and highly significant, suggesting a strong sorting of high income households into municipalities with low tax rates.

2.2 Why focusing on the Canton of Bern?

When studying local public equilibria consistently, it appears natural to focus on a State or Canton. First, it insures that the institutional setting is homogeneous. For instance, the income tax scheme – up to the municipal multipliers, as described above – and the task division between the municipal and the cantonal levels are constant across all municipalities, and so are incentives induced by the equalization system. At the same time, focusing on a smaller region than a State or Canton – such as a metropolitan area – forbids representing the equalization system accurately. Indeed, averages in tax revenues and transfers towards and from the equalization pools are computed using all municipalities in a Canton.

Given these considerations, we focus on the Canton of Bern (out of all Swiss Cantons) for the following reasons:

1. Largest Swiss Canton: With 352 municipalities, Bern is the largest Swiss Canton in terms of municipalities. The Canton thus allows us to maximize the number of local jurisdictions within an homogeneous institutional setting.

2. Local direct democracy in the Canton of Bern: A well-known feature of Switzerland's political system is its direct democratic institutions. Citizens regularly vote on a wide range of issues – e.g., public finances or infrastructure projects – at the municipal, cantonal and federal levels via referenda. What is less well-known is that the executive body of a large number of Swiss municipalities is a

citizen's assembly.[10] Even in that context, the Canton of Bern stands out. 94% of its municipalities (329/352) host an assembly as executive body. In practice, this means that Swiss residents vote via show of hands at annual assemblies on all important local decisions. As shown by Figure 1a parliaments tend to be the predominant legislative authority in larger municipalities and vice versa, although the two legislative types coexist across a large range of municipal population size. Assemblies are operative in many municipalities above 1,000 inhabitants and up to 11,000 inhabitants.

3. The Canton of Bern is particularly heterogeneous: As apparent from Figure 1e, the canton is separated by the German-French language border that has been frequently exploited to isolate cultural from institutional differences.[11] Moreover, cantonal and national referenda regularly reveal large differences in preferences on a wide range of topics.[12] Hence, even though total population count (1,026,513 inhabitants) and the geographic space (5959,5 km^2) of the Canton is small, a high level of preference heterogeneity for public outcomes is present.

3 Quantifiable framework of local fiscal autonomy

Consider an area that occupies a closed and bounded subset N of a two-dimensional space. The area comprises $\{n, i\} \in N$ municipalities whose location is given by their centroid, where n refers to municipaties as residential locations and i as commercial locations. Each municipality is endowed with a fixed amount of residential floor space, L_n. The area is populated by a total of Ω households of $\theta \in \Theta$ exogenous types. Θ is only required to be finite. Within the framework, household heterogeneity finds a translation in five dimensions: (i) joint income, (ii) location of origin, (iii) distance elasticity of migration, (iv) distance elasticity of commuting, (v) municipal-composition-specific preference for a tax-benefit linkage.[13]

[10] See Brülhart and Jametti (2019) for a detailed description of direct-democratic participation rights in Swiss municipalities.

[11] See for instance Eugster et al. (2011) (on the demand for social insurance), Gentili et al. (2017) (on elderly care), Steinhauer (2018) (on working mothers) and Eugster and Parchet (2019) (on the level of public services and taxes).

[12] See Political Atlas of Switzerland from the Swiss Federal Statistical Office for maps displaying the voting results on federal referenda.

[13] The last dimension allows households to have different preferences on tax-benefit linkages depending on the composition of the municipality.

3.1 Households

The utility of household $\omega \in \Omega$ of type θ residing in n and working in i is defined as follows:

$$U_{ni}^{\theta}(\omega) = \frac{b_n}{d_n^{\theta}} \left(\frac{c_{ni}^{\theta}}{\alpha}\right)^{\alpha} \left(\frac{l_{ni}^{\theta}}{1-\alpha}\right)^{1-\alpha} g_{ni}^{\theta}(\omega) \tag{1}$$

where c_{ni}^{θ} and l_{ni}^{θ} denote the consumption of a single final good and of housing by a household of type θ in municipality n. d_n^{θ} refers to the migratory costs incurred by a household of type θ by locating in n.

The utility derived by household ω of type θ in n from consuming public goods is denoted by $g_{ni}^{\theta}(\omega)$. It is defined as the product of an idiosyncratic household specific component, $z_{ni}^{\theta}(\omega)$; a component determining the utility from self-determination, O_n; and the average utility derived from public goods g_n^{ι}. ι shapes the utility responses to the level of public goods enjoyed. z_{ni}^{θ} is drawn from a Fréchet distribution such that $\Pr(z_{ni}^{\theta}(\epsilon) \leq z) = \exp(-z^{-\epsilon})$, with $\epsilon > 1$ being the shape parameter.

Gains from self-determination are modelled using a procedural utility approach such that: $O_n = 1 + \mathbb{1}(n)\rho$. It captures utility that residents derive due to a feeling of influence and self-determination when participating in the political decision-making process itself, irrespective of the outcome (Frey et al., 2004; Frey and Stutzer, 2005). Intuitively, O_n acts as an amenity shifter specific to a municipality and its household composition. The indicator function $\mathbb{1}(n)$ takes on one if citizens of municipality n decide autonomously on municipal spending or tax levels, and zero otherwise, while ρ denotes the respective percentage gains in procedural utility.

Formally, we then have:

$$g_{ni}^{\theta}(\omega) = z_{ni}^{\theta}(\omega) O_n g_n^{\iota}, \quad \text{with} \quad O_n \geq 1 \quad \text{and} \quad g_n > 0, \quad \forall n \in N. \tag{2}$$

Households' expenditures on the consumption good and housing can be expressed as a function of their wages (w_{ni}^{θ}); the share of revenues spent on either of the consumption good (α) or housing ($1-\alpha$); the municipal income tax rate levied on type θ ($t_n^{\theta} = m_n t^{\theta}$) which is the product of a municipal multiplier m_n and the average cantonal tax rate for households of type θ, t^{θ}; and the property tax rate, r_n. Type-specific commuting costs, d_{ni}^{θ}, are modeled as an effective reduction in final wages, such that: $w_{ni}^{\theta} = \frac{w^{\theta}}{d_{ni}^{\theta}}$. Q_n is the residential land rent in n. The price of the consumption good is the numeraire. Hence, indirect utility can be expressed as:

$$u_{ni}^{\theta}(\omega) = \frac{(1-t_n^{\theta})w_{ni}^{\theta}b_n}{d_n^{\theta}((1+r_n)Q_n)^{1-\alpha}}g_{ni}^{\theta}(\omega) \tag{3}$$

3.2 Household preferences for tax multipliers

Since Black (1948), a long tradition in local public finance has approached the local decision process on tax rates and public spending using the median voter theorem. Whereas the simplicity and flexibility of this approach explain its wide use, it suffers from important limitation when wanting to model accurately local choices. First, voters' preferences must follow strict functional forms: single-peaked preferences. Second, voters heterogeneity is limited to a few – one in most instances – number of dimensions. Third, it does not allow the preference of voters to differ depending on the composition of the municipality.[14] Finally, its predictive accuracy remains low.

We propose a novel Data Science-based approach to local decision making which does not suffer from the above-mentioned limitations. The proposed approach remains *a priori* agnostic on the form of voters' preferences. Voters' preferences may be defined in any (finite) number of dimensions, and may differ based on the demographic composition of the municipality. The approach is notably suitable to study cases with large voter heterogeneity θ and relatively small number of municipalities n.[15] Given the increased availability of individual level data, such flexibility is likely to become increasingly relevant. Finally, the predictive accuracy of this approach is high.[16]

Formally, allowing for two tax bases at the municipal level, the equilibrium income tax and the property tax multipliers (respectively, m_n and r_n) chosen in a given municipality are functions of the share of household types ($\frac{H_n^1}{H_n}, \frac{H_n^2}{H_n}, ..., \frac{H_n^{\Theta}}{H_n}$) residing in the municipality:

$$m_n = f_m\left(\frac{H_n^1}{H_n}, \frac{H_n^2}{H_n}, ..., \frac{H_n^{\Theta}}{H_n}\right), \quad \text{and} \quad r_n = f_r\left(\frac{H_n^1}{H_n}, \frac{H_n^2}{H_n}, ..., \frac{H_n^{\Theta}}{H_n}\right). \tag{4}$$

Importantly, (4) implicitly assumes that households are rational enough to recog-

[14] Among others, Alesina and Ferrara (2000) have documented that such interaction effects are relevant in local decision making.

[15] In the application, we study the local decisions of 250 household types located in 352 municipalities (Section 6).

[16] For instance, in the calibration exercise, we show that using this approach the mean absolute deviation of the out-of-sample prediction is around 40% below the measure for the ordinary least squares regression (Section 5.3).

nize the implications of municipal tax multiplier and property tax rate on public goods expenditure.

3.3 Public finance

Public budget clearing requires that total public good spending equals the effective revenue raised plus net transfers from the equalization mechanism. Therefore, public spending in municipality n amounts to:

$$g_n = y_n + \delta\Big(y - y_n^h\Big) \tag{5}$$

$y_n = \frac{Y_n}{H_n^\chi}$ denotes the effective tax income per household in municipality n. In this equation $Y_n = \sum_\theta^\Theta \pi_n^\theta H t_n^\theta w_n^\theta$ denotes the total tax income of municipality n and H_n denotes the total number of households in municipality n, with $w_n^\theta = \mathbb{E}[w_i^\theta|n]$. The parameter χ captures economies of scale in providing the public good and ranges from $\chi = 0$ (infinite economies of scale) to $\chi = 1$ (no economies of scale). Following the same logic, $y = \frac{Y}{H^\chi}$ denotes the average effective municipal tax income per household in the canton of Bern, where $Y = \sum_n^N Y_n$ is the aggregated tax income of all municipalities in the canton of Bern and H is the total number of households. $\delta \in [0, 1]$ denotes the equalization parameter. A larger value of δ implies a more redistributive equalization system.

The transferred amount is calculated based on the municipalities' harmonized, rather than absolute tax incomes.[17] Therefore, $y_n^h = \frac{Y_n^h}{H_n^\chi}$ with $Y_n^h = \sum_\theta^\Theta \pi_n^\theta H \overline{m} t^\theta w_n^\theta$ denotes the effective harmonized tax income of municipality n. It is the hypothetical effective tax income per household of municipality n if households were taxed at a population weighted average municipal tax rate. The amount that municipality n receives out of the equalization system thus primarily depends on its tax base, and only indirectly on its choice for the tax multiplier.

3.4 Mobility and commuting

Let us denote the attractiveness of a pair of municipalities $\{n, i\}$ for a given household type as v_{ni}^θ. Given that z_{ni}^θ enters multiplicatively in the indirect utility function, we

[17]This modelling choice follows the design of most existing tax equalization systems, including the one in place in the Canton of Bern.

can define the probability that a household of type θ working in i locates in n as follows:

$$\pi_n^\theta = \frac{\sum_i^N v_{ni}^\theta}{\sum_n^N \sum_i^N v_{ni}^\theta}, \quad \text{where} \quad v_{ni}^\theta = \left(\frac{(1-t_n^\theta) w_{ni}^\theta g_n^\iota B_n}{d_n^\theta ((1+r_n) Q_n)^{1-\alpha}} \right)^\epsilon. \tag{6}$$

In turn, the expected utility of households from type θ in the region is then:

$$\mathbb{E}[u]^\theta = \Gamma\left(\frac{\epsilon-1}{\epsilon}\right) \left[\sum_n^N \sum_i^N v_{ni}^\theta \right]^{1/\epsilon} \tag{7}$$

where Γ is the Gamma function.

3.5 Sorting, effective location choice, and spatial utility differences

In line with descriptive evidence in Section 2, we propose a "rich choose first" effective allocation mechanism. Without loss of generality, household types can be sorted based on their respective (pre-commuting) income, such that $\theta = 1$ is the type with the largest income, $\theta = 2$ the type with the second largest income, etc. Then, household types sequentially choose a residential location – starting with household type $\theta = 1$ – based on the probability that a household of type θ locates in n, π_n^θ, and under the condition that $H_n = \sum_\theta^\Theta H_n^\theta \leq L_n$. As locations get filled up, the effective location choice set offered to households from lower income groups shrinks. At the end, households from the last group do not actually choose a location, but are instead forced to locate where there is still housing space. At the end of the allocation process, local housing clearing then requires:

$$L_n = \sum_\theta^\Theta H_n^\theta, \quad \forall n. \tag{8}$$

3.6 Land clearing

Clearing of the housing market requires that demand and supply must equate. Denote the exogenous residential land mass and floor space available in n as $\tilde{L}_n$ and L_n, respectively. Further denoting the local intensity of residential development, referring to land regulations, as ϕ_n, we have:

$$L_n = \phi_n \tilde{L}_n = (1-\alpha) \frac{\sum_\theta^\Theta \mathbb{E}[w_i^\theta | n] H_n^\theta}{Q_n}. \tag{9}$$

3.7 Equilibrium

Brouwer's Fixed Point Theorem (1911) may be used to prove equilibrium existence and uniqueness. Formally, the following proposition holds:

Proposition 1. *Assuming strictly positive, finite, and exogenous characteristics, there exist unique general equilibrium vectors* $\{\pi^{\theta}, ..., \pi^{\Theta}, m, g, Q, O\}$ *solving the model.*

PROOF: See the proof of Proposition 1 in the Online Appendix C. *Q.E.D.*

4 Data description

For the calibration of our model we primarily draw on administrative data on the universe of individuals and households in the canton of Bern. We complement this data with survey and municipal level data. We provide a short summary of the different data sources and the data construction process in this section and refer to Online Appendix D for a complete description.

Matched earning-census data. We construct a data-set on the universe of households in the Canton of Bern. For this end, we merge the register-based population census data of Switzerland for the years between 2012 and 2016 via a social security number to 100% of the Social Security Earnings Records (SSER) from the Old-Age and Survivors' Insurance (OASI).[18] The SSER dataset contains the near universe of individual annual wages.[19] The register-based census data contain information about the individuals age, gender, the municipality of residence, and the municipality of origin. Moreover, individuals of the same households are linked through a household identification variable. Through this variable, we construct a data-set on the universe of households in the Canton of Bern during the five year period between 2012 and 2016. Finally, we add average tax rates per household based on the cantonal tax act and the municipal tax multiplier.[20]

[18]See, for instance, Martinez et al. (2021) who merge the same data sources, albeit for an earlier time period.

[19]The SSER dataset does not include information about individuals capital income and contributions from pension plans. However, the dataset covers the earnings record of individuals back to 1981. This allows us to approximate the income of retired individuals based on i) their full wage history before retirement and ii) distributions of pension plan income (first and second pillar) by household category.

[20]The tax act of the Canton of Bern applicable in 2016 (BSG 661.11 - Steuergesetz (StG)) is available at https://www.belex.sites.be.ch/frontend/versions/1040.

We provide summary statistics of the household variables for the year 2016 (our main year as described in detail in Section 5) in Panel A of Table A6 and descriptive statistics for the entire time range from 2012 to 2016 in Table A7. In 2016, the canton counted 445'393 households. Over the period of five years, the mean nominal household wage increased by about 1.2% from CHF 98,825 in 2012 to CHF 100,002 in 2016. Variation in the average age of adults, the number of adults and the number of children across years is small.

Survey data. We complement our data on the universe of households by annual survey data form the Swiss households Panel (SHP) (Voorpostel et al., 2020) covering the ten annual waves between 2000 and 2009. The SHP is a ongoing longitudinal survey within the framework of the Swiss Foundation for Research in Social Sciences (FORS). It contains a wide range of variables from interviews of all household members of a representative sample of private households in Switzerland. Over our 10 year analysis period the total sample size amounts to 121'622 individuals from 43'693 households. Both survey attrition and the non-response bias are moderate (Voorpostel and Lipps, 2011).

Among others, the survey covers questions on the respondents satisfaction with their political influence and their life in general (on a scale from 0 to 10). In addition, the data includes information on the interviewees age, gender, nationality, civil status, years of education, canton of residence, residence permit and the year of immigration to Switzerland. We use this data to infer utility that residents derive from municipal fiscal autonomy. A detailed description of the methodology is provided in Subsection 5.2.

Municipality data. We match data on municipal tax multipliers with data on municipal expenditure and the type of the legislative body (parliament or assembly) for a time period spanning the years from 1998 to 2016. Up to 2015 municipal tax multipliers have been compiled by Parchet (2019). Tax multipliers of more recent years are publicly available from the Swiss Federal Tax Administration (SFTA). Municipal expenditure data and data on the type of the legislative body are publicly accessible at the website of the cantonal administration.[21]

[21]We thank Raphael Parchët for the generous provision of data on local tax multipliers. For the years since 2016 data are available: here. Municipal expenditure data are available: here. Data on the type of the legislative body are available: here.

Panel B in Table A6 provides descriptive statistics. As generally in Switzerland, most municipalities of the Canton of Bern are very small in international comparison.[22] However, municipalities are strongly heterogeneous in the number of households. While 77 of the 352 municipalities count less than 200 households, 45 consist of more than 2000. Consequently, absolute expenditure levels also differ widely. Finally, we observe substantial differences in municipal tax multiplier and therefore tax rates. Note that the average tax rate for the same household income is lower for married couples in order to counter the effect of a joint assessment of spouses in the progressive tax system.

5 Calibration

In this section, we provide a detailed description of our empirical approach. Generally, we focus on the year 2016, the last year for which complete household level data are available to us; although some of our identification strategies rely on panel and cross-sectional data-sets comprising multiple years. We start by showing how we cluster the universe of households in the Canton of Bern into groups with similar socio-economic characteristics. In the subsequent subsections, we present our identification strategies for the estimation of the different model parameters. In Subsection 5.2, we show how we exploit a natural experiment to estimate procedural utility, i.e., peoples gains from participating in the political decision-making process. In Subsection 5.3, we show how we predict the tax multipliers of counterfactual municipalities based on the spatial distribution of household types. In Subsection 5.4, we describe our estimation approach for the household-type-specific mobility cost parameters, κ^{θ}. These parameters shape the utility response of the different household types when locating away from their original residence municipality. Finally, we estimate the parameter δ that controls the extent to which larger jurisdictions profit from economies of scale when providing public services in Subsection 5.5. Table A1 summarizes the calibration exercise.

We borrow from the literature three sets of parameters. First, we follow Desmet et al. (2018), who set the shape parameter of the idiosyncratic utility shock $\epsilon = 2$, based on studies by Diamond (2016), Ortega and Peri (2016), Monte et al. (2018),

[22] See Figure 1d and Brülhart et al. (2015) for an international comparison of jurisdictional fragmentation in federally organized countries.

and Fajgelbaum et al. (2019). This value is informed by data of the European Union. Second, we suppose that the elasticity of utility to public goods ι equals 0.2 following Aronsson and Johansson-Stenman (2008). This estimate is in line with values obtained and applied in Luttmer (2005) and Fajgelbaum et al. (2019). Finally, we build on the Swiss transportation literature to calibrate household-type-specific commuting costs, as a reduction in effective residential wages (see, Becker, 1965; Axhausen et al., 2008; Börjesson and Eliasson, 2014). Precisely, we follow evidence on Swiss workers' commuting in Axhausen et al. (2008), and consider $d_{ni}^{\theta} = \tau_{ni}^{\eta_1} + (\tau_{ni} * \frac{w^{\theta}}{\overline{w}})^{\eta_2}$, with $\overline{w}$ being the average wage in the Canton. We set $\eta_1 = 0.15$ and $\eta_2 = 0.17$.

5.1 Household types (θ)

We group the households of the Canton of Bern into 250 homogeneous types pooling all years from 2012 to 2016. We define these types by k-mean clustering, a standard machine learning approach (Hartigan and Wong, 1979; Alpaydin, 2020). This approach proceeds by defining groups such that the variance within each group is minimized. Formally, the method solves the following problem:

$$\arg\min_{\theta} \sum_{\theta}^{\Theta} \sum_{i \in \theta} ||\boldsymbol{x}_i - \boldsymbol{\mu}_{\theta}||^2 \tag{10}$$

$\boldsymbol{x}_i$ is a matrix of household characteristics consisting of the variables described in Panel A of Table A6: average age of adults, the number of adults, the number of children, aggregated household wage and the geographic coordinates of the original residence municipality of the household member with the highest income.[23] $\boldsymbol{\mu}_{\theta}$ are the averages of the characteristics within type θ.

When selecting the optimal number of groups, we face a trade-off between a high accuracy of household types, that is minimal within-cluster variances, and sufficient households within each type to estimate their underlying distributions across municipalities. After experimenting with different numbers we specify 250 household types. Given the roughly 445'000 households in the Canton of Bern, this choice provides suf-

[23]Our household variables have different units of measurements. We therefore standardize the data to have a mean of zero and a standard deviation of one as an essential pre-processing step. Moreover, we use the center coordinates of municipalities. These are coordinates of centers of the municipalities, such as railway stations or main squares as specified by the Federal Office of Topography. We also experimented with centroid coordinates. However, center coordinates appear superior when calculating routing distance.

ficient observations per group – in average about 1780 per year and 8900 over the 5 year period between 2012 and 2016. Moreover, our choice ensures a high accuracy of household types as apparent from Figure A2 and the standard accuracy measures reported in Table A2.

In average, household wages deviate by about 18% from their respective group means. Moreover, types are homogeneous in terms of household members with mean absolute errors in the number of adults and children amounting to about 0.07 and 0.06, respectively. By contrast, household types tend to be more heterogeneous in terms of the average age of adults with a mean absolute error of 6.5 years. Finally, we observe mean lateral and horizontal deviations to the households' actual locations of origin of around 5km. Figure A2 provides additional support for our clustering strategy by plotting kernel density estimates of the deviations of the household variables from their respective group means. As shown by the different graphs, the method ensures a high degree of homogeneity within groups.

5.2 Benefits of self-determination as procedural utility (ρ)

In the baseline version of our model, we specify the benefits from self-determination as procedural utility gains. That is to say that individuals *value the possibility to make decisions on their own, irrespective of the actual outcome* (Frey et al., 2004; Frey and Stutzer, 2005). In this section, we present our empirical approaches to identify procedural utility gains.

The public finance and decentralization literature has also approached the benefits from self-determination as an informational advantage. That is to say that individuals have an informational advantage when making decisions for themselves. In Section B of the Online Appendix, we exploit the difference in local institutions (assembly vs parliament) to estimate the size of the informational advantage. Results are qualitatively and quantitatively similar using both definitions of self-determination benefits.

A natural experiment. To retrieve the parameter governing the benefit from self-determination (ρ), we exploit a natural experiment in the Canton of Geneva. In 2005, the Canton of Geneva gave the *right to vote locally* to all foreigners who have been residing in Switzerland for 8 years or more.[24]

Combining this natural experimental setting with data from the annual Swiss Households Panel (SHP) survey, we first study how foreigners' perception of their own political influence varies around the 8 years threshold. To obtain ρ, we subsequently multiply this effect on perceived political influence with the effect of perceived political influence on overall life satisfaction.

Estimation strategy. We propose two complementary identification strategies – as well as their combination – to estimate how political participation rights at the municipal level affect peoples perception about their overall political influence. In the Regression Discontinuity (RD) approach, we focus on foreigners just above and just below the 8 years threshold in and after 2005. In the Difference-in-Difference (DiD) approach, we compare foreigners in Geneva versus in other cantons without local voting rights (the Canton of Bern is one of them), before and after 2005. Both approaches are presented in greater detail below.

Regression Discontinuity (RD) approach. We refer to $Dist_{i,t}$ as the running variable which measures the number of years since receiving the voting rights. $Dist_{i,t}$ takes a negative value for years before receiving local voting rights. $Treat_{i,t}$ is a treatment dummy equal to one if $Dist_{i,t} \geq 8$, and 0 otherwise. τ_1 is the coefficient of interest. h is the estimated MSE-optimal bandwidth following Calonico et al. (2018). To maximize sample size, and thus estimation power, we pool together all years between 2005 and 2009 and capture time specific effects by including year dummies. Formally, we estimate the following model:

$$\begin{aligned} Y_{i,t} = \alpha + \tau_1 Treat_{i,t} + f(Dist_{i,t}) + \mu_t + \varepsilon \\ s.t. - h \leq Dist_{i,t} \leq h. \end{aligned} \tag{11}$$

where $f(Dist_{i,t})$ are different polynomial orders of $Dist_{i,t}$ and $Treat_{i,t} \times Dist_{i,t}$ (see Lee and Lemieux, 2010; He et al., 2020). The AIC is used for the polynomial order choice. μ_t are year fixed effects.

Difference-in-Difference (DiD). Consider the panel of all foreigners residing in Switzer-

[24] In the 2000s, several Swiss cantons allowed foreigners to vote locally (Neuchatel, Jura, Vaud, and Fribourg). However, we focus on Geneva, as it is the only Canton which based the vote capacity criterion solely on length of the residency in Switzerland (which we observe).

land for 8 or more years.[25] Define the indicator variable $Geneva_i$ which equals one if an individual is living in Geneva, and zero otherwise. $Post_t$ amounts to one for $t \geq 2005$. We estimate the following model, with τ_2 as the coefficient of interest:

$$Y_{it} = \alpha + \tau_2(Post_t \times Geneva_i) + \boldsymbol{X'\beta} + \mu_t + \eta_i + \varepsilon \tag{12}$$

where $\boldsymbol{X'\beta}$ are a set of individual-level covariates: age, gender, educational level, years of education, civil status, and region of origin. μ_t and η_i are year and individual fixed effects, respectively.

Furthermore, we also combine both estimation strategies and estimate the triple difference model which we report under the label DiDiD. By combining the two identifying variations presented above, we view this approach as a robustness check.

Validity of the approaches. A key asset of Regression Discontinuity designs is that identifying assumptions are empirically testable. First, Table A3 provides balancing tests for all observed individual level covariates around the 8 residency year threshold. These include demographic covariates (i.e., age, gender, civil status), education level covariates (i.e., type and length of education), as well as origin covariates based on the nationality of foreigners. Overall, individuals just above and just below are not statistically different in any of the observed dimensions.

Second, Figure A3 assesses the smooth variation of the running variable – years of residence in Switzerland – around the 8 years threshold following McCrary (2008). A smooth variation is indeed observed.

Validity of the Difference-in-Difference model (12) requires parallel trends in the control and treatment groups prior to the year of the treatment. Figure A4 shows that the reported perceived political influence followed similar trends between the treatment and control groups prior to the reform.

Results. Table 1 presents the estimation results. It is organized in three pairs of columns, within each the first reports an OLS specification whereas the second adopts a Poisson specification. The first pair reports the results of the RD design (11) for τ_1. The second pair reports the results for the DiD (12) for τ_2. Finally, the third pair

[25]We further restrict the panel to all foreigners residing in Switzerland for not more than 30 years (more than twice the time needed to ask for citizenship, i.e., 12 years). Estimating the interest in political participation on such individuals is likely to be impacted by sorting on unobservables.

reports the results of the triple interaction model (DiDiD). Overall, we obtain similar effects of a magnitude between 0.997 and 2.248. We interpret the similarity in the obtained values under the different approaches as a signal that our result is robust to the type of identification strategy employed. Figure A4 provides graphical evidence of the effect estimated under both the RD and the DiD approaches. Figure A4a shows a RD plot obtained following Calonico et al. (2014), whereas Figure A4b plots the year-specific treatment effect when interacting the DiD treatment dummy with year indicators.

Table 1: Estimation of the procedural utility parameter

Dependent variable: Perceived political influence (0: low; 10: high)						
Identification strategy	RDD		DiD		DiDiD	
Estimation model	Linear	Poisson	Linear	Poisson	Linear	Poisson
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	1.802**	2.248*	-	-	-	-
	(0.826)	(0.850)				
Geneva × Post	-	-	1.062**	1.919***	-	-
			(0.467)	(0.469)		
Treat × Geneva × Post	-	-	-	-	0.997**	1.549**
					(0.449)	(0.286)
Obs.	87	87	1,870	1,870	1,387	1,387
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes	No	No
Covariates	No	No	Yes	Yes	Yes	Yes
Bandwidth	Opt.	Opt.	-	-	Opt.	Opt.

Notes: This table presents the estimation of (11) in Columns (1) and (2), using respectively a linear and a Poisson specification. The optimal bandwidth size is 7 years. Using the same order, Columns (3) and (4) report the estimation of (12). Columns (5) and (6) report a triple interaction model combining the two sources of variation in (11) and (12). Incidence-rate ratios are reported for the Poisson regressions. Robust standard errors are in parentheses.

From perceived political influence to overall utility. To bridge the gap between the obtained effect on perceived political influence provided in 1 and the effect on overall utility as modelled in (1), we correlate the level of perceived political influence and overall satisfaction with life. To do so, we use individual responses on both dimensions from the annual Swiss Households Panel (SHP) survey. Figure A5 illustrates the significant and plausible correlation. The resulting coefficient of 0.055 indicates that a one unit increase in perceived political influence (on a scale from 0 to 10) is associated with a 0.055 increase in overall life satisfaction. Using this elasticity with the estimates from Table 1, we obtain $\rho = 2.248 \times (1/11) \times 0.055 = 0.011$.

5.3 Preferences for tax multiplier (m_n)

(4) states that municipal tax multiplier, and hence local income tax rates, depend on the distribution of household types in a given municipality:

$$m_n = f\left(\frac{H_n^1}{H_n}, \frac{H_n^2}{H_n}, ..., \frac{H_n^\Theta}{H_n}\right).$$

Our goal in this subsection is to predict the tax multiplier of counterfactual municipalities based on the distribution of household types. We could restrict ourselves to traditional econometric methods. However, a rapidly increasing literature shows that the predictive performance of traditional models is poor relative to common methods of the machine learning literature (for instance, see Athey and Imbens, 2019; Athey et al., 2019; Varian, 2014). This is particularly true in the case at hand with a large number of 250 explanatory variables which may interact in complex ways that are hard to specify in advance. As stated by Mullainathan and Spiess (2017): "the appeal of machine learning is that it manages to uncover generalizable patterns". The methods manage to fit complex functional forms to the data that work well out-of-sample. Certain machine learning methods have a strong tendency to over-fit and thus overstate in-sample performance. Therefore, we first fit different machine learning methods and an ordinary least squares regression to a training sample containing the pooled municipality observations from 2012 to 2015. Subsequently, we predict the tax multiplier of 2016 based on the shares of household types in each municipality.

Table A8 provides the out-of-sample performance accuracy measures for the different methods. As apparent from the table, the machine learning algorithms largely outperform the ordinary least squares regression. This is especially true for the Gaussian Process Regression and, to a lesser extend, the Support Vector Machine. The fifth row in Table A8 shows the performance of the Gaussian Process Regression: the mean absolute deviation of the out-of-sample prediction amounts to 3.75%, which is around 40% below the measure for the ordinary least squares regression.[26]

An important feature in the machine learning literature to improve out-of-sample

[26] Note that averaged over the 250 household types, the standard deviation *within* municipalities (0.0024) amounts to 24% of the standard deviation *across* municipalities (0.0102). The out-of-sample predictions are thus based on both the within and between variation in our predictor variables.

Table 2: Tax multiplier: accuracy of out-of-sample prediction

	(1) **MAPE**	(2) **RMSPE**	(3) **p10**	(4) **p90**	(5) **Weight in Ensemble**
Model					
Mean	9.77%	13.59%	1.15%	19.08%	0.00
Linear Regression	6.59%	8.68%	1.02%	14.69%	0.00
Random Forest	5.61%	8.06%	0.68%	10.99%	0.00
Support Vector Machine	4.96%	7.01%	0.72%	11.19%	0.10
Gaussian Process Regression	3.84%	5.86%	0.46%	8.48%	0.90
Ensemble	3.83%	5.80%	0.35%	8.61%	

Notes: The training period is: 2012 to 2015. Each year contains the entire set of 352 municipalities. All goodness-of-fit measures refer to the out-of-sample year: 2016. For the Support Vector Machine Regression and the Gaussian Process Regression we select the Kernel Function with the lowest in-sample RMSPE (exponential for the Gaussian Process Regression and cubic for the Support Vector Machine Regression). The Random forest is averaged over 1000 trees; p10 and p90 are calculated on the absolute percentage error.

predictions of single models is the use of ensemble methods or model averaging. A combination of our models in Table A8 thus may perform better than our best single method. The key question concerning the use of these methods is how to weight the different models. We follow the guidelines from Athey and Imbens (2019) and chose weights by minimizing the sum of squared residuals in the test sample, imposing the restriction that the weights, p^{lr}, p^{rf}, p^{sv} and p^{gp} are non-negative and sum to one:

$$(\hat{p}^{lr}, \hat{p}^{rf}, \hat{p}^{sv}, \hat{p}^{gp}) = \min_{p^{lr}, p^{rf}, p^{sv}, p^{gp}} \sum_{n=1}^{N^{test}} (m_n - p^{lr}\hat{m}_n^{lr} - p^{rf}\hat{m}_n^{rf} - p^{sv}\hat{m}_n^{sv} - p^{gp}\hat{m}_n^{gp})^2, \quad (13)$$

subject to $p^{lr} + p^{rf} + p^{sv} + p^{gp} = 1$ and $p^{lr}, p^{rf}, p^{sv}, p^{gp} \geq 0$.

Column (5) in Table A8 shows the estimated weights. Our final model consists of a weighted average of the prediction from the Gaussian Process Regression ($p^{gp} = 0.9$) and the Support Vector Machine ($p^{sv} = 0.1$). As apparent from the table, this combination of methods slightly outperforms the Gaussian Process Regression, although only by a narrow margin.

Finally, we apply the same procedure as described in this subsection to predict the housing tax rate of counterfactual municipalities. The corresponding out-of-sample performance accuracy measures are provided in Table A8. As apparent from the Table, the Gaussian Process Regression again provides the most accurate out-of-sample predictions.

5.4 Preferences for location of origin (κ^θ)

We estimate the heterogeneous preferences of different household types for their municipality of origin following McFadden (1974) and Eaton and Kortum (2002).[27]

The probability that a household of type θ locates in n is defined in (6) as:

$$\pi_n^\theta = \frac{\sum_i^N v_{ni}^\theta}{\sum_n^N \sum_i^N v_{ni}^\theta}.$$

We model the dis-utility incurred by a household of type θ by locating away from its origin municipality through an iceberg mobility cost $d_n^\theta = e^{\psi^\theta \bar{d}_n^\theta}$. It increases with travel time between the origin location of a household of type θ and municipality n.[28] Taking logs, a key prediction from (6) is a semi-log gravity equation[29] describing the decisions of households of type θ to locate in municipality n.

$$\ln \pi_n^\theta = \kappa^\theta \bar{d}_n^\theta + \nu_n + \mu_\theta + v_n^\theta. \tag{14}$$

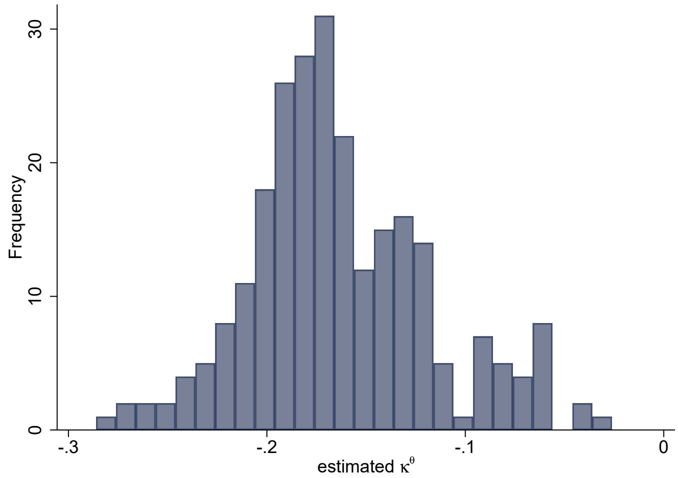
We account for the endogenous construction of the road network by instrumenting travel time with the euclidean distance between the two places. $\bar{d}_n^\theta$ thus denotes the residual travel time between the original residence location of a household of type θ and the municipality of residence n. The parameters $\kappa^\theta = -\psi^\theta \epsilon$ denote the household-type specific semi-elasticities of the moving flows with respect to the residual travel times. They are combinations of the type-specific mobility cost parameters ψ^θ and the location choice heterogeneity parameter ϵ. $\pi_n^\theta = H_n^\theta / H$ denote the households of type θ living in municipality n as a share of all households in the Canton of Bern. The municipality fixed effects ν_n capture municipality specific characteristics $\{m_n, g_n, b_n, O_n, Q_n, \sum_{i=1}^N d_{ni}\}$. Wages as well as the denominator in (6) are absorbed through the household-type fixed effects μ_θ. v_n^θ denotes the disturbance term.

[27] See for instance Ahlfeldt et al. (2015) and Donaldson and Hornbeck (2016) for more recent related applications.

[28] We use the georoute command from Weber and Péclat (2017) Travel time is in minutes and computed based on the road network.

[29] See in particular Head and Mayer (2014) for comprehensive guidelines on the application of gravity-equations.

Figure 2: Household-type specific semi-elasticities of residential mobility



Notes: The histogram shows the distribution of the estimated semi-elasticities (κ^{θ}) of the moving flows with respect to the residual travel times from (6). Minimum: -0.2861, 10^{th} percentile: -0.2131, 90^{th} percentile: -0.0901, maximum: -0.0284, mean -0.1632. All coefficients are statistically significant at the one percent level; Width of bins: 0.01; Number of groups: 250.

We estimate (14) using a Poisson Pseudo Maximum Likelihood estimator and provide the distribution of the found semi-elasticities in Figure 2. The mean of the 250 estimates amounts to -0.163 and all coefficients are statistically significant at the one percent level. We find large heterogeneity in the semi-elasticities across household types, ranging from -0.286 to -0.028. Therefore, depending on the household type, each additional minute away from a types' original residence location decreases the number of households by 3 to 29 percent.

5.5 Economies of scale (χ)

In this subsection, we aim to estimate the parameter χ, which captures economies of scale in providing public services.[30] A parameter value of zero implies infinite economies of scale or the provision of perfectly non-rival public services. By contrast, a parameter value of one implies perfectly rival public services or a complete absence of economies of scale.

Rewriting the equation on the public budget clearing (5) and taking logs provides

[30] See Fajgelbaum et al. (2019) for a recent paper that models the degree to which public goods are rival similarly, albeit without actually estimating a model parameter. For an empirical test on the existence of economies of scale in municipal government expenditures see Holcombe and Williams (2009).

$$\log(Y_n) = \mu + \chi \log(H_n) + \epsilon_n. \tag{15}$$

where Y_n denotes the part of public spending in municipality n that is not directly affected by geographical characteristics, as explained in more detail below. μ denotes the constant, which absorbs the component of the municipal budget that is determined through the cantonal equalization scheme.

Table 3: Estimation for economies of scale in the provision of municipal services

	(1) IV-Second Stage	(2) IV-First Stage	(3) OLS
Log number of households	0.989***	-	0.967***
	(0.018)		(0.017)
Log area	-	1.106***	
		(0.070)	
Log altitude	-	-2.191***	
		(0.166)	
Observations	352	352	352
R-squared	0.953	0.477	0.953
F-statistic	-	127.81	2971.29

Notes: This table presents the OLS and 2SLS estimates of the log number of households on log municipal expenditure for public services without direct geographical dependence for the year 2016. Dependent variable in (1) and (3) is log municipal expenditure public services without direct geographical dependence. Dependent variable in (2) is the log number of households. Robust standard errors are in parentheses.

We refrain from estimating χ by OLS as our baseline due to potential simultaneous causality of public expenditure and the number of households. In addition (15) may suffers from an omitted variable bias since the number of households is likely affected by the level of effective public expenditure per household, g_n, which is part of the error term. We address these endogeneity issues as follows. First, we instrument H_n with municipal area and altitude using a two stage least squares regression. Second, we exclude expenditure categories that are potentially directly affected by our instruments, such as transport services, the maintenance of road networks or water supply, from our dependent variable.[31] In addition, we provide a robustness check in which we only use municipal expenditure on purely administrative matters as our

[31] We exclude the following public services from our dependent variable, since they may be directly affected by our instruments: public transport services, maintenance of road networks, water supply, sewage disposal, waste disposal, promotion of culture, social aid and expenses for environmental protection.

outcome variable in Section A.

The first two columns of Table A5 present the results from the two stage least squares regressions. As apparent from the first stage regression, the two instruments are highly significant. A larger municipal area naturally results in a higher population while more rural and mountain regions tend to have a lower population density. Column (1) reports the estimate of our second-stage regression. We find an elasticity of absolute public expenditure with respect to the number of households of $\chi = 0.989$. Therefore, we find relatively modest economies of scale in the provision of public goods. Finally, we run a simple OLS regression as a sensitivity check, which we report in Column (3).

5.6 Municipal fundamentals (B_n, ϕ_n)

The framework features two types of municipal fundamentals: fundamental residential amenities (B_n) and density of development (ϕ_n).

Proposition 2. *Given known values for the parameters and the observed data, there exist unique vectors of residential amenities $\{\boldsymbol{B}\}$, and density of development $\{\boldsymbol{\phi}\}$ that close the model.*

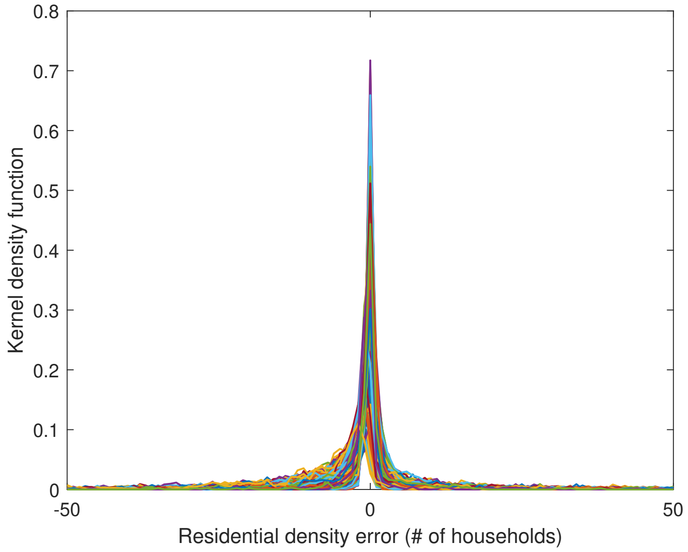
PROOF: See the proof of Proposition 2 in the Online Appendix E.2. *Q.E.D.*

5.7 Model validation

The distribution of the 250 household types across the 352 municipalities is a key endogenous moment of the framework, and one from which the other endogenous variables are derived (see Online Appendix C). However, whereas the overall residential density is directly constrained by available land space, L_n, this is not the case of household-type-specific densities.

Hence, to assess the validity of the model, Figure 3 proposes over-identification checks in the form of household-type-specific distributions of residential density errors. Formally, denoting the baseline predicted number of household of type θ in municipality n by $\hat{H}_n^\theta$, Figure 3 plots the difference $H_n^\theta - \hat{H}_n^\theta, \forall n, \theta$. On the horizontal axis, the error is measured in the number of households. For each of the 250 household types, we plot a kernel density function with a smoothing bandwidth of 0.5. Overall, across all 250 household types, the type-specific errors are centered around zero, with

Figure 3: Type-specific distributions of residential density errors



Notes: The histogram shows the distribution of the error in the residential density at the municipal level for all household types. i.e., $H_n^\theta - \hat{H}_n^\theta, \forall n, \theta$. The number of municipalities is 352, and the number of household types is 250. For smoothing, the selected bandwidth size is 0.5.

less than 5% of all errors larger than 0.005% of the average household type size. This confirms that the framework performs well at predicting the location of household types.

6 Counterfactual analysis

We simulate two types of reforms limiting municipalities in their fiscal autonomy. First, on the revenue side, we simulate different levels of tax harmonization across jurisdictions (Section 6.1). Second, on the spending side, we suppose that different spending mandates are imposed to local jurisdictions (Section 6.2). In each case, we study both the aggregate and distributional effects of limitations to fiscal autonomy. In what follows, we present and discuss the two simulation sets successively.

Outcomes of interest

We focus on the endogenous variables of the model measured from the household-group perspective. We measure total welfare as the sum of household-group specific welfare levels, $\mathbb{E}[u]^\theta$, hence: $\mathbb{E}[u] = \sum_{\theta \in \Theta} \frac{H^\theta}{H} \mathbb{E}[u]^\theta$. Overall and household-group specific housing prices, public good provision, and local tax multiplier are, respectively, aggregated as follows: $x = \sum_{\theta \in \Theta} \sum_{n \in N} \frac{H_n^\theta}{H} x_n, \ \forall x \in \{Q, g, m\}$, and $x^\theta = \sum_{n \in N} \frac{H_n^\theta}{H^\theta} x_n, \ \forall x \in \{Q, g, m\}$

6.1 Tax harmonization

Definition

In the first set of simulation scenarios, we suppose that local tax multipliers are harmonized by imposing symmetric lower and upper bounds on equilibrium tax multipliers. We set bounds in percentage points relative to the 2016 observed average tax multiplier (i.e., 1.719 with a standard deviation of 0.202). The resulting bandwidth size, γ^{TH}, within which tax multipliers may lie is defined on the interval $\gamma^{TH} \in [0, 0.5]$ with 0.01 increments. This leads to 51 different scenarios.

Aggregate effects

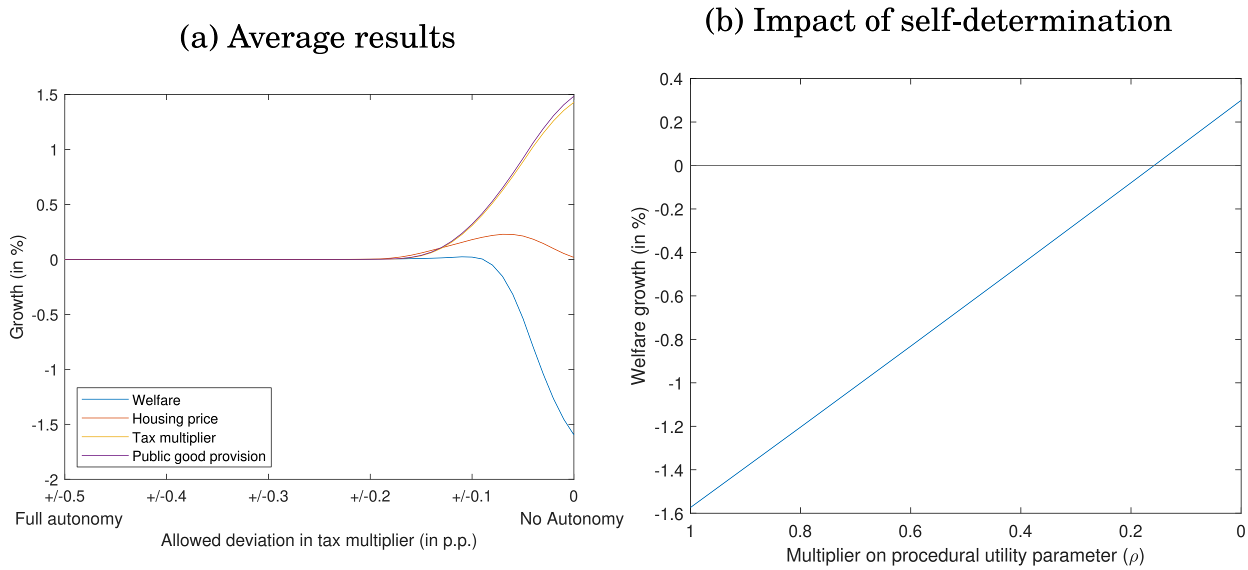
Figure 4a reports the aggregate effects of harmonizing local tax multipliers. The horizontal axis shows the absolute value of the different bounds imposed on local tax multipliers. The vertical axis displays overall growth in percent.

Overall welfare growth decreases as the allowed space for the tax multiplier is constrained. Average welfare starts to drop substantially as the allowed bandwidth decreases to 0.1 percentage point. Full tax harmonization is then associated with a welfare decline of 1.55%. This loss in average welfare is primarily driven by the negative effects through the increase in the average tax multiplier (by 1.5%), which outweighs the welfare gains from a similar increase in public good provision. The obtained effects on housing prices are quantitatively less important.

Figure 4b shows how the parameter governing the utility from self-determination (ρ) impacts the average welfare effect. The welfare effect of fiscal harmonization only turns positive at a 80% reduction in ρ. Beyond that point, the negative externalities of decentralization – e.g., spatial misallocation (Fajgelbaum et al., 2019), or "the poor chases the rich"-type inefficiencies (Hamilton, 1975, 1976; Wheaton, 1993; Fernandez and Rogerson, 1996; Hoyt and Lee, 2003; Calabrese et al., 2012) – dominate.

Overall, we interpret the gains from fiscal autonomy as arising from (i) a better sorting of household groups across municipalities, which permits a better match of tax-benefits to household preferences; and (ii) benefits from self-determination.

Figure 4: AGGREGATE EFFECTS OF TAX HARMONIZATION



Notes: Aggregate results from tax harmonization. Symmetric bounds (from below and above) are imposed on the local tax multiplier around the 2016 observed average tax multiplier (i.e., 1.719). The resulting bandwidth size, γ^{TH}, within which tax multipliers may lie is defined on the interval $\gamma^{TH} \in [0, 0.5]$ with 0.01 increments. Growth is expressed in percent relative to full tax autonomy. Panel (a) focuses on the average effect (across household types and municipalities) on welfare, tax multipliers, local spending and housing prices. Panel (b) studies how the welfare effect from moving to full harmonization is impacted by a reduction in the procedural utility parameter (ρ).

Distributional effects

Figure 5 shows the distributional effects of harmonizing local taxes by way of contour plots.[32] Panel (a) focuses on heterogeneity in households' income, and Panel (b) in households' number of children. To highlight differential impacts of tax harmonization, we normalize growth such that it equals zero for full tax autonomy.

The distributional effects of harmonizing local tax multipliers confirm the aggregate effects. Panel (a) reveals that all income brackets benefit from more freedom in setting local tax multipliers. However, whereas the gains are minimal for low income brackets, welfare growth is substantial for higher income brackets (up to 2%). The picture is slightly more complicated when looking at the households' number of children. Households without children or with just 1 child gain substantially from less

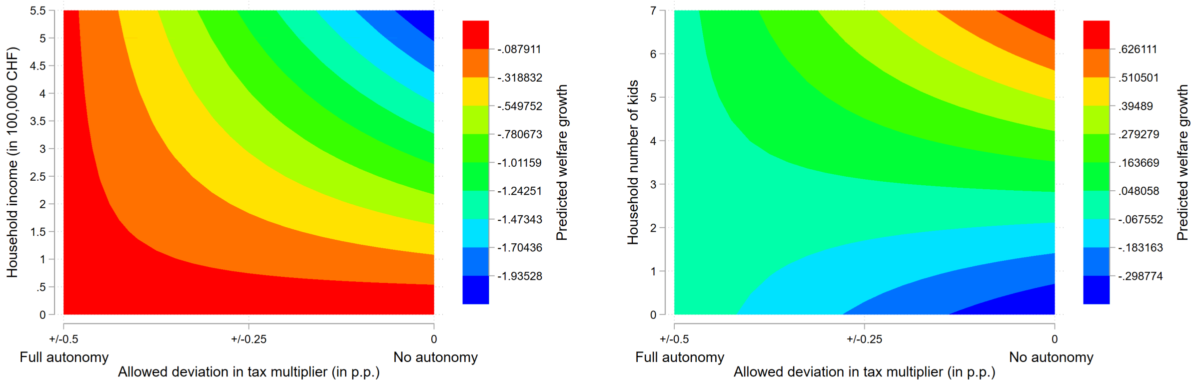
[32] Denoting a given simulation by the subscript $_s$, contour plots in Figure 5 are formally derived by estimating the following model:

$$y_{\theta s}^{TH} = \beta_1^{TH}\gamma_s^{TH} + \beta_2^{TH}(\gamma_s^{TH} \times income_\theta) + \beta_4^{TH}(\gamma_s^{TH} \times kids_\theta) + \beta_\theta^{TH} + \boldsymbol{X'\beta} + \epsilon_{\theta s}^{TH}. \quad (16)$$

$y_{\theta s}^{TH}$ is the θ-specific outcome of interest in scenario s (e.g., $\mathbb{E}[u]^\theta$ under s). β_θ^{TH} controls from household-group specific effects. Figure 5 then reports the estimated interaction effects. Importantly, note that all interaction effects are jointly estimated.

Figure 5: Distributional effects of tax harmonization

(a) Household income (b) Household number of children



Notes: Distributional effects of harmonizing local taxes by way of contour plots following (16). Panel (a) focuses on heterogeneity in households' income, and Panel (b) in households' number of children. Growth is normalized such that is equal to zero for an allowed deviation in tax multiplier equal to 0.

tax harmonization (with a welfare growth of 0.5% for households without children). Yet, the effect is reversed for households with more than three children. Households with seven children lose up to 0.6% in welfare when local tax multipliers are not harmonized. In general, these findings are in line with the optimal tax and sorting literature: more mobile households (i.e., richer, with fewer children) have more to gain from increased freedom in setting local tax multipliers. However, the gains from increased freedom in setting tax multipliers are widespread and concern the vast majority of households.

6.2 Mandatory spending

Definition

In the second set of simulation scenarios, we suppose that mandates are imposed on the level of public spending. Formally, we set different minimum (per capita) spending levels around the observed average 2016 municipal spending level (i.e., CHF 4,739 per capita). We then define the different spending minimum, γ^M, on the interval $\gamma^M \in [4000, 5000]$ with increments of 50. This leads to 21 different scenarios.

Aggregate effects

Figure 6a reports the aggregate effects of imposing mandatory spending levels to municipalities. The horizontal axis shows the imposed minimum spending levels. The vertical axis displays overall growth in percent.

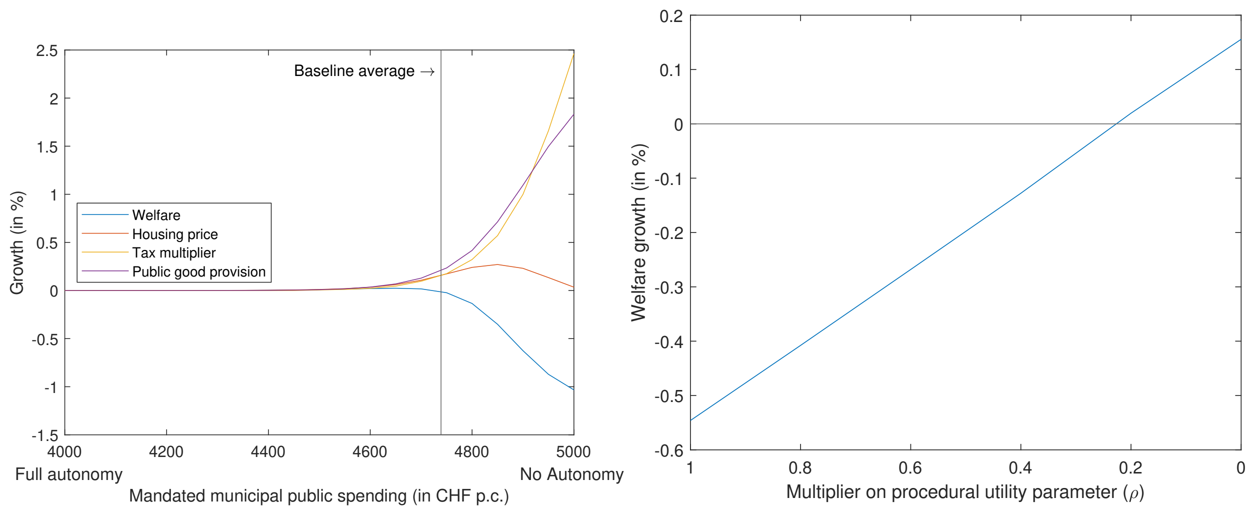
Below CHF 4,600 per capita, we observe no average effect in the endogenous variables of interest. Beyond that level, overall welfare growth decreases with the size of the mandatory spending level to reach a 1% decline for a minimum mandatory spending level of CHF5,000. Average per capita public spending mechanically increases with mandates. To finance such increase, local tax multipliers also increase by up to 2.5% to finance an average 1.8% increase in spending. The effect on housing prices is ambiguous, but remains close to 0%.

As for tax harmonization, Figure 6b reveals that the decline in welfare following mandatory spending (here of CHF 5,000 per capita) resists up to a 80% decline in the parameter governing the utility from self-determination, ρ.

Imposed minimum spending levels generate the overall decline in welfare through two channels: first, imposed mandates reduces utility that local constituents gain from self-determination. Second, higher mandatory spending implies higher local tax rates and reduces sorting in the area. In turn, group-specific preferences for public spending levels are less well matched in equilibrium.

Figure 6: Aggregate effects of mandatory spending

(a) Average results (b) Impact of self-determination



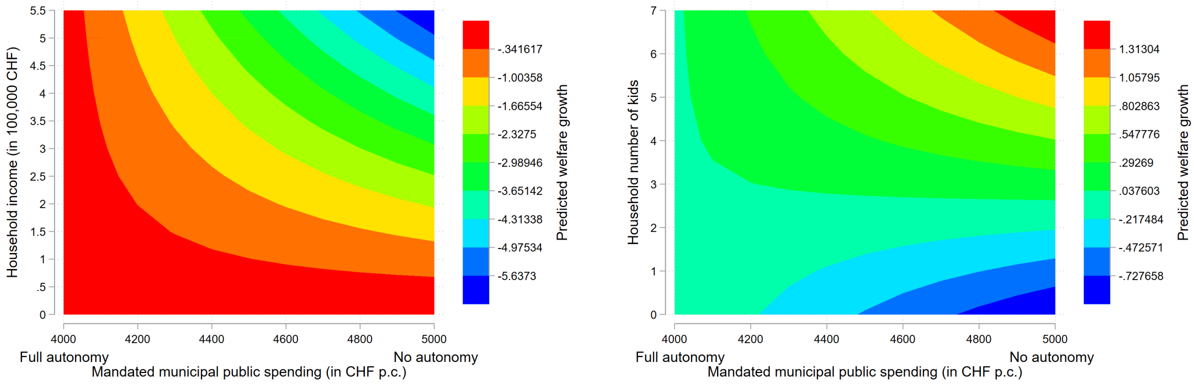
Notes: Consequences of different minimum mandatory spending (per capita) are plotted. Minimum are chosen around the observed average 2016 municipal spending (i.e., CHF 4,739 per capita). The different spending minimum, γ^M, are defined on the interval $\gamma^M \in [4000, 5000]$ with increments of 50. Growth is expressed in percent relative to a minimum mandatory spending of CHF 4,000. Panel (a) focuses on the average effect (across household types and municipalities) on welfare, tax multipliers, local spending and housing prices. Panel (b) studies how the welfare effect from imposing a high mandatory spending (CHF 5,000 per capita) is impacted by a reduction in the procedural utility parameter (ρ).

Distributional effects

Figure 7 shows the distributional effects of imposing mandatory spending levels by way of contour plots.[33] Panel (a) focuses on heterogeneity in households' income, and Panel (b) in households' number of children. To highlight differential impacts of imposing mandatory spending levels, we normalize growth such that it equals zero for a municipal spending level of CHF 4,000 per capita.

Figure 7: Distributional effects of mandatory spending

(a) Household income (b) Household number of children



Notes: Distributional effects of imposing mandatory spending by way of contour plots following (17). Panel (a) focuses on heterogeneity in households' income, Panel (b) in households' average age, and Panel (c) in households' number of children. Growth is normalized such that is equal to zero for a minimum spending of CHF 4,000 per capita.

The results are qualitatively similar to the distributional effects following from tax harmonization and presented in Figure 5. The main difference is that gains and losses under mandatory spending are larger in magnitude (e.g., maximum gain from no spending mandate of 6% in overall welfare for the richest households). In line with the optimal tax and sorting literature: more mobile households (i.e., richer, with less children) have more to gain from increased freedom in choosing the level of public spending. However, the gains from increased freedom in setting public spending levels are widespread and concern most households in Bern. Conditional on all other

[33] Denoting a given simulation by the subscript $_s$, contour plots in Figure 7 are formally derived by estimating the following model:

$$y^M_{\theta s} = \beta^M_1 \gamma^M_s + \beta^M_2 (\gamma^M_s \times income_\theta) + \beta^M_4 (\gamma^M_s \times kids_\theta) + \beta^M_\theta + \boldsymbol{X'\beta} + \epsilon^M_{\theta s}. \quad (17)$$

$y^M_{\theta s}$ is the θ-specific outcome of interest in scenario s (e.g., $\mathbb{E}[u]^\theta$ under s). β^M_θ controls for household-group specific effects. Figure 7 reports the estimated interaction effects. Importantly, note that all interaction effects are jointly estimated.

group characteristics, the only household groups benefiting from higher public spending levels are households with 3 or more children.

6.3 Robustness to parameter governing the elasticity of utility to public goods (ι)

We test the robustness of the results above to perturbations in the parameter governing the elasticity of utility to public goods (ι). This central parameter was not estimated in a Swiss context. As such, it appears important to ensure that small perturbations in its value do not impact the overall qualitative results.

Figure A6 presents the results from this robustness exercise. Panel a focuses on the results from tax harmonization, whereas Panel b focuses on the results from mandatory spending. In each case, we decrease/increase the parameter of interest by $\pm 80\%$ (with 40 p.p. intervals). As summarizing moment, we focus on aggregate welfare growth. Overall, whereas the quantitative results are naturally impacted, the welfare effect of reducing fiscal autonomy (either via tax harmonization or mandatory spending levels) remains largely negative.

6.4 Policy implications: Welfare versus fairness?

The obtained results presented in Figure 5 and Figure 7 reveal that for the vast majority of household types – except the less than 1% of households with more than 3 children – restricting fiscal autonomy implies a decline in welfare. Hence, allowing for more fiscal autonomy is (almost) Pareto optimal.

Yet, the welfare growth effect of unrestricted fiscal autonomy is highly heterogeneous and leads to rising welfare inequalities from above. This fact has received increased media attention, and some have advocated to give more weight to relative inequalities when designing local institutions. As such, it appears informative to study the relative marginal evolution of welfare and relative inequalities for different levels of fiscal autonomy. To measure inequalities, we compute the maximum spread in welfare for each measure and level of fiscal autonomy.

We display the results in Figure A7, where Panels a and b study different levels of tax harmonization and mandatory spending, respectively. In line with results in Figure 4, marginal welfare and spread growth decreases with tax harmonization; especially for tight bounds below 0.1 percentage points. However, the net effect (in red) exhibits an interesting pattern: When reducing fiscal autonomy with tax harmonization, we observe that for bounds of 0.16 p.p. (or lower) the decline in welfare is less important than the decline in the spread. Hence, policy makers favoring a relative social welfare metric may still favor some tax harmonization. For instance, with bounds between 0.1 and 0.16 p.p., the negative welfare effects remain small and we observe a decline in the spread.

Marginal welfare and inequality growth in the case of varying mandatory spending levels offer a more united picture (Figure A7b). Imposing mandatory spending levels leads to negative marginal welfare growth beyond CHF4,500 per capita (and null below), in line with average results in Figure 6. The welfare spread is also declining beyond that mark. However, the decline in marginal welfare growth always exceeds the decline in the spread. Marginal benefits in terms of spread reduction are therefore always outweighted by a larger decline in welfare.

7 Conclusion

This paper quantifies the equilibrium effects of limiting local fiscal autonomy accounting for the benefits of self-determination. We combine data on the universe of households, individual survey data and precise municipal information for the Canton of Bern to inform a spatial equilibrium framework of local public good provision. The framework features imperfectly mobile households – defined by joint income, location of origin, average age and number of children – who decide on their residential location given endogenous local tax multiplier, public goods, and housing prices. It accounts for the benefits from self-determination as the utility derived from citizens' participation in the political decision-making process. Quasi-natural variation implied by a voting rights reform is used to inform the calibration of these benefits in both a regression discontinuity design and a difference-in-difference approach.

After calibrating the framework to fit the Canton of Bern's economic and geographic characteristics and assessing its validity with over-identification checks, we simulate two sets of reforms affecting local fiscal autonomy: (i) harmonization of the local tax multipliers via various upper and lower bounds, (ii) spending mandates by imposing various minimum thresholds to public spending per capita. We find that the benefits of self-determination largely outweigh the negative externalities induced by decentralization. Limiting fiscal autonomy appears to decrease aggregate welfare for almost all household groups. Yet, in line with the optimal tax and sorting literature, more mobile households (richer and older with fewer children) benefit relatively more from local fiscal autonomy.

The optimal allocation of competences between government levels and the value of autonomous decision-making has long received strong interest among scholars, policy makers and media outlets.[34] With the revolution in communication technology and rapidly changing requirements on public services provision, the topic will likely remain at the top of the policy agenda of democracies. By building a bridge between spatial equilibrium frameworks and the literature on self-determination, this paper offers an original contribution to the ongoing discussion.

References

Gabriel M Ahlfeldt, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf. The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6):2127–2189, 2015.

David Albouy. The unequal geographic burden of federal taxation. *Journal of Political Economy*, 117(4):635–667, 2009.

Alberto Alesina and Eliana La Ferrara. Participation in heterogeneous communities. *The Quarterly Journal of Economics*, 115(3):847–904, 2000.

Alberto Alesina and Enrico Spolaore. On the number and size of nations. *The Quarterly Journal of Economics*, 112(4):1027–1056, 1997.

Ethem Alpaydin. *Introduction to machine learning*. MIT press, 2020.

Thomas Aronsson and Olof Johansson-Stenman. When the joneses' consumption hurts: Optimal public good provision and nonlinear income taxation. *Journal of Public Economics*, 92(5):986–997, 2008.

Susan Athey and Guido W Imbens. Machine learning methods that economists should know about. *Annual Review of Economics*, 11:685–725, 2019.

Susan Athey, Mohsen Bayati, Guido Imbens, and Zhaonan Qu. Ensemble methods for causal effects in panel data settings. In *AEA Papers and Proceedings*, volume 109, pages 65–70, 2019.

Kay W. Axhausen, Stephane Hess, Arnd König, Georg Abay, John J. Bates, and Michel Bierlaire. Income and distance elasticities of values of travel time savings: New swiss results. *Transport Policy*, 15(3):173–185, 2008.

H. Spencer Banzhaf and Randall P. Walsh. Do people vote with their feet? an empirical test of tiebout. *American Economic Review*, 98(3):843–63, June 2008.

Iwan Barankay and Ben Lockwood. Decentralization and the productive efficiency of government: Evidence from swiss cantons. *Journal of Public Economics*, 91(5):1197–1218, 2007.

[34] See e.g. the following articles as examples for the ongoing discussion in national and regional media outlets: Wo die freisten Gemeinden der Schweiz liegen (Der Bund, 2012); Gemeinden verlieren ein Stück Autonomy (Berner Zeitung, 2017); So wollen Gemeinden das Milizsystem retten (Tagesanzeiger, 2019); Steuerwettbewerb oder Steuerharmonisierung?; Lusthemmender Finanzausgleich (NZZ, 2001, 2015).

Christoph Basten, Maximilian von Ehrlich, and Andrea Lassmann. Income taxes, sorting and the costs of housing: Evidence from municipal boundaries in switzerland. *The economic journal*, 127(601):653–687, 2017.

Gary S. Becker. A theory of the allocation of time. *The Economic Journal*, 75(299):493–517, 1965.

Timothy Besley and Anne Case. Political institutions and policy choices: evidence from the united states. *Journal of Economic Literature*, 41(1):7–73, 2003.

Duncan Black. On the rationale of group decision-making. *Journal of Political Economy*, 56 (1):23–34, 1948.

Maria Börjesson and Jonas Eliasson. Experiences from the swedish value of time study. *Transportation Research Part A: Policy and Practice*, 59:144–158, 2014.

L.E.J. Brouwer. Uber abbildung von mannigfaltigkeiten. *Math. Ann.*, (71):97–115, 1911.

Jan Brueckner. A test for allocative efficiency in the local public sector. *Journal of Public Economics*, 19(3):311–331, 1982.

Marius Brülhart and Mario Jametti. Does tax competition tame the leviathan? *Journal of Public Economics*, 177:104037, 2019.

Marius Brülhart, Sam Bucovetsky, and Kurt Schmidheiny. Taxes in cities. In *Handbook of regional and urban economics*, volume 5, pages 1123–1196. Elsevier, 2015.

Marius Brülhart, Jayson Danton, Raphaël Parchet, and Jörg Schläpfer. Who bears the burdean of local taxes? Technical report, CEPR, 2021.

Allen Buchanan. Self-determination and the right to secede. *Journal of International Affairs*, 45(2):347–365, 1992.

Stephen M. Calabrese, Dennis N. Epple, and Richard E. Romano. Inefficiencies from metropolitan political and fiscal decentralization: Failures of tiebout competition. *Review of Economic Studies*, 79(3):1081–1111, 2012.

Sebastian Calonico, Matias D. Cattaneo, and Rocio Titiunik. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326, 2014.

Sebastian Calonico, Matias D Cattaneo, and Max H Farrell. On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association*, 113(522):767–779, 2018.

Robert A Dahl. *On democracy*. Yale university press, 2020.

Hamid Davoodi and Heng-Fu Zou. Fiscal decentralization and economic growth: A cross-country study. *Journal of Urban Economics*, 43(2):244–257, 1998. ISSN 0094-1190.

Klaus Desmet, David Nagy, and Esteban Rossi-Hansberg. The geography of development. *Journal of Political Economy*, 126(3):903 – 983, 2018.

Rebecca Diamond. The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524, 2016.

Dave Donaldson and Richard Hornbeck. Railroads and american economic growth: A âœmarket accessâ approach. *The Quarterly Journal of Economics*, 131(2):799–858, 2016.

Jonathan Eaton and Samuel Kortum. Technology, geography, and trade. *Econometrica*, 70 (5):1741–1779, 2002.

Ruben Enikolopov and Ekaterina Zhuravskaya. Decentralization and political institutions. *Journal of Public Economics*, 91(11):2261–2290, 2007. ISSN 0047-2727.

Dennis Epple and Thomas Nechyba. Fiscal decentralization. In J. V. Henderson and J. F. Thisse, editors, *Handbook of Regional and Urban Economics*, volume 4, chapter 55, pages 2423–2480. Elsevier, 1 edition, 2004.

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

Beatrix Eugster and Raphaël Parchet. Culture and taxes. *Journal of Political Economy*, 127 (1):296–337, 2019.

Beatrix Eugster, Rafael Lalive, Andreas Steinhauer, and Josef Zweimüller. The demand for social insurance: does culture matter? *The Economic Journal*, 121(556):F413–F448, 2011.

Jean-Paul Faguet. Does decentralization increase government responsiveness to local needs?: Evidence from bolivia. *Journal of Public Economics*, 88(3):867–893, 2004. ISSN 0047-2727.

Pablo D Fajgelbaum, Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar. State taxes and spatial misallocation. *The Review of Economic Studies*, 86(1):333–376, 2019.

Lars P Feld and Gebhard Kirchgässner. The political economy of direct legislation: direct democracy and local decision–making. *Economic policy*, 16(33):330–367, 2001.

Raquel Fernandez and Richard Rogerson. Income distribution, communities, and the quality of public education. *The Quarterly Journal of Economics*, 111(1):135–164, 1996.

Sarah Flèche. The Welfare Consequences of Centralization: Evidence from a Quasi-Natural Experiment in Switzerland. *The Review of Economics and Statistics*, 103(4):621–635, October 2021.

Bruno S Frey and Alois Stutzer. Beyond outcomes: measuring procedural utility. *Oxford Economic Papers*, 57(1):90–111, 2005.

Bruno S Frey, Matthias Benz, and Alois Stutzer. Introducing procedural utility: Not only what, but also how matters. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, pages 377–401, 2004.

Patricia Funk and Christina Gathmann. Does direct democracy reduce the size of government? new evidence from historical data, 1890–2000. *The Economic Journal*, 121(557): 1252–1280, 2011.

Patricia Funk and Christina Gathmann. Voter preferences, direct democracy and government spending. *European Journal of Political Economy*, 32:300–319, 2013.

Elena Gentili, Giuliano Masiero, and Fabrizio Mazzonna. The role of culture in long-term care arrangement decisions. *Journal of Economic Behavior & Organization*, 143:186–200, 2017.

Bruce W. Hamilton. Zoning and property taxation in a system of local governments. *Urban Studies*, 12(2):205–211, 1975.

Bruce W. Hamilton. Capitalization of intrajurisdictional differences in local tax prices. *The American Economic Review*, 66(5):743–753, 1976.

John A Hartigan and Manchek A Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, 28(1):100–108, 1979.

Guojun He, Shaoda Wang, and Bing Zhang. Watering down environmental regulation in china. *The Quarterly Journal of Economics*, 135(4):2135–2185, 2020.

Keith Head and Thierry Mayer. Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of international economics*, volume 4, pages 131–195. Elsevier, 2014.

Louis Henkin. *Federalism and Decentralization: Constitutional Problems of Territorial Decentralization in Federal and Centralized States*, chapter Federalism, Decentralization and Human Rights. Routledge, 1987.

Randall G Holcombe and DeEdgra W Williams. Are there economies of scale in municipal government expenditures? *Public Finance and Management*, 9(3):416, 2009.

William H. Hoyt and Kangoh Lee. Subsidies as sorting devices. *Journal of Urban Economics*, 53(3):436–457, 2003.

Robert E Lane. Procedural goods in a democracy: How one is treated versus what one gets. *Social Justice Research*, 2(3):177–192, 1988.

David S. Lee and Thomas Lemieux. Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355, 2010.

E Allan Lind and Tom R Tyler. *The social psychology of procedural justice*. Springer Science & Business Media, 1988.

Erzo F. P. Luttmer. Neighbors as negatives: Relative earnings and well-being. *The Quarterly Journal of Economics*, 120(3):963–1002, 2005.

Isabel Z Martinez, Emmanuel Saez, and Michael Siegenthaler. Intertemporal labor supply substitution? evidence from the swiss income tax holidays. *American Economic Review*, 111(2):506–46, 2021.

John G Matsusaka. Direct democracy works. *Journal of Economic perspectives*, 19(2):185–206, 2005.

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

Daniel McFadden. The measurement of urban travel demand. *Journal of public economics*, 3 (4):303–328, 1974.

Ferdinando Monte, Stephen J. Redding, and Esteban Rossi-Hansberg. Commuting, migration, and local employment elasticities. *American Economic Review*, 108(12):3855–3890, December 2018.

Sendhil Mullainathan and Jann Spiess. Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106, 2017.

Wallace Oates. The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the tiebout hypothesis. *Journal of Political Economy*, 77(6):957–71, 1969.

Wallace E. Oates. An essay on fiscal federalism. *Journal of Economic Literature*, 37(3):1120–1149, 1999.

Francesc Ortega and Giovanni Peri. The effect of income and immigration policies on international migration. In *The Economics of International Migration*, chapter 11, pages 333–360. World Scientific Publishing Co. Pte. Ltd., 2016.

Ugo Panizza. On the determinants of fiscal centralization: Theory and evidence. *Journal of Public Economics*, 74(1):97–139, 1999.

Raphaël Parchet. Are local tax rates strategic complements or strategic substitutes? *American Economic Journal: Economic Policy*, 11(2):189–224, 2019.

Stephen J. Redding and Esteban Rossi-Hansberg. Quantitative spatial economics. *Annual Review of Economics*, 9(1):21–58, 2017.

Jennifer Roback. Wages, rents, and the quality of life. *Journal of Political Economy*, 90(6): 1257–1278, 1982.

Andrés Rodríguez-Pose and Roberto Ezcurra. Is fiscal decentralization harmful for economic growth? evidence from the oecd countries. *Journal of Economic Geography*, 11(4):619–643, 07 2010.

Sherwin Rosen. *Wage-based indexes of urban, quality of life*. Johns Hopkins University Press, 1979.

Stephen Ross and John Yinger. Sorting and voting: A review of the literature on urban public finance. In P. C. Cheshire and E. S. Mills, editors, *Handbook of Regional and Urban Economics*, volume 3, chapter 47, pages 2001–2060. Elsevier, 1 edition, 1999.

Daniel L. Rubinfeld. The economics of the local public sector. In A. J. Auerbach and M. Feldstein, editors, *Handbook of Public Economics*, volume 2, chapter 11, pages 571–645. Elsevier, 1 edition, 1987.

Lukas Rühli. Gemeindeautonomie zwischen illusion und realität, 2012.

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

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

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

Kurt Schmidheiny and Michaela Slotwinski. Tax-induced mobility: Evidence from a foreigners' tax scheme in switzerland. *Journal of Public Economics*, 167:293–324, 2018.

Amartya Kumar Sen. Democracy as a universal value. *Journal of democracy*, 10(3):3–17, 1999.

Andreas Steinhauer. Working moms, childlessness, and female identity. 2018.

Charles M. Tiebout. A pure theory of local expenditures. *Journal of Political Economy*, 64, 1956.

Hal R Varian. Big data: New tricks for econometrics. *Journal of Economic Perspectives*, 28 (2):3–28, 2014.

Marieke Voorpostel and Oliver Lipps. Attrition in the swiss household panel: Is change associated with later drop-out? *Journal of Official Statistics*, 27(2):301–318, 2011.

Marieke Voorpostel, Robin Tillmann, Florence Lebert, Ursina Kuhn, Oliver Lipps, Valérie-Anne Ryser, Erika Antal, Gian-Andrea Monsch, Nora Dasoki, Hannah S Klaas, et al. Swiss household panel user guide (1999–2018). *FORS: Lausanne, Switzerland*, 2020.

Sylvain Weber and Martin Péclat. Georoute: Stata module to calculate travel distance and travel time between two addresses or two geographical points. 2017.

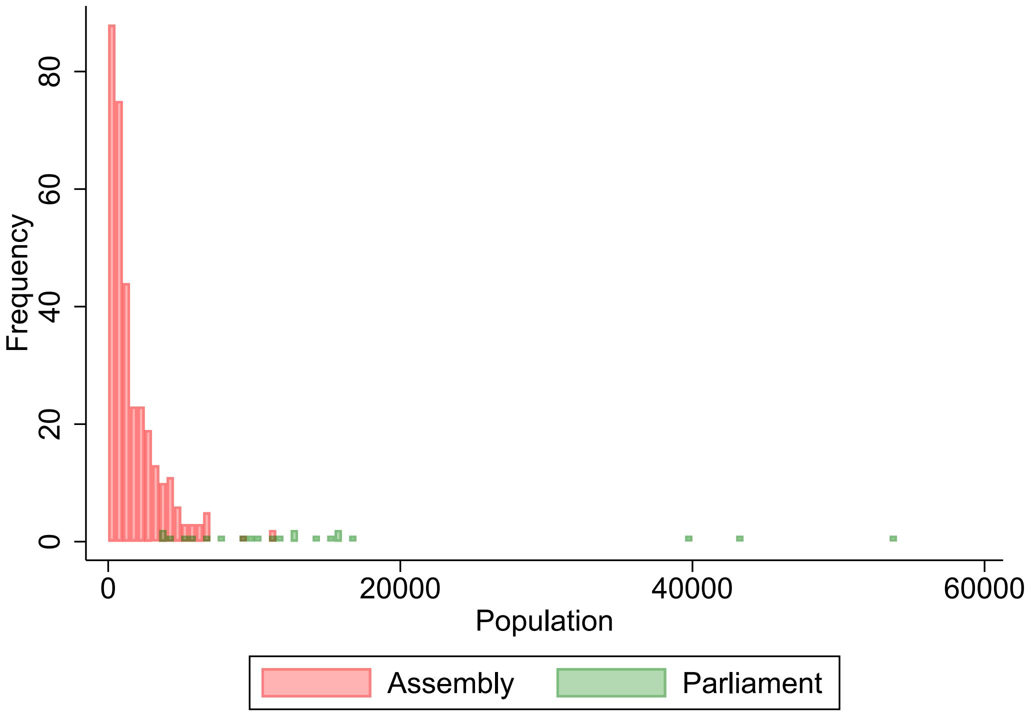
William C. Wheaton. Land capitalization, tiebout mobility, and the role of zoning regulations. *Journal of Urban Economics*, 34(2):102–117, 1993.

Online Appendix

A Online appendix: Supporting material

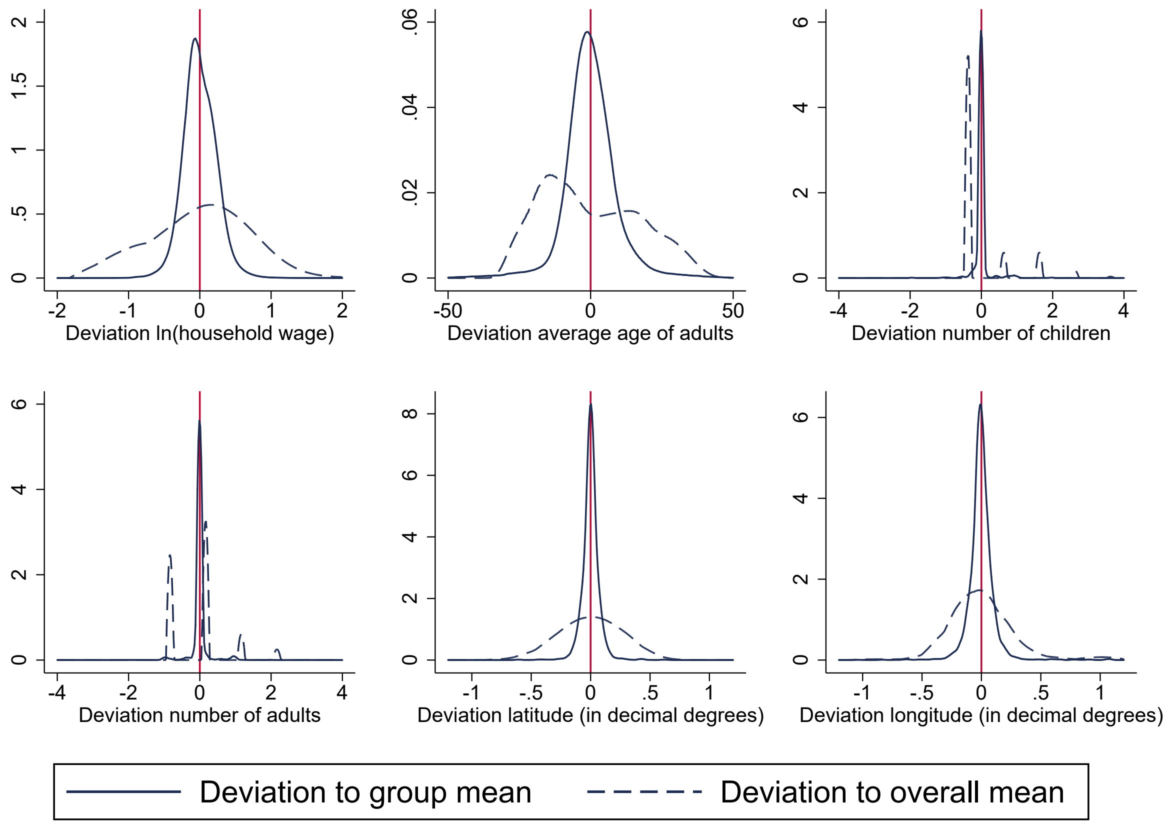
Figures

Figure A1: Histogram of local institutions in Bern: Assembly and parliament by population



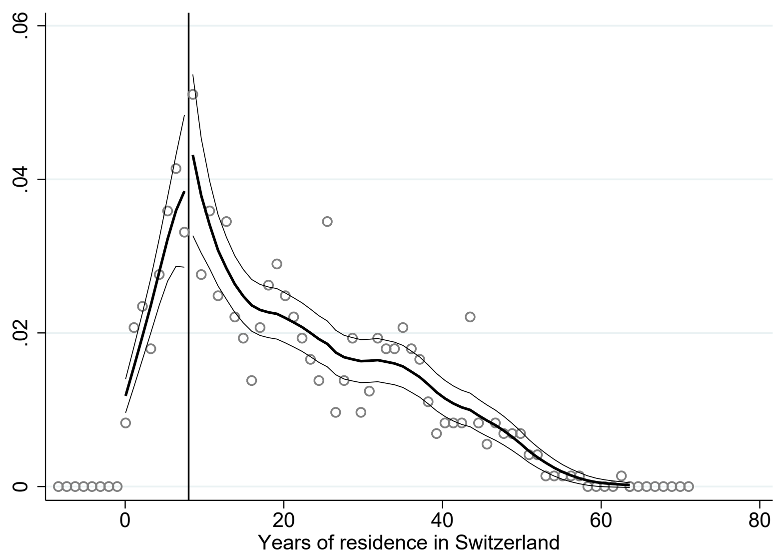
Notes: The figures provides a distribution of the municipalities in the canton of Bern across population size. The figure excludes the city of Bern with 130'015 inhabitants (parliament). Figure 1 provides a zoom to the population interval [3000, 14500].

Figure A2: Deviation of actual values to group and overall means



Notes: The figures compares kernel density estimates of the deviations between the actual values from their respective group means (solid line) with kernel density estimates of the deviations between the actual values from overall means (dashed line) by household variable.

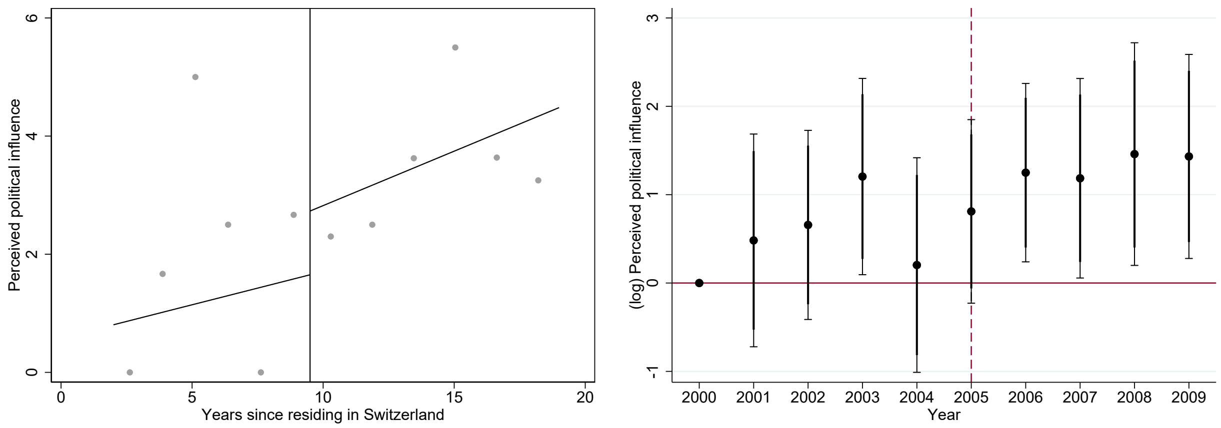
Figure A3: McCrary's smooth test (2008) of the running variable: Years of residence in Switzerland



Notes: The figure assesses the smooth variation of the running variable (i.e., years of residence in Switzerland) around the threshold (i.e., 8 years) following McCrary (2008).

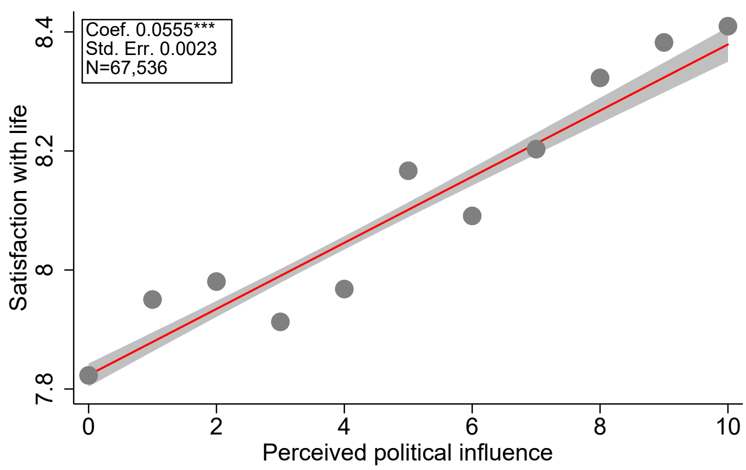
Figure A4: Graphical analysis of the procedural utility Parameter

(a) RD plot (b) Difference in Difference



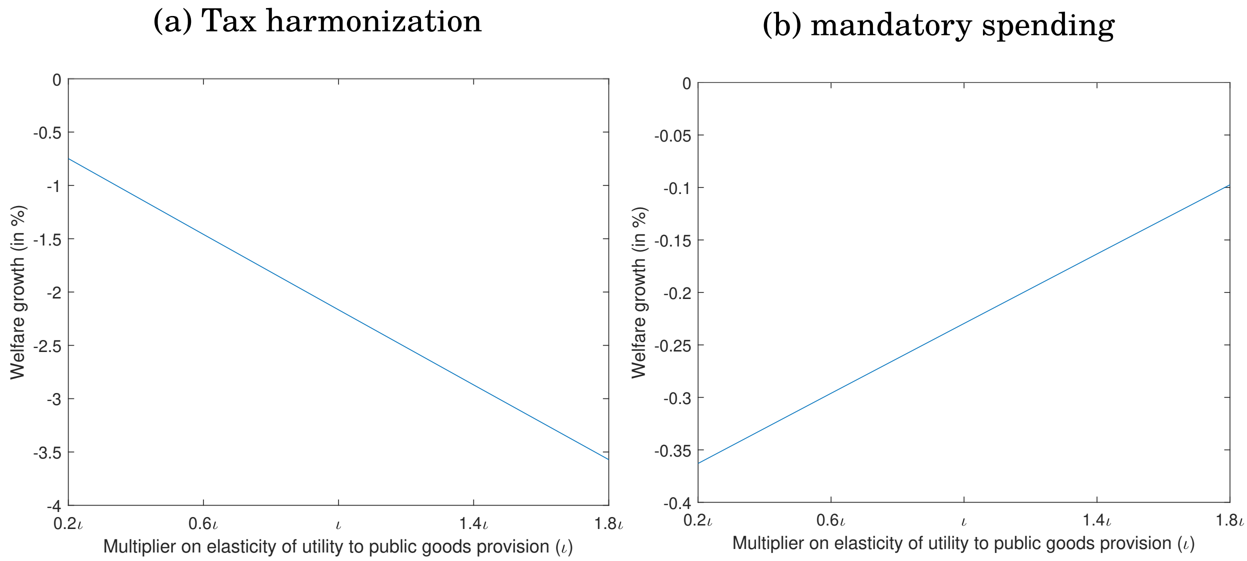
Notes: Panel (a) displays a RD plot – following Calonico et al. (2014) – of the effect of voting locally on perceived political influence. Threshold is 8 years of residency in Switzerland. Optimal bandwidth (of 7 years) is chosen. Panel (b) studies the same effect but exploits the difference across individuals in different cantons before and after the reform in Geneva.

Figure A5: Satisfaction with life and perceived political influence



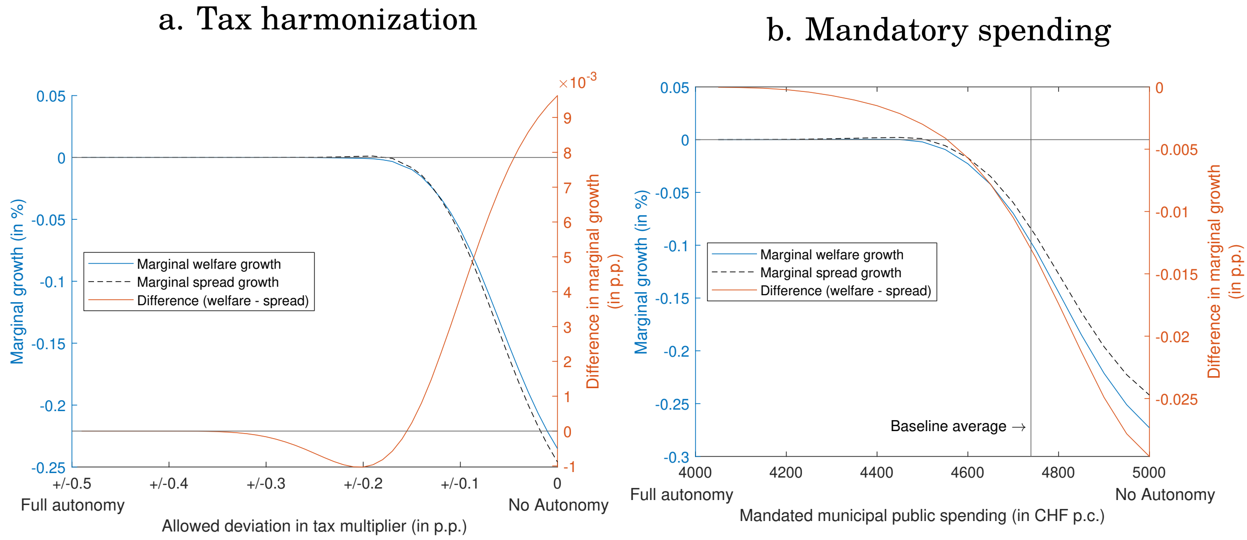
Notes: The Figure plots the average satisfaction with life by the individuals reported perception of influence into the political decision making process. We consider all individuals of age 18 and older that were interviewed between 2000 to 2009. The red line shows the predicted values from a linear regression (robust 95 percent confidence intervals are in gray). Number of observations: 67,536.

Figure A6: ROBUSTNESS OF RESULTS TO PERTURBATIONS OF THE PARAMETER GOVERNING THE VALUATION OF PUBLIC GOOD (ι)



Notes: We evaluate the robustness of the results to perturbations of the parameter governing the Valuation of public good (ι). Panel (a) focuses on tax harmonization; Panel (b) on mandatory spending. As summarizing moment, we focus on aggregate welfare growth.

Figure A7: THE EFFECT OF FISCAL AUTONOMY ON MARGINAL WELFARE AND INEQUALITY GROWTH



Notes on Figure A7a: Symmetric bounds (from below and above) are imposed on the local tax multiplier around the 2016 observed average tax multiplier (i.e., 1.719). The resulting bandwidth size, γ^{TH}, within which tax multipliers may lie is defined on the interval $\gamma^{TH} \in [0, 0.5]$ with 0.01 increments.

Notes on Figure A7b: Minimum spending levels are chosen around the observed average 2016 municipal spending level (i.e., CHF 4,739 per capita). The different spending minimum, γ^{M}, are defined on the interval $\gamma^{M} \in [4000, 5000]$ with increments of 50.

Tables

Table A1: CALIBRATION OVERVIEW

PARAMETERS COMMON TO ALL LOCATIONS AND HOUSEHOLD TYPES		
1. Preferences		
$\alpha = 0.70$	Consumption share in utility	FSO, Swiss consumer price index (2019)
$\epsilon = 2$	Shape parameter	Desmet et al. (2018); Diamond (2016), Ortega and Peri (2016), Monte et al. (2018) Fajgelbaum et al. (2019)
2. Benefits from self-determination		
$\rho = 0.011$	Procedural utility	Own estimation (Section 5.2);
$\tilde{\rho} = 0.017$	Information advantage	Own estimation (Section B)
3. Public good provision characteristics		
$\chi = 0.989$	Economies of scale parameter	Own estimation (Section 5.5)
$\delta = 0.37$	Equalization parameter	Official cantonal parameter
$\iota = 0.2$	Elasticity of utility to public goods	Aronsson and Johansson-Stenman (2008); Luttmer (2005), Fajgelbaum et al. (2019).
$f_m\left(\frac{H_n^1}{H_n}, \ldots, \frac{H_n^\Theta}{H_n}\right)$	Comp.-specific pref. for tax multip.	Own estimation (Section 5.3)
$f_r\left(\frac{H_n^1}{H_n}, \ldots, \frac{H_n^\Theta}{H_n}\right)$	Comp.-specific pref. for housing tax	Own estimation (Section E.1)
4. Mobility		
ψ^θ	HH-type-specific semi-elast. of migr.	Own estimation (Section 5.4)
$\eta_1 = 0.15$	Elast. of travel time in commuting	Axhausen et al. (2008); Becker (1965),
$\eta_2 = 0.17$	... interacted with income	Börjesson and Eliasson (2014)
HOUSEHOLD-TYPE-SPECIFIC CHARACTERISTICS		
θ	Household types	Own derivation (Section 5.1)
LOCATION-SPECIFIC CHARACTERISTICS		
b_n	Residential amenities	Own derivation (Section 5.6)
ϕ_n	Residential density	Own derivation (Section 5.6)

Notes: FSO is the Federal Statistical Office ("Bundesamt für Statistik" in German). "Comp.", "elas.", "HH", "migr.", "multip." and "pref." stands for composition, elasticity, household, migration, multiplier and preference, respectively. The legal base of the financial equalization parameter forms Article 10 of the cantonal law on fiscal equalization schemes ("Gesetz über den Finanz- und Lastenausgleich (FILAG)" in German).

Table A2: ACCURACY OF HOUSEHOLD TYPES (2012 TO 2016)

	(1)	(2)	(3)	(4)
	MAE	**RMSE**	**p10**	**p90**
Ln of total household wage (in CHF)	0.1785	0.2303	0.0310	0.3606
Average age of adults	6.5147	9.3001	0.8932	13.5927
Number of adults	0.0740	0.2341	0.0010	0.1165
Number of children	0.0642	0.2318	0.0005	0.0998
Latitude (in decimal degrees)	0.0463	0.0765	0.0038	0.1035
Longitude (in decimal degrees)	0.0696	0.1266	0.0088	0.1445

Notes: (1) Mean Absolute Error: $MAE = \frac{1}{M}\sum_{\theta}^{\Theta}\sum_{i\in\theta}|x_i^{\theta} - \mu_{\theta}|$; (2) Root Mean Square Error $RMSE = \sqrt{\frac{1}{M}\sum_{\theta}^{\Theta}\sum_{i\in\theta}(x_i^{\theta} - \mu_{\theta})^2}$; (3) 10^{th} percentile; (4) 90^{th} percentile. The length of 0.0463 degrees of latitude at the longitude of Bern is equivalent to about 5.1km. The length of 0.0251 degrees of longitude at the latitude of Bern is equivalent to about 5.4km. Observation window: 2012 to 2016. Number of households: 2,092,834; Number of groups (θ): 250.

Table A3: BALANCING TESTS FOR RDD ANALYSIS ON PROCEDURAL UTILITY

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Demographic characteristics			Educational level		Origin country	
Outcome	Age	Gender	Civil status	Educ. cat.	Educ. years	France	Neighboring
Treatment	3.068	-0.157	0.134	0.886	1.108	0.094	-0.033
	(7.403)	(0.216)	(0.523)	(2.025)	(1.985)	(0.157)	(0.178)
Bandwidth	Opt.	Opt.	Opt.	Opt.	Opt.	Opt.	Opt.
Obs.	202	202	202	202	202	202	202
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	No	No	No	No	No	No

Notes: This table presents the non-parametric estimation of balancing tests at the individual level following Calonico et al. (2014). Threshold is 8 years of residency in Switzerland. France and neighboring outcomes are indicator variables equal to unity if the foreign individual is from France or a Swiss neighbor (France, Germany, Italy, and Austria) and zero otherwise, respectively. Robust standard errors are in parentheses.

Table A4: Estimation for informational advantage

	(1) Full sample	(2) Only 2016	(3) Fully sample
Population	-0.101***	-0.092***	-0.033***
	(0.010)	(0.031)	(0.009)
Assembly	-0.768***	-0.654**	-0.513***
	(0.099)	(0.318)	(0.091)
Assembly $\times$ Population	0.084***	0.070**	0.052***
Adj. R^2	0.05	0.02	0.20
Obs.	5280	352	5280
Year FE	Yes	-	Yes
Covariates	No	No	Yes
Information Advantage ($\tilde{\rho}$)	0.0170	0.0179	0.0174

Notes: This table presents the estimation of informational advantage by exploiting the difference in access to local decision across local assembly and parliaments as population increases. Dependent variable is the log municipal tax multiplier. The information advantage is calculated as the population weighted marginal effect from having an assembly. Robust standard errors are in parentheses.

Table A5: Estimation for economies of scale in administrative spending

	(1) IV-Second Stage	(2) IV-First Stage	(3) OLS
Log number of households	0.892***	-	0.898***
	(0.023)		(0.021)
Log area	-	1.106***	
		(0.070)	
log altitude	-	-2.191***	
		(0.166)	
Constant	7.970	11.30	7.933
	(0.133)	(0.46)	(0.123)
Observations	352	352	352
R-squared	0.946	0.477	0.946
F-statistic	-	127.81	1841.5

Notes: This table presents the OLS and 2SLS estimates of the log number of households on log municipal expenditure on administrative matters for the year 2016. Dependent variable in (1) and (3) is log municipal expenditure on administrative matters. Dependent variable in (2) is the log number of households. Robust standard errors are in parentheses.

B Online Appendix: Benefits from self-determination due to an informational advantage ($\tilde{\rho}$)

Aside from procedural utility benefits, a main argument in favor of local autonomy and decentralised decision making is concerned with an information advantage of local residents (Feld and Kirchgässner, 2001; Matsusaka, 2005; Funk and Gathmann, 2011, 2013). As Oates (1999) puts it: *"[individuals] possess knowledge of both local preferences and cost conditions that a central agency is unlikely to have."*

To assess the robustness of our results to the definition of self-determination benefits, we exploit institutional variation across municipalities in the Canton of Bern. As revealed in Figure 1, 93% of municipalities in the Canton have a citizen's assembly as executive body, including municipalities up to 11,000 inhabitants. At the same time, some municipalities with as little as 3,500 inhabitants already have a parliament. This rare setting in international comparison can be used as a source of variation in citizen's access to decision making. The underlying assumption made here is that one's access to decision making – and thus, one's capacity in exercising any informational advantage – is more rapidly decreasing with population when the executive body is an assembly, as opposed to a parliament.

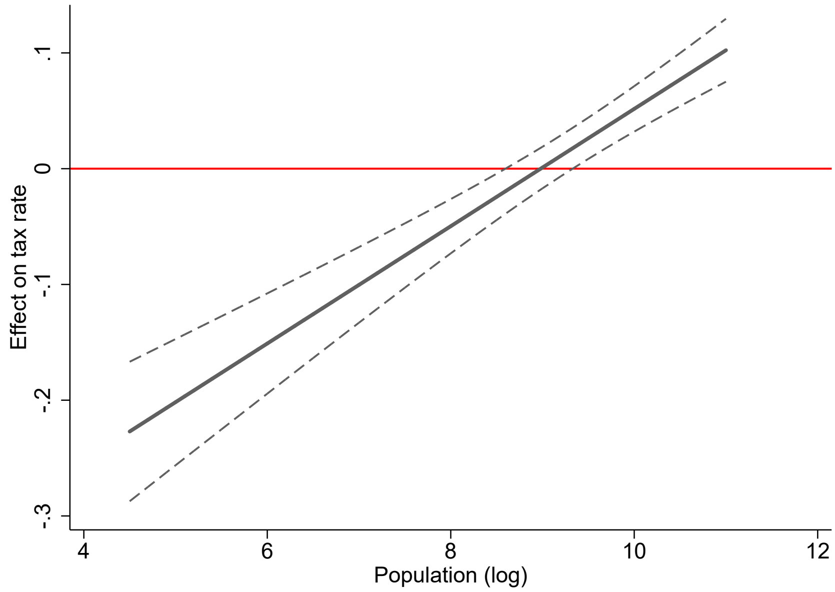
Formally, this setting may be used to estimate:

$$ln(m_{n,t}) = \beta_1 assembly_n + \beta_2 pop_{n,t} + \beta_3 (assembly_n \times pop_{n,t}) + \mu_t + \epsilon_{n,t}, \tag{18}$$

where $m_{n,t}$ is the tax multiplier of municipality n in year t. $assembly_n$ is a dummy variable. It is equal to one if the legislative body of a municipality is an assembly, and zero if it is a parliament. $pop_{n,t}$ is the population of municipality n in year t. β_3 is the coefficient of interest, measuring how population size affects the choice of local multiplier in municipalities with an assembly, relative to those with a parliament. Finally, μ_t denotes year fixed effects and $\epsilon_{n,t}$ the error term. In our preferred specification, we draw on a panel including all municipalities of the Canton of Bern between between 2002 to 2016. We present the obtained coefficient values from the underlying regression in Column 1 of Table A4. Moreover, we present the obtained coefficient values from two alternative models in Columns 2 (focusing only on 2016) and 3 (including municipal-time covariates) of Table A4 as a robustness check.

Figure A8 plots the obtained marginal effects of having an assembly depending on municipal population size. Compared to parliamentary municipalities, tax rates tend

Figure A8: GRAPHICAL ANALYSIS OF INFORMATION ADVANTAGE

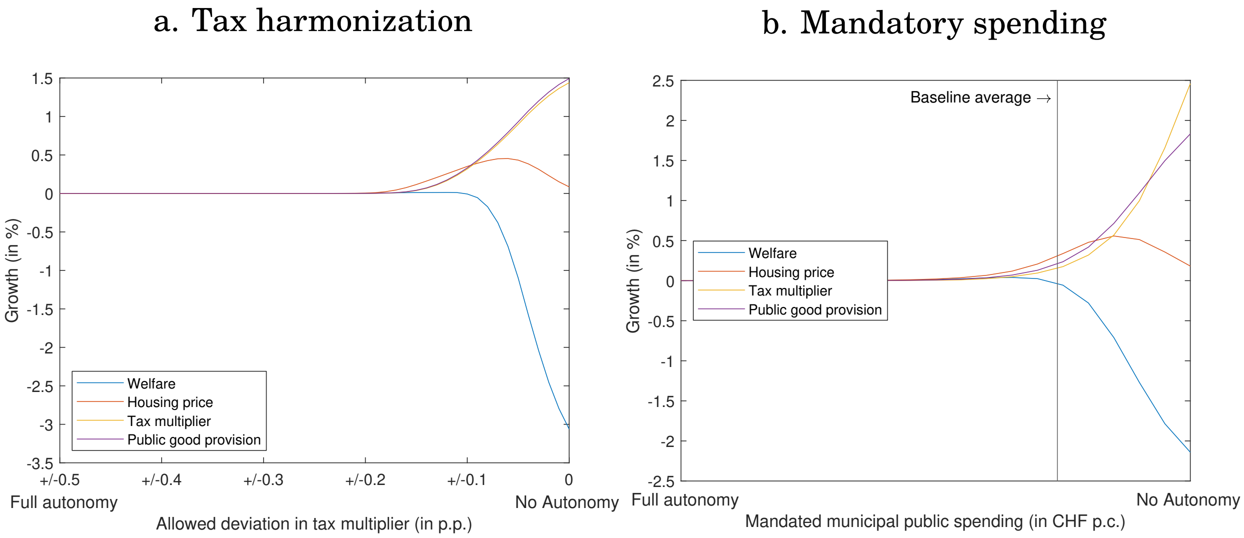


Notes: The figure plots the marginal effects of having an assembly on log municipal tax multipliers by log municipal population size (see Column 1 of Table A4 for the coefficient values of the underlying regression). The dashed lines represent 95% confidence intervals.

to be significantly lower in assembly municipalities when the municipal population is small, and vice versa. These results relate to the findings on the the deviation from actual to preferred levels of government services under heterogeneous levels of direct democratic involvement (Besley and Case, 2003; Brülhart and Jametti, 2019; Funk and Gathmann, 2011). We interpret them with heterogeneity in the effectiveness of assemblies across municipal population size. Town hall meetings, in which citizens gather at a particular place and time to make public decisions are the most effective institution of democratic control in smaller municipalities, but become increasingly inefficient as the population gets larger. Based on these results we calculate the canton wide benefits of self-determination that arise from an informational advantage as the population weighted mean marginal effect from having an assembly. The calculated coefficient values are displayed in Table A4 and vary little across specifications. We then model the gains from self-determination due to an information advantage by defining $O_n = 1 - \mathbb{1}(n)\tilde{\rho}$. Again, the indicator function $\mathbb{1}(n)$ takes on one if citizens of municipality n decide autonomously on municipal spending or tax levels, and zero otherwise, while $\tilde{\rho}$ denote the respective percentage gains from an information advantage.

The two graphs of Figure A9 presents the obtained simulation results from our

Figure A9: Equilibrium effects when defining benefits from self-determination as an informational advantage



Notes: The graphs show the aggregate results from our two counterfactual scenarios defined in 6 when the benefits of self-determination result form an information advantage ($O_n = 1 - \mathbb{1}(n)\tilde{\rho}$). Panel (a) shows the results when imposing Symmetric bounds (from below and above) on the local tax multiplier around the 2016 observed average tax multiplier (i.e., 1.719). Growth is expressed in percent relative to full tax harmonization; Panel (b) shows the results when imposing a minimum level of mandatory spending (per capita) around the observed average 2016 municipal spending (i.e., CHF 4,739 per capita). Growth is expressed in percent relative to a minimum mandatory spending of CHF 4,000.

two counterfactual scenarios defined in 6 when focusing on the information advantage channel of self-determination. A9a reports the aggregate effects of harmonizing local tax multipliers. As in 4a the horizontal axis shows the absolute value of the different bounds imposed on local tax multipliers, while the vertical axis displays overall growth in percent.

Consistent with our analysis that focuses on the procedural utility channel of self-determination, overall welfare growth increases with the allowed bandwidth around the 2016 observed average tax multiplier. The calculated gains from self-determination through the information advantage channel slightly exceed the ones through the procedural utility channel. Consequently the maximum welfare loss from an abolition of municipalities tax-setting autonomy tends to exceed the one obtained in our section on the procedural utility channel (3% compared to around 2% in Subsection 6.1).

As shown in 4a, the same is true when imposing a lower bound on mandatory spending levels. Again, we observe no average effect on welfare up to a lower bound of around CHF 4,600 per capita. However, maximum welfare losses for a minimum spending requirement of CHF 5,000 reach around 0.8% compared to around 0.5% in Subsection 6.2 of the main paper.

C Online appendix: Proof of equilibrium existence and uniqueness

In this Section, we prove the existence and uniqueness of the equilibrium defined by the system of equations in Section 3.

The elements of the equilibrium vector are determined by the following system of equations:

$$\pi_n^\theta = \frac{\sum_i^N v_{ni}^\theta}{\sum_n^N \sum_i^N v_{ni}^\theta},$$

$$O_n = 1 + \mathbb{1}(n)\rho,$$

$$m_n = f_m\left(\frac{H_n^1}{H_n}, \frac{H_n^2}{H_n}, ..., \frac{H_n^\Theta}{H_n}\right),$$

$$r_n = f_r\left(\frac{H_n^1}{H_n}, \frac{H_n^2}{H_n}, ..., \frac{H_n^\Theta}{H_n}\right),$$

$$g_n = y_n + \delta\Big(y - y_n^h\Big)$$

$$Q_n = (1-\alpha)\frac{\sum_\theta^\Theta \mathbb{E}[w_i^\theta|n]H_n^\theta}{L_n},$$

Step 1: Using the probability to locate in n, we have:

$$\pi_n = \sum_\theta^\Theta \pi_n^\theta = \sum_\theta^\Theta \frac{\sum_i^N v_{ni}^\theta}{\sum_n^N \sum_i^N v_{ni}^\theta}. \tag{19}$$

Let us define $V_n = \Big(g_n^\iota(1-m_n)O_n\Big)^\epsilon \Big(\big((1+r_n)Q_n\big)^{1-\alpha}\Big)^{-\epsilon}$. To insure separability of m_n and t^θ, we approximate $1-m_n t^\theta$, by $(1-m_n)(1-t^\theta)$. Given land clearing (9), we can rewrite this condition as a system of N equations for the N unknown V_n as follows:

$$D_n(V) = \phi_n \tilde{L}_n - H\sum_\theta^\Theta \frac{V_n \sum_i^N ((1-t^\theta)w_{ni}^\theta)^\epsilon (d_n^\theta)^{-\epsilon}}{\sum_n^N \sum_i^N V_n((1-t^\theta)w_{ni}^\theta)^\epsilon (d_n^\theta)^{-\epsilon}} = 0. \tag{20}$$

Lemma 3: The system in (20) exhibits the following properties:

Property 1: $D(V)$ is continuous.

Property 2: $D(V)$ is homogeneous of degree zero.

Property 3: $\sum_{n=1}^{N} D_n(V) = 0$.

Property 4: $D(V)$ exhibits gross substitution:

$$\frac{\partial D_n(V)}{\partial V_m} > 0, \quad \forall \text{ n, m, n} \neq \text{m} \tag{21}$$

$$\frac{\partial D_n(V)}{\partial V_n} < 0, \quad \forall \text{ n} \tag{22}$$

PROOF: Properties 1 and 2 of Lemma 3 follow directly from an inspection of (20). Property 3 is satisfied by noting:

$$\begin{aligned}\sum_{n=1}^{N} D_n(V) =& \sum_{n=1}^{N} \phi_n \tilde{L}_n - \sum_{n=1}^{N} \left[\sum_{\theta}^{\Theta} \frac{\sum_i^N V_n (B_n w_{ni}^{\theta})^{\epsilon} (d_n^{\theta}(1-t^{\theta})^{(1-\alpha)})^{-\epsilon}}{\sum_n^N \sum_i^N V_n (B_n w_{ni}^{\theta})^{\epsilon} (d_n^{\theta}(1-t^{\theta})^{(1-\alpha)})^{-\epsilon}} \right] \\ =& 1 - 1 \\ =& 0.\end{aligned} \tag{23}$$

Property 4 can be established by noting:

$$\frac{\partial D_n(V)}{\partial V_m} = \frac{V_n[(w_{ni}^{\theta})^{\epsilon}(d_n^{\theta}(1-t^{\theta})^{(1-\alpha)})^{-\epsilon}]^2}{[V_n(w_{ni}^{\theta})^{\epsilon}(d_n^{\theta}(1-t^{\theta})^{(1-\alpha)})^{-\epsilon}]^2} > 0. \tag{24}$$

Using property 2, which implies $\nabla D_n(V)V = 0$, it follows that:

$$\frac{\partial D_n(V)}{\partial V_n} < 0, \quad \forall \text{ n}. \tag{25}$$

Thus, gross substitution is established.

Q.E.D.

Lemma 4: There exists a unique vector V which solves (20).

PROOF: We proceed in two steps. First, we show that there exists at most one (normalized) vector V which solves (20). Second, we show a vector V that solves (20) exists.

Gross substitution requires that $D(V) = D(V')$ cannot occur if V and V' non-collinear vectors. By homogeneity of degree zero, we can assume that $V' \geq V$ and $V_n = V'_n$ for some i. Now suppose that we lower (or keep constant) V' in all locations except in n one at a time. By gross substitution, V_n will increase in at least one step. Hence, $D(V) > D(V')$ which is a contradiction.

By homogeneity of degree zero, the search for an equilibrium vector can be restricted to the unit simplex $\Delta = \{\sum_{n=1}^{N} V_n = 1\}$. Define on Δ the function $D^+(\cdot)$ by $D_n^+(V) = \max\{D_n(V), 0\}$. $D^+(\cdot)$ is continuous. Denote $\alpha(V) = \sum_{n=1}^{N}[V_n + D_n^+(V)]$ with $\alpha(V) \geq 1, \forall V$. Then define the function $f(\cdot)$ from the closed convex set Δ into itself as:

$$f(V) = [1/\alpha(V)][V + D^+(V)]. \tag{26}$$

By Brouwer's Fixed Point Theorem (1911), there exist a $V^* \in \Delta$ such that $V^* = f(V^*)$. Since $\sum_{n=1}^{N} D_n(V) = 0$, it follows that at the fixed point for amenity, $V^* = f(V^*)$ and $D_n(V) = 0$ for all i.

Q.E.D.

Step 2: Given V_n, and under the "rich chooses first" allocation mechanism (see Section 3), we can derive π_n^θ.

Step 3: Solve for the tax multiplier m_n using machine learning approach presented in Section 5.3. Note that the municipal composition matters to derive m_n which in turns affects the location choice of each group.

Step 4: For each municipality n, given the equalization parameter (δ), household wages (w^θ) and residential densities (H_n^θ), we can solve for the local public good spending g_n using:

$$g_n = y_n + \delta\left(y - y_n^h\right),$$

Step 5: Solve for O_n given preferred m_n and g_n as well as the fiscal limitations imposed to municipalities across the different simulation scenarios.

Step 6: Finally, housing prices per housing unit (Q_n) can then be derived using the land market clearing condition:

$$Q_n = (1-\alpha)\frac{\sum_{\theta}^{\Theta} w^{\theta} H_n^{\theta}}{L_n},$$

Q.E.D.

D Online appendix: Data

The calibration of our model is based on both, household and municipal level data. In this section we provide a detailed description of the data sources and the data preparation process.

Matched earning-census data

The calibration of our model is primarily based on individual level data from the universe of inhabitants in the Canton of Bern for the period spanning the years between 2012 and 2016. However, our paper also profits from individual level data of earlier years and the other cantons in Switzerland. Therefore, we merge the register-based population census data of Switzerland for the years between 2012 and 2016 via a social security number to 100% of the Social Security Earnings Records (SSER) from the Old-Age and Survivors' Insurance (OASI) for the period between 1981 and 2016.[35]

The register-based census data contain information about i) the individuals gender, age, nationality and marital status, ii) the municipality of residence, and iii) the municipality of residence 1, 2 and 5 years ago. The SSER data-set contains the near universe of individual annual wages. However, it does not include i) information about individuals capital income and ii) contributions from pension plans and the social security system. Therefore, our wage data contains mostly zeros for individuals that passed the statutory retirement age.[36] However, as the individual earnings record dates back to 1981, our data covers almost the full wage history of individuals that were retried in 2016. This allows us to complete the missing income of retired individuals as follows: We first calculate the decile rank of an individual's mean wage between the age of 50 and 59 within their age and gender cohort. We then replace the zero income of these individuals by matching their cohort rank with the contribution

[35] See for instance Martinez et al. (2021) who merge the same data, albeit for an earlier time period.

[36] This age was 65 for men and 64 for women throughout our sample period, albeit retirement at an earlier age is not uncommon.

deciles of the Old Age and Survivors Insurance (first pillar, data provided by Swiss Federal Department of Social Security (FDSS)) and of the occupational pension plans (second pillar, data provided by the Swiss Federal Office of Statistics (FSO)).

Table A6: Descriptive statistics (2016)

	(1) mean	(2) sd	(3) p10	(4) p90	(5) N
Panel A: Households					
Total household wage (in CHF)	100,002	73,726	31,977	181,184	445,393
Average age of adults	51.932	17.470	30.5	77	445,393
Number of adults	1.814	0.852	1	3	445,393
Number of children	0.372	0.807	0	2	445,393
Latitude of place of origin	46.967	0.188	46.717	47.194	445,393
Longitude of place of origin	7.535	0.281	7.247	7.822	445,393
Panel B: Municipalities					
Number of households	1265	3785	105	2271	352
Tax multiplier	1.719	0.202	1.49	1.95	352
Tax rate at CHF 100,000 (unmarried)	0.079	0.009	0.068	0.089	352
Tax rate at CHF 100,000 (married)	0.067	0.008	0.058	0.076	352
Expenditure (in million CHF)	17.415	77.874	1.149	29.141	352

Notes on Panel A: Adults are defined as aged 18 or above and children as aged 17 or below; Coordinates of the place of origin refer to the center of the municipality in which a household resided 5 years ago (2011). Center coordinates of municipalities are defined by the Swiss Federal Office of Topography based on different criteria. They are generally placed on a main road on the principal place of a municipality. If household members resided in different places 5 years ago, we take the mean coordinates of household members. *Notes on Panel B*: Tax rates are average municipal tax rates at a household income of CHF 100,000. Expenditure corresponds to total expenditure in million CHF.

Table A7: Descriptive statistics of household variables in Bern by year

	(1) mean	(2) sd	(3) p10	(4) p90	(5) N
2012					
Total household wage (in CHF)	98,825	77,747	29,658	179,760	374,329
Average age of adults	52.988	17.316	32	78	374,329
Number of children	0.376	0.813	0	2	374,329
Number of adults	1.868	0.886	1	3	374,329
2013					
Total household wage (in CHF)	99,441	74,767	30,718	180,613	394,759
Average age of adults	52.550	17.395	31	77.5	394,759
Number of children	0.372	0.809	0	2	394,759
Number of adults	1.854	0.874	1	3	394,759
2014					
Total household wage (in CHF)	99,702	77,242	31,977	180,541	437,736
Average age of adults	51.613	17.461	30.333	77	437,736
Number of children	0.376	0.810	0	2	437,736
Number of adults	1.826	0.859	1	3	437,736
2015					
Total household wage (in CHF)	100,245	74,987	31,977	181,466	440,617
Average age of adults	51.801	17.465	30.5	77	440,617
Number of adults	0.374	0.808	0	2	440,617
Number of adults	1.824	.856	1	3	440,617
2016					
Total household wage (in CHF)	100,002	73,726	31,977	181,184	445,393
Average age of adults	51.932	17.470	30.5	77	445,393
Number of children	0.372	0.806	0	2	445,393
Number of adults	1.814	0.852	1	3	445,393

Notes: We define adults as aged 18 or above and children as aged 17 or below. For the years 2014 to 2016 the data comprise the universe of household in the Canton of Bern. For the years 2012 and 2013 around 11%, and 9% respectively, of households are excluded due to the missing household number.

Through a household identification variable we construct a data-set for the universe of households in Bern for the time period spanning the years 2012 to 2016.

After accounting for pension incomes, around 9% of households still record a wage below the basic financial minimum guaranteed by the cantonal Social Assistance Act.[37] We replace these low wage values by the lower bound of the guaranteed amount (dependent on household size) that is outlined in the Social Assistance Ordinance.[38] We define a household's place of origin, as the center coordinates of the municipality in which a household resided in $t-5$. This center coordinates are defined by the Swiss Federal Office of Topography and generally placed on the main road or main square of the principal place of a municipality. We also experiment with centroid coordinates. However, center coordinates appear superior when calculating routing distances since centroid coordinates are often located away from main roads. If household members resided in different places in $t-5$, we take the mean coordinates of household members. Finally, We define adults as aged 18 or above and children as aged 17 or below. Descriptive statistics for our final data-set at the household level are provided in Table A7.

E Online appendix: Calibration

E.1 Preferences for housing tax rate

Equation (4) states that, as local income tax multipliers, local housing tax rates, depend on the distribution of household types in a given municipality:

$$r_n = f_r\left(\frac{H_n^1}{H_n}, \frac{H_n^2}{H_n}, ..., \frac{H_n^\Theta}{H_n}\right).$$

We predict housing tax rates following the same methodology as in the prediction of municipal tax multiplier. Table A8 provides the out-of-sample performance accuracy measures for the different methods. As apparent from the table, the Gaussian Process Regression largely outperforms the other methods and receives a weight of 1 following the model averaging procedure described in Section 5.3 of the main paper.

[37] https://www.belex.sites.be.ch/frontend/versions/1213

[38] https://www.belex.sites.be.ch/frontend/versions/1616?locale=de

Table A8: Housing tax rate: accuracy of out-of-sample prediction

	(1) MAPE	(2) RMSPE	(3) p10	(4) p90	(5) Weight in Ensemble
Model					
Mean	13.73%	17.00%	1.01%	21.22%	0.00
Linear Regression	10.15%	13.11%	2.03%	21.26%	0.00
Random Forest	6.98%	9.13%	0.77%	15.30%	0.00
Support Vector Machine	7.52%	11.12%	0.87%	16.84%	0.00
Gaussian Process Regression	4.69%	6.57%	0.43%	11.00%	1.00

Notes: The training period is: 2012 to 2015. Each year contains the entire set of 352 municipalities. All goodness-of-fit measures refer to the out-of-sample year: 2016. For the Support Vector Machine Regression and the Gaussian Process Regression we select the Kernel Function with the lowest in-sample RMSPE (exponential for the Gaussian Process Regression and cubic for the Support Vector Machine Regression). The Random forest is averaged over 1000 trees; p10 and p90 are calculated on the absolute percentage error.

E.2 Calibration of municipal fundamentals

Density of development

We retrieve the municipality-specific density of development using the land market clearing condition, as follows:

$$\phi_n = (1-\alpha)\frac{\sum_{\theta}^{\Theta} w^{\theta} H_n^{\theta}}{Q_n \tilde{L}_n} \tag{27}$$

Fundamental residential amenities

Existence and (up-to-scale) uniqueness of B_n can be shown following the same reasoning as the proof for V_n in Section C.