Highways, Taxes, and Voting

Very preliminary - Please do not cite

Paolo Campli and Raphaël Parchet (USI)

October 21, 2018

1 Introduction

In this paper we study the aggregation of heterogeneous preferences of Swiss residents regarding municipal tax rates, using the change in the composition of municipal population after the opening of nearby highway accesses and their subsequent tax rate choice to measure the political weight of individuals belonging to different income classes. The fundamental problem of aggregation of heterogeneous preferences has been addressed by a large theoretical literature, with many models of policy choice in political economy (Persson and Tabellini, 2000). There has also been some experimental approach to the question, for example Ambrus, Greiner and Pathak (2015), who show some interesting deviations from the median voter result, which can be interpreted as a non-zero weight given to the preferences of agents close to median. We aim to study a particular example of this general problem in a real world context.

We start by building a simple spatial equilibrium model (á la Redding and Rossi-Hansberg 2017), in which residents optimally locate given their heterogeneous preferences for local amenities and public goods in a set of peripheral municipalities, taking the cost of transportation along the network into account. Local taxes are set as a linear combination of the agents preferences and used to finance a local public good. We then follow Fretz, Parchet and Robert-Nicoud (2017) who show that the construction of the Swiss highway network increased the number of residents in all income classes in newly connected municipalities, but disproportionately so in high-income classes. We show that the opening of a new highway connection is followed by a decrease in local tax rates. We combine

our theoretical model and the empirical results to recover, via structural estimation, the "political weight" of individuals belonging to different income classes and their private cost of commuting.

The paper proceeds as follows: in section 2 we introduce and develop the theoretical model, in section 3 we go over the identification strategy and the reduced form empirical results, in section 4 we conclude with some preliminary results on the structural estimation and mention further steps we will take in the near future.

2 Theoretical model

In this section we introduce a theoretical model that features location decisions by individuals in multiple income classes and a tax setting mechanism for each municipality. We later use the reduced form results and a log-linearization of the model to estimate a parameter related to the agents' individual political weight and, for each income class, one parameter connected to private commuting costs.

Setting and location decision

We assume the following form for the indirect utility function:

$$V_{ijm} = \kappa + \log\left[(1-\tau_j)w_m\right] + \delta_m \log g_j + A_{ijm} \tag{1}$$

where the triplet (i, j, m) stands for the individual, the municipality and the income class, respectively. We indicate with w_m the income of income class m and with g_j the public good available in municipality j. The term A_{ijm} is an idiosyncratic preference shifter that will include a commuting cost part in the following. The public good is financed through taxes with a balanced budget: $g_j = C_j \bar{w}_j$, where $\bar{w}_j$ is average income in the municipality (i.e. we assume full congestion for the public good).

The idiosyncratic preference shifter comprises of a location-specific component (which measures the commuting cost: $-\mu_m \log(w_m C_j)$, where C_j is the yearly commuting time from location j to the center, and a location-specific idiosyncratic preference ξ_{ij}, which we assume is distributed as an i.i.d.

Gumbel distribution with mean zero, variance σ^2 and scale parameter $\beta = \frac{\pi}{\sigma\sqrt{6}}$.

The individual will choose the municipality j that maximizes his indirect utility function. We use the notation $V_{ijm} = u_{jm} + \xi_{ij}$. The share of of households with income m choosing to reside in municipality j is given by:

$$N_{jm} = \mathbb{P}(V^*_{ijm} > V^*_{ij'm} \forall j' \neq j) = \frac{\exp \beta u_{jm}}{\sum_{j'} \exp \beta u_{j'm}} \tag{2}$$

Tax setting

Tax rates in the municipality are chosen by some form of democratic process. We avoid an explicit modeling assumption at this stage and use a functional form that can be estimated with our data. We simply assume that the tax rate is a weighted average of the preferred tax rates of the various income classes, with the weights depending both on the fraction of citizens belonging to the class and a parameter (to be empirically determined) summarizing the per capita political weight conditional on income.

$$\tau_j = \sum_m \lambda_m s_{jm} \tau^*_m \tag{3}$$

where, from the utility function, $\tau^*_m = \frac{\delta_m}{1+\delta_m}$, and we use the notation $s_{jm} = \frac{N_{jm}}{N_j}$. More in general, the mechanics of our tax setting equation embeds many of the benchmark models of policy choice:

- Utility maximization $\rightarrow \lambda_m$'s all roughly one
- Median voter $\rightarrow \lambda_{med} = \frac{1}{s_{j,med}}$, others are zero
- Lobbying $\rightarrow \lambda_{rich} \gg \lambda_{poor}$
- Rawlsian $\rightarrow \lambda_{rich} \ll \lambda_{poor}$
- Revenue maximization $\rightarrow \lambda_m = f(\epsilon_m)$, ϵ_m tax base elasticity

Log-linearization of the model

After developing the full model, we log-linearize it to obtain a matrix that connects the variations in the endogenous variables, N_{jm}, τ_j to shocks in the values of commuting time caused by new highway connections. We assume throughout that municipalities are small, so that changes in the number of

residents in one locations only have a negligible impact on residents of other municipalities in the region.

Full model

For a general number of income classes, and indicating with $\gamma_j = \sum_m w_m N_{jm}$ the total tax base in municipality j, we get:

$$dN_{jm} = \beta N_{jm} \left[\frac{\partial V_{ijm}}{\partial \tau_j} d\tau_j + \frac{\partial V_{ijm}}{\partial g_j} dg_j \right] \tag{4}$$

Use that:

$$\hat{g}_j = \hat{\tau}_j + \sum_m \frac{y_m N_{jm}}{\gamma_j} \hat{N}_{jm} \tag{5}$$

$$\frac{1}{\beta} \hat{N}_{jm} = \frac{1}{\beta} \frac{dN_{jm}}{N_{jm}} = -\frac{\tau}{1-\tau} \hat{\tau} + \delta_m (\hat{\tau}_j + \sum_m \frac{\gamma_{jm}}{\gamma_j} \hat{N}_{jm}) \tag{6}$$

$$\left(\frac{1}{\beta} - \frac{\gamma_{jm}}{\gamma_j} \right) \hat{N}_{jm} = \sum_{k \neq m} \frac{\gamma_{jk}}{\gamma_j} \hat{N}_{jk} + \left(\delta_m - \frac{\tau_j}{1-\tau_j} \right) \hat{\tau}_j - \mu_m \hat{C}_j \tag{7}$$

Let us see the log-linearization of the tax rate equation now. As intuitive, it only gets contributions from changes in the N_{jm}'s that change the "political equilibrium":

$$\hat{\tau}_j = \frac{1}{\tau_j} \sum_m \lambda_m \tau_m^* s_{jm} \left(\hat{N}_{jm} - \hat{N}_j \right) \tag{8}$$

where the optimal tax for income class m is given by $\tau_m^* = \frac{\delta_m}{1+\delta_m}$.

$$\hat{\tau}_j = \frac{1}{\tau_j} \sum_m \lambda_m \tau_m^* s_{jm} \left(\hat{N}_{jm} - \sum_k s_{jk} \hat{N}_{jk} \right) = \sum_m \tilde{\tau}_{jm} \hat{N}_{jm} \tag{9}$$

where the $\tilde{\tau}_{jm}$ have the following expression:

$$\tilde{\tau}_{jm} = \frac{1}{\tau_j} \left[\lambda_m \tau_m^* s_{jm} (1 - s_{jm}) - \sum_{k \neq m} \lambda_k \tau_k^* s_{jk} s_{jm} \right] \tag{10}$$

More on the λ's

The λ's are a measure of political weight for individuals in income class m. The tax is then a weighted average of individual preferences ($\sum_m s_{jm} \lambda_m = 1$). For the case with only 2 income classes:

$$\lambda_1 \in [0, \frac{1}{s_1}] \qquad \lambda_2 \in [0, \frac{1}{s_2}]$$

$$\lambda_2 = \frac{1}{s_2}(1 - s_1\lambda_1)$$

Matrix form

For brevity and clarity of exposition, in the following we present first the simplest case: a two income-classes model with myopic agents, that is individuals do not anticipate any change in residents' income distribution and assume that public good provision will only change as per the direct effect of changes in the tax rate. In this case, the variation in the number of residents is given by:

$$\begin{bmatrix} \hat{N}_1 \\ \hat{N}_2 \\ \hat{\tau} \end{bmatrix} = \begin{bmatrix} \beta^{-1} & 0 & \frac{t}{1-t} - \delta_1 \\ 0 & \beta^{-1} & \frac{t}{1-t} - \delta_2 \\ \tilde{\tau}_1 & \tilde{\tau}_2 & -1 \end{bmatrix}^{-1} \begin{bmatrix} \mu_1 \\ \mu_2 \\ 0 \end{bmatrix} \hat{C}_j \tag{11}$$

where the coefficients $\tilde{\tau}_i$ can be read from the previous equation and take a simple form in the two classes case:

$$\tilde{\tau}_1 = \frac{s_{j1}s_{j2}}{\tau}\left[\lambda\tau_1^* + (1-\lambda)\tau_2^*\right] \tag{12}$$

and

$$\tilde{\tau}_2 = -\tilde{\tau}_1 \tag{13}$$

In this myopic version an analytic expression for the theoretical moments of the model can be obtained. Let's call $|A|$ the determinant of the previous matrix:

$$|A| = -\frac{1}{\beta}\left[\frac{1}{\beta} + \frac{s_{j1}s_{j2}}{\tau}\left[\lambda\tau_1^* + (1-\lambda)\tau_2^*\right](\delta_2 - \delta_1)\right] \tag{14}$$

Now we can write the theoretical moments in closed form as follows:

$$\begin{bmatrix} \tilde{N}_1 \\ \tilde{N}_2 \\ \tilde{\tau}_3 \end{bmatrix} = \frac{1}{\beta|A|}\begin{bmatrix} (1+\beta\tilde{\tau}_2\alpha_2)\mu_1 - \beta\tilde{\tau}_1\alpha_2\mu_2 \\ -\beta\tilde{\tau}_2\alpha_1\mu_1 + (1+\beta\tilde{\tau}_1\alpha_1)\mu_2 \\ \alpha_1\mu_1 + \alpha_2\mu_2 \end{bmatrix} \tag{15}$$

where we used the shorthand α_i for $\frac{t}{1-t} - \delta_i$.

Non-myopic version

A further step is to avoid the assumption of myopic agents. We do this in the next step, where the structure of the matrix gets slightly more complex to take into account the agents' responses in anticipation of population shifts.

$$\begin{bmatrix} \hat{N}_1 \\ \hat{N}_2 \\ \hat{\tau} \end{bmatrix} = \begin{bmatrix} \beta^{-1} - \frac{\gamma_1}{\gamma} & -\frac{\gamma_2}{\gamma} & \frac{\tau}{1-\tau} - \delta_1 \\ -\frac{\gamma_1}{\gamma} & \beta^{-1} - \frac{\gamma_2}{\gamma} & \frac{\tau}{1-\tau} - \delta_2 \\ \tilde{\tau}_1 & \tilde{\tau}_2 & -1 \end{bmatrix}^{-1} \begin{bmatrix} \mu_1 \\ \mu_2 \\ 0 \end{bmatrix} \hat{C}_j \tag{16}$$

Multiclass version

In the case with e.g. 4 income classes the form of the matrix immediately generalizes:

$$\begin{bmatrix} \beta^{-1} - \frac{\gamma_1}{\gamma} & -\frac{\gamma_2}{\gamma} & -\frac{\gamma_3}{\gamma} & -\frac{\gamma_4}{\gamma} & \frac{\tau}{1-\tau} - \delta_1 \\ -\frac{\gamma_1}{\gamma} & \beta^{-1} - \frac{\gamma_2}{\gamma} & -\frac{\gamma_3}{\gamma} & -\frac{\gamma_4}{\gamma} & \frac{\tau}{1-\tau} - \delta_2 \\ -\frac{\gamma_1}{\gamma} & -\frac{\gamma_2}{\gamma} & \beta^{-1} - \frac{\gamma_3}{\gamma} & -\frac{\gamma_4}{\gamma} & \frac{\tau}{1-\tau} - \delta_3 \\ -\frac{\gamma_1}{\gamma} & -\frac{\gamma_2}{\gamma} & -\frac{\gamma_3}{\gamma} & \beta^{-1} - \frac{\gamma_4}{\gamma} & \frac{\tau}{1-\tau} - \delta_4 \\ \tilde{\tau}_1 & \tilde{\tau}_2 & \tilde{\tau}_3 & \tilde{\tau}_4 & -1 \end{bmatrix} \tag{17}$$

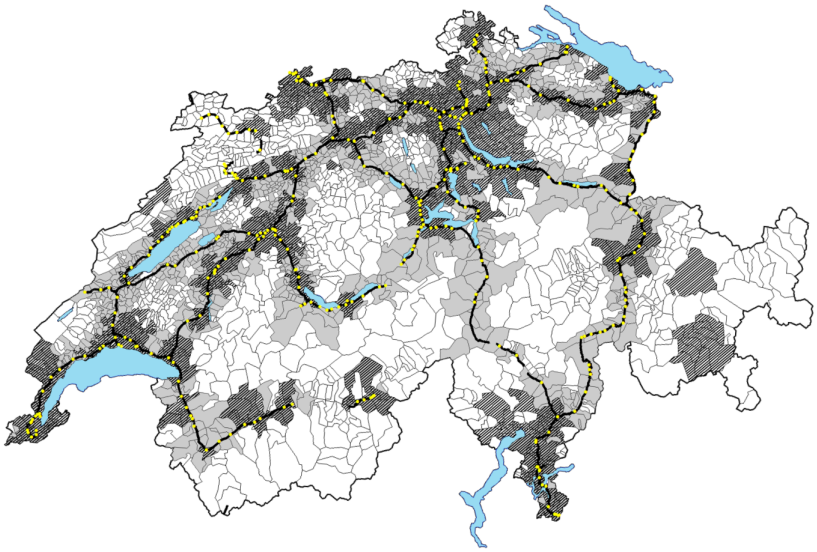
In the case with two income classes, it is possible to use the relationship $\sum_m s_{jm}\lambda_m = 1$ to have only one λ. Now it is not possible anymore so we assume that the political weights are in a linear relationship: $\lambda_k = \lambda k$, which means that empirically we can identify the slope λ. Plugging this in the expression for $\tilde{\tau}_1$:

$$\tilde{\tau}_1 = \frac{\lambda}{\tau} s_1 \left[\tau_1^*(1 - s_1) - 2\tau_2^* s_2 - 3\tau_3^* s_3 - 4\tau_4^* s_4\right] \tag{18}$$

As before, the $\tilde{\tau}_m$'s satisfy the relationship $\sum_m \tilde{\tau}_m = 0$

3 Reduced-form empirical moments

The Swiss highway system was planned in the 1950s to connect Switzerland's biggest cities, and as a consequence a number of small municipalities out of metropolitan areas got connected by accident along the way. We focus on the last group, where we can consider the timing of the opening of new



Note: Shaded areas denote municipalities that are part of a metropolitan area as defined by the Federal Statistical Office in 2000. Municipalities with highway access within 10km road distance are in light gray. Access points open at a given year are in yellow.

Figure 1: Map showing the Swiss highway network, the position of highway access points and the subdivision of the territory in municipalities and metropolitan areas.

ramps of access to the system as exogenous.

Identification strategy

Our identification employs the strategy used in Fretz, Parchet and Robert-Nicoud (2017), exploiting the spatial variation and the long panel dimension of the data. We restrict the sample to municipalities that get a highway access over our observation period and exploit the heterogeneity in the opening time of the access. Thus, all municipalities in our sample are eventually treated and are all located within the same distance to the next highway, increaseing the homogeneity of the sample and minimizing a potential source of bias, which arises if heterogeneous municipalities follow heterogeneous growth paths. We include year dummies to control for country-level shocks. Finally, as the highway system was built with the intent of connecting the main cities and urban areas, we eliminate all municipalities in metropolitan areas to avoid this source of selection bias.

Empirical results: seemingly unrelated regressions

Our main econometric model for the reduced-form results is the following system of seemingly-unrelated regressions (where $i = 1, \ldots, 4$ for the first equations involving population shares), with municipality- and year-fixed effects:

$$\log(N_{ijt}) = \gamma_i \text{Acc}_{jt} + \sum_{s=1}^{10} \beta_{is}(\text{Acc}_{j,t-s} - \text{Acc}_{jt}) + \alpha_j + \rho_t + \epsilon_{ijt}$$
$$\log(\tau_{jt}) = \gamma_5 \text{Acc}_{jt} + \sum_{s=1}^{10} \beta_{3s}(\text{Acc}_{j,t-s} - \text{Acc}_{jt}) + \alpha_j + \rho_t + \epsilon_{5jt}$$

where the τ's are consolidated cantonal+municipal+ church tax rate on a married households without children belonging to the top 10% of the nation-wide income distribution. Access is a dummy variable that turns 1 when a highway access is opened within a road distance of 10km, and stay 1 for all subsequent years; N is the number of taxpaying households in different income percentiles, calculated using the individual taxpayers data for the years 1973 onwards and aggregated data for all years before. The time spans the period $t = 1947, \ldots, 2010$. The α_j's are municipal fixed effects, while the ρ_t's are time fixed effects. Finally, all the error terms are potentially correlated across the five equations.

The results for the estimation of the system of equation are presented in Table 1. The first 4 columns reproduce the results obtained in Fretz, Parchet and Robert-Nicoud (2017). Population in all income classes increased with the highway connection disproportionally more in high-income classes, shifting the income distribution to the right. The last column shows the decrease in the municipal tax rate.

	Top10	75-90	50-75	Bottom50	Tax
Long-term effect ($\hat{\gamma}$)	.133***	.170***	.118***	.043***	−.034***
	(.033)	(.018)	(.013)	(.009)	(.005)

Standard deviations in parentheses - Homoskedastic standard errors

Table 1: Results from the SUREGs

Connecting theory and empirics

In section 2 we obtained the theoretical moments of the model, which depend on:

	Parameter value
β	2
δ_1	0.3
δ_2	0.4
δ_3	0.5
δ_4	0.6

Table 2: Table of calibrated parameters.

- Calibrated parameters (β, δ_m, etc), taken from other studies (e.g. from Brülhart, Parchet, Danton and Schläpfer (2017))
- Parameters that we want to estimate: λ and the μ_m's

These moments should match the coefficients γ_i of the reduced-form, seemingly-unrelated regressions showed in section 2.

The numerical algorithm that we use, the Classical Minimal Distance (CMD) estimator, minimizes the distance in 5-dimensional space of the theoretical moments (vector function of 5 parameters) from the empirical moments. Formally, we minimize the quantity:

$$(\boldsymbol{m} - \boldsymbol{\gamma})' V^{-1} (\boldsymbol{m} - \boldsymbol{\gamma})$$

where V is the variance matrix from the SUREGs, an essential piece of the analysis which makes it important to run the regressions in an inter-dependent format. In this analysis we also exclude cities and municipalities that never get connection and, as said, the error terms can be correlated across equations.

4 Preliminary results and further steps

Empirical results: structural estimation

By minimizing the distance between theoretical and empirical moments, we can get an estimate of the five parameters of interest: λ, the political weight of individuals belonging to different income classes, and the four μ's, the private costs of commuting in the two classes. Preliminary results are presented below. As can be seen, we obtain a positive value for the parameter λ, which corresponds

for a bigger political weight for high income classes. We are now working on improving this analysis obtaining suitable standard errors for these estimates, and repeating it in the new setting with a meaningful network notion of connectedness, which would allow a better interpretation of the parameters μ_m.

	λ	μ_1	μ_2	μ_3	μ_4
Parameter estimate	.337	.058	.030	.050	.083

Table 3: Structural parameter estimation

Instead of the simple definition of connectedness used before, we organized spatial data on the Swiss road and highway network, and its evolution year by year since the 1950's. This will enable the study of important network effects, which can result in dramatically heterogeneous gains for connected municipalities. At this stage, we are computing, for every year in our database, a so-called origin-destination cost matrix, which contains the road distance and driving time between each pair of Swiss municipalities. We are then going to summarize the information for each municipality in a single measure of market access, a continuous variable which will take the place of our previous dummy. We plan to reproduce all reduced-form results in this new setting. In this way, we can also connect our empirics to the theoretical model better, being now able to measure the shocks to commuting time which are part of the model.

References

[1] Torsten Persson and Guido Tabellini. *Political Economics - Explaining Economic Policy.* The MIT press (2000)

[2] Stephen J. Redding and Esteban Rossi-Hansberg. *Quantitative Spatial Economics.* Annual Review of Economics Vol. 9:21-58 (August 2017)

[3] Stephan Fretz, Raphaël Parchet and Frederic Robert-Nicoud. *Highways, Market Access and Spatial Sorting.* DP12437, 2018.

[4] Attila Ambrus, Ben Greiner, Parag A. Pathak. *How individual preferences are aggregated in groups: An experimental study.* Journal of Public Economics, 1-120 (September 2015)

[5] Marius Brülhart, Raphaël Parchet, Jayson Danton, Jörg Schläpfer. *Who Bears the Burden of Local Income Taxes?*. Working paper 2018