Public Attitudes Towards the Role of the State and the Private Provision of Training: Evidence from the Swiss Apprenticeship System

Andreas Kuhn, Swiss Federal Institute for Vocational Education and Training, University of Bern, and IZA*

Jürg Schweri, Swiss Federal Institute for Vocational Education and Training,

Stefan C. Wolter, University of Bern, Swiss Coordination Centre for Research in Education, CESifo and IZA

PRELIMINARY DRAFT – PLEASE DO NOT CIRCULATE

June 2018

Abstract

Existing empirical research shows that a substantive fraction of Swiss training firms is willing to incur considerable net training costs. One potential explanation for such behavior is that firms act in accordance with the norms and expectations they are faced with in the local labor market they are operating in. In the research described in this paper we ask whether the norm towards the private, rather than the public, provision of public goods influences the probability that a firm is willing to offer apprenticeship positions. In line with our hypothesis, we find that the training incidence is higher in communities which are characterized by a stronger norm towards the private, rather than the public, provision of public goods, which we measure using local results from several national-level votes which explicitly dealt with the proper role of the state in the context of the apprenticeship system. This finding turns out to be robust to a series of robustness checks as well as to an instrumental-variable strategy that tackles the issue of potential endogeneity of public attitudes.

JEL classification: D22; D63; H41; I22; J24
Keywords: public goods; private provision of training; social norms; normative attitudes towards the role of the state; vocational education and training; apprenticeship training

---

*We thank Manuel Aepli, Uschi Backes-Gellner, Giuliano Bonoli, Jürg Markowitsch, Sandra McNally and Ludger Woessmann, seminar participants in Bern, as well as participants at the 2017 CVER conference at the LSE, the 2017 INAP conference in Washington, and at the 2017 VET Congress at SFIVET for many helpful comments and constructive suggestions. We also thank Mirjam Strupler for compiling the survey data on the costs and benefits of apprenticeship training for the purpose of this study; Romain Douard, Stéphane Henriod and Dominik Ullmann from the Swiss Federal Statistical Office for helping us with various data issues; Manuel Aepli, Marcel Hageman and Marianne Müller for great research assistance. Financial support by the State Secretary for Education, Research and Innovation (SERI) is gratefully acknowledged (contract number 1315000817).
Corresponding author: Andreas Kuhn, Swiss Federal Institute for Vocational Education and Training, Kirchlindachstrasse 79, 3052 Zollikofen, Switzerland, andreas.kuhn@ehb.swiss.

1 Introduction

> *"No matter how cleverly designed (...), incentives alone cannot provide the foundations of good governance."* Bowles (2016, p.2)

The private provision of training by firms looks back on a centuries-old tradition. Yet, the numbers and shares of firms that provide training vary markedly between countries: while firm-based apprenticeships prevail in German-speaking and some other European countries, they are nowadays less important in Anglo-Saxon countries and almost inexistent in Latin countries. Even within countries, firms' training incidence varies substantially. Since these regional patterns are hard to explain by classical profit-maximizing behavior, we investigate the role of social norms in favor of the private provision of training.

The recent empirical literature on social norms has focused on the effects of individuals' beliefs and norms on their own behavior and that of others. The effects of norms on firm behavior have been less investigated. More specifically, we are not aware of any empirical study which has tried to estimate the effect of norms on the private provision of education or training.

In this paper, we show that local social norms favoring the private provision of public goods increase local firms' training incidence in Switzerland. Such a norm may be internalized by firms, or enforced by consumers that favor training firms and sanction non-training firms. In line with this argument, Swiss firms actively communicate their training efforts, e.g. by placing newspaper and online ads in which they congratulate their apprentices for successfully passing their final exams.

Switzerland is particularly well suited to study the effect of social norms on firms' provision of apprenticeship training. It is the country with the highest share of youth attending firm-based apprenticeships (OECD, 2008). A first fact calling for explanation is that this education and training system works without obligation for firms to provide training places. A second fact calling for explanation is that firms' training incidence differs substantially between regions. Training incidence rises from the eastern to the western part of the country. The lowest share of training firms is found in the French-speaking region in Western Switzerland, but the regional variation is high even within the French-speaking and the German-speaking region. We postulate that social norms play an essential role in securing a sufficient number of training places as well as in explaining the regional patterns of training incidence.

To identify the effect of social norms on firms' training behavior, we profit from the system of direct democracy in Switzerland. Swiss citizens' are asked to express their preferences on federal laws and amendments to the Swiss constitution at the voting booth several times a year. We use municipality-level results from national popular votes on the public or private provision of apprenticeships. Such voting results are particularly well suited to measure social norms as citizens are likely to reveal their true attitudes on the issue at stake. Furthermore, the voting outcomes are reliably collected by the authorities and available in high regional detail. In 1986 and 2003, there were two popular initiatives that asked for an amendment to the constitution stipulating a stronger involvement of the state in the provision of training places. These votes provide us with unique regional measures of people's preferences on the public or private provision of training and thus their expectations towards state and firms to provide training places.

The data on firms' willingness to provide apprenticeship places comes from three national surveys on firms' training costs and benefits carried out in 2000, 2004 and 2007. The surveys contain detailed information about firms' training behavior and their characteristics. Pooling the data sets allows us to exploit regional variation in training patterns and match them to voting outcomes on municipality level. An additional advantage of the cost-benefit data set is that it allows us to disentangle the role of monetary benefits to training from the role of social norms.

The main result of our analysis is that regional norms on the private provision of public goods are highly correlated with firms' training incidence and that this correlation is robust to different specifications. Controlling for firm and regional characteristics does not change the results. However, a major concern is that the social norm for the private provision of training might be stronger in regions where firms' training incidence is already high. This would result in a reverse causality problem where social norms themselves are influenced by firm behavior, interfering with the hypothesized influence of social norms on firm behavior. Therefore, we instrument voting outcomes on new training laws with other voting outcomes that also deal with the public or private provision of public goods, but not with apprenticeship training. Using such votes, e.g. on the public or private provision of health insurance, as instruments for the social norm confirms our findings and results in a significant and sizeable effect of the norm on

firms' training incidence.

Our work contributes to a growing body of literature on the effects of social norms on the private provision of public goods. There is ample experimental evidence suggesting that internalized norms of cooperation are able to sustain a high level of cooperation between individuals in settings where there are strong countervailing incentives (Chaudhuri, 2011; Fehr and Fischbacher, 2004; Ledyard, 1995; Ostrom, 2000). However, as laboratory experiments necessarily operate in an artificial setup, it has been argued that the results from such experiments may not carry over to real contexts. Experiments trying to elicit social preferences and norms may be particularly affected by such problems. Levitt and List (2007) argue that "real-world markets operate in ways that make pro-social behavior much less likely (p. 168)", e.g. because actors on markets are more anonymous than in an experimental lab situation. Several empirical studies found that individuals' behavior in laboratory experiments can deviate substantially from their behavior in the field (e.g. Benz and Meier, 2008; Frey and Meier, 2004; Laury and Taylor, 2008). Evidence from the field on the importance of social norms for contributions to public goods by individuals is still rare and restricted to applications such as charitable giving (Shang and Croson, 2009), tipping (Azar, 2004), labor effort under different pay schemes (Bandiera *et al.*, 2005), or investments into "sin stocks" (Hong and Kacperczyk, 2009). Evidence on the importance of social norms for contributions to public goods by firms is even rarer and focusses mainly on corporate social responsibility (Schmitz and Schrader, 2015). An exception is the paper by Bassanini *et al.* (2017), who investigate the effects of local social pressure and show that firms dismiss fewer workers in secondary establishments that are closer to the headquarters.

The remainder of this paper is structured as follows. The following section presents some background information on the institutional setting, focusing on the key features of the Swiss apprenticeship system. We will push the argument that the system represents a rare example of community governance in the educational context. Section 3 discusses the different data sources used in our empirical analysis, focusing on the data containing information about firms' training behavior and on the measurement of the norm towards the private provision of public goods and, more generally, the role of the state. Our econometric framework is discussed in section 4, and the resulting estimates are presented and discussed in section 5. In that section, we also provide a series of robustness checks and we test several secondary hypotheses to further

strengthen the credibility of our main estimates. Section finally 6 concludes.

2 The Swiss apprenticeship system

This section provides some background information on the Swiss educational system and the institutional setting of the Swiss apprenticeship system.[1] The main message that we push here is that the Swiss apprenticeship system has several unique features that make it a rare, yet at the same time very successful, real-world example of community governance in the realm of education and training.

2.1 General education and vocational education and training at the upper-secondary level

The Swiss educational system is, first and foremost, characterized by it's exceptionally strong emphasis on vocational education and training (VET, henceforth) at both the upper-secondary and tertiary level. After completion of mandatory schooling, about 64% of the most recent cohorts of adolescents eventually enter some kind of apprenticeship training (SERI, 2014). The remainder mainly chooses further general education (taught at a "Gymnasium") that prepares for, and grants access to, university studies. Thus, in Switzerland, VET is by far the most often chosen educational track at the upper secondary level.

Among those entering some kind of apprenticeship training, the most frequent choice by far is to enter a firm-based apprenticeship program lasting from two to up to four years, depending on the occupation learned.[2] During their training, apprentices spend most of their time in their training firm, where they are involved in both practical exercises and actual work from the start of their apprenticeship. In addition, apprentices spend one or two days per week in vocational school, where they acquire both occupation-specific knowledge as well as general human capital (such as native and foreign languages). Apprentices get paid by their employer, but their wages are considerably lower than those of fully-trained workers in the same occupation, even taking

[1] More information about the Swiss VET system, and how it fits into Switzerland's educational system as a whole, is available in Wettstein *et al.* (2017). See also Wolter and Ryan (2011) for a more general discussion of apprenticeship training beyond the Swiss case.

[2] Among these, about 91% enter a dual apprenticeship which combines practical training in a firm with vocational school. The remainder (about 9%) attends full-time school-based VET programme.

their lower productivity into account, which implies that apprentices share the costs of training with their employers.

2.2 Key features of the Swiss apprenticeship system

Voluntary participation of both employers and apprentices

The most obvious feature of the Swiss apprenticeship system is that it is based on the voluntary participation not only of apprentices but also of employers. Indeed, there is no direct regulation of the number of apprenticeship positions, except perhaps that public employers may also train apprentices (e.g. hospitals training nurses), and, in fact, there is no explicit regulation of wages paid to apprentices neither.[3] Thus there essentially exists a market for apprentices, regulated to a large extent by the supply of and demand for apprentices.

Moreover, various formal associations and informal cooperations among (training) firms within the same occupation or industry (e.g. Agell, 1999; Busemeyer and Trampusch, 2011) play a key role in the Swiss apprenticeship system. Indeed, employers and their associations ("Organisationen der Arbeitswelt") are not only responsible for the (further) development of the training curricula, they are also responsible for the preparation and the implementation of the final practical examinations, which are decisive for obtaining the diploma, and they can even call for a change in the duration of an apprenticeship or the introduction of a new learnable occupation (e.g. because technological innovations change the demand for competencies and skills on the labor market).

The financing of firm-based vocational education and training

Another distinguishing feature of the Swiss apprenticeship system, and one closely related to the voluntary participation of both employers and apprentices, is that the the costs accruing from apprenticeship training (within the firm) is almost fully borne by the firms actually providing the training positions as well as by the apprentices. Because training within the firms is organized privately, there are no official or comprehensive statistics on the costs of the training provided by employers. However, approximate estimates suggest that employers incur about

[3]There is the possibility of implementing legally binding, usually sector specific, funds ("Berufsbildungsfonds") that collect payments from all employers to support those employers providing training positions. However, the initiation of such funds must come from the employers and their associations, not from the legislator.

2.7 billions of direct training costs per year.[4]

In contrast, however, vocational schooling is almost fully funded publicly (by both the federal government as well as by the cantons). The costs of vocational schooling amount to about 2.5 billions per year, according to official statistics (SERI, 2014). The total public costs of VET at the upper secondary level, which mainly includes the costs for vocational schooling but also other costs (e.g. costs for the final practical exams), amount to almost 3 billions Swiss francs per year. Thus approximately 47.5% (= 100% · 2.7/(2.7+3)) of the annual overall costs of the apprenticeship system are born by the employers actually providing the training positions.[5]

Specificity of firm-based vocational training, external certification of training, and poaching

One might argue that the setting just described should imply that the training provided must be specific to the training firm to a significant degree. However, empirical evidence suggests that a substantial part of the human capital aquired through apprenticeship training is Switzerland is transferable across firms – and often even across different occupations (e.g. Mueller and Schweri, 2015). Moreover, Switzerland's labor market is comparatively unregulated and flexible, undermining the argument that imperfections on the labor market may explain the high ratio of training firms in Switzerland (e.g. Muehlemann *et al.*, 2013).[6]

Further, indirect evidence on the transferability of the competencies acquired through apprenticeship is given by the observation that apprentices get poached by other employers once they have completed their training (e.g. Muehlemann and Wolter, 2011). If the competencies acquired during apprenticeship training were fully or mainly firm-specific, however, we should not observe such behavior on the labor market.

In addition, there are several institutional features in place that explicity aim at ensuring that mobility across employers is possible for apprentices after the completion of their training

[4]This estimate is taken from the Swiss federal office of statistics and is based on the same firm-level survey data that we use in this paper (see section 3 for details).

[5]These numbers neglect the costs born by the apprentices by accepting low wages (simply because there are no estimates available). To put these numbers into perspective, note that public spending on mandatory schooling (tertiary-level education, research) amounts to about 16.3 (8.5, 3.7) billions Swiss francs per year, respectively (figures are taken from the Federal Statistical Office, FSO, and relate to the year 2015).

[6]Consistent with this, comparisons between Switzerland and Germany (e.g. Muehlemann *et al.*, 2010) and between Switzerland and Austria (Moretti *et al.*, 2017) argue that the observed differences in the net benefits to employers from training apprentices is partly explained by corresponding differences in labor market regulation.

(for example, there are centralized examinations at the end of the apprenticeship and there is an external certification of the competencies acquired during the apprenticeship by the federal administration; cf. Acemoglu and Pischke (2000)).

Short-run benefits of apprenticeship training to the training firm

A final key feature of the Swiss apprenticeship system is that there are not only considerable short-run costs from apprenticeship training, but also often substantial monetarized gains to the training firm, as discussed in considerable detail in, for example, Strupler and Wolter (2012). Employers may benefit from training apprentices because apprentices, at least towards the end of their training, are able to perform skilled work (i.e. work that needs otherwise to be done by a trained worker) to a lower cost than when performed by a fully trained worker.

Indeed, one of the main results of the empirical literature on the costs and benefits of apprenticeship training in Switzerland is that a large fraction of the training firms (about two-thirds in the year 2009; but this share varies quite a lot over time) is able to realize a net benefit from training apprentices within the training period, the sometimes high costs of training notwithstanding (in the year 2009, for example, training costs per apprenticeship averaged almost 90,000 Swiss francs; which is considerably higher than the annual wage of an average worker in that year).[7]

At the same time, however, many training firms incur substantial net costs from training apprentices.[8] Moreover, even if a training firm covers its costs until the end of training, it can not be sure about that at the time they start a new apprenticeship because there is considerable uncertainty in both the costs and benefits of training from an ex-ante point of view. This can be inferred from the large variation of net benefits observed within the same training occupation. Indeed, with a few exceptions, there is a substantive fraction of firms realizing net costs from training in apprentices that are charcterized by on average positive net benefits.

[7]It has further been shown that training firms may save recruiting costs that they would otherwise have to spend if they (are able) retain apprentices that have completed their training (e.g. Blatter *et al.*, 2012). Relatedly, apprenticeship training may also serve as a (costly) screening device for employers (Mohrenweiser *et al.*, 2017).

[8]Net benefits are typically negative for the more technical and the more demanding apprenticeships (e.g. electrician or polymechanic).

3 Data and key variables

In this section we discuss the different sources of data used in the empirical analysis and how we will measure the local norm towards the public (versus the private) provision of apprenticeship positions using national-level voting results. We also present a short summary discussion of the spatial structure of the final data set at the end of this section.

3.1 Firm-level survey data

Our first and most important data source are three consecutive surveys that were specifically designed to elicit detailed information about the costs and the monetarized benefits of apprenticeship training from the point of view of the employers. The first of these surveys was administered in the year 2000, the second in 2004, and the third in 2009 (see Strupler and Wolter, 2012, for details and many additional references). Taken together, the three surveys cover more than 21,000 firm-level observations, containing both training and non-training firms. Moreover, the sample of firms is representative of almost the entire population of firms in Switzerland.[9] While it is possible that the same firm appears more than once in the combined data because the same firm has been sampled in more than one wave of the survey, it is not possible for us to follow the firms across time for reasons of data protection. Because we will use clustered standard errors throughout, however, this feature of the data is taken into account with regard to statistical inference.

Because all three surveys cover both training and non-training firms, and because we know wheter a specific firm currently trains apprentices or not, the data can be used to model the incidence of apprenticeship training – which is our variable of main substantive interest. Moreover, the data cover not only detailed additional variables related to the costs and benefits of apprenticeship training, but also emplyoers' assessment of their motives to offer apprenticeship positions (see Muehlemann and Wolter, 2014, for an overview). The richness of information available in the data allows us to test a whole series of ancillary hypotheses, e.g. implementing an empirical test on norm internalization by the employers (see section 5.6 below for details).

[9] In all three years of the survey, each cross-section of firms is representative of the universe of all firms in the year of the corresponding survey, except for the very smallest firms as well as employers from the primary sector (which have been excluded from the sampling frame in all three surveys). Additional details on the sampling procedure, for the most recent wave of the survey, are given in Potterat (2011).

Finally, a key feature of the survey data is that they contain the postal code indicating the physical address of the firms, which allows us to merge data from other sources, such as community-level voting results or additional variables from the census or the busniess census (see section 3.4 below for details).

3.2 Community-level voting results

As one of the main pillars of the direct-democratic political system of Switzerland, citizens are regularly asked to cast their vote on various policy topics, such as environmental policy, gender issues or, of course, educational policy. Votes take place both at the national and the subnational level (i.e. at the cantonal and the communal level), depending on the level(s) at which the corresponding legislation takes place. As mentioned in section 2 above, the VET system is regulated at the national level in Switzerland, in contrast to most other educational domains which are regulated at either the cantonal and/or the communal level. This opens up the possibility to use national-level voting results related to VET policy to measure individuals' attitudes towards the role of the state in a consistent way across whole Switzerland and, ultimately, the level of civic virtue prevailing within a given community.

Using voting results to measure normative attitudes towards the role of the state

Quite undisputedly, voting results are a direct measure of voters' attitudes towards specific policy issues. We believe that the use of the voting results has some distinct advantages compared to the use of attitudinal survey data.[10] A first important advantage of the use of voting results is that the outcome of a given vote usually has real consequences, and thus voters have a strong incentive to reveal their true preferences. In contrast, corresponding survey questions necessarily remain hypothetical, providing virtually no incentive at all to respondents to reveal their true attitudes. Moreover, because voting is strictly anonymous, there is no pressure towards expressing socially desirable opinions (e.g. Bertrand and Mullainathan, 2001). Therefore, focusing on those votes that dealt specifically with the question of whether the state or private actors should take responsibility, we believe that we are able to measure public

[10]The main downside of using the voting results to measure attitudes is that there are no individual-level data available (up to now there is no electronic voting in Switzerland, thus voting is strictly anonymous, and the communities only record aggregate-level results). However, since we are only interested in the aggregated data anyway, this restriction has no bearing at all for our analysis.

attitudes towards the role of the state in a convincing yet very straightforward way.

Moreover, at the aggregate level (e.g. at the community level), we will interpret differences in the voting results across communities as reflecting spatial differences in civic virtue, which we understand as a local norm towards the private – rather than public – provision of training in the context of our study.[11] Indeed, note that aggregate-level voting results fulfill the two conditions noted by Brennan *et al.* (2013) for the existence of a (social) norm. First, and obviously, a significant fraction of individuals must hold a certain normative attitude towards a given subject. Second, it must also be the case that a significant fraction of individuals is aware that there exists a normative attitudes that is shared among a significant part of their community. In other words, people within a community must be aware that there exists a shared norm in the community they live in. The voting data that we use to measure [xxx] in our empirical analysis (discussed in detail in the following subsection) fit this definition almost perfectly – they directly measure the fraction of people sharing a given normative attitude on a specific subject and, after the vote has taken place, the result of the vote is public knowledge because the results are discussed in the media and published in national and/or local newspapers, implying that the strength of the norm within a given community becomes evident to the members of a community, as well as everyone else.

Implicitly, we also have to assume that we use data that are aggregated at the "correct" spatial level, i.e. the level at which the voting data are aggregated should reflect the level at which social norms are expected to have an effect on individual behavior. We believe that the spatial units used in our empirical analysis are small enough that we can plausibly expect social norms to be effective within these units (cf. section 3.4 below).

A further advantage of the voting data is that they are virtually complete, i.e. votes represent kind of a full census of attitudes on a specific subject among voters, which allows us to measure mean attitudes even for scarcely populated communities; something that would not be possible with usual attitudinal survey data.[12]

[11] Community-level voting results have been used before in various contexts to measure cultural and/or social norms. For example, Stutzer and Lalive (2004) and Eugster *et al.* (2017) use regional voting results to measure work attitudes, while Lalive and Stutzer (2010) and Janssen *et al.* (2016) use them to measure the local norm towards gender equality.

[12] At the same time, however, voting results do not necessarily represent attitudes among the whole local population. First, participation rates are usually far below 100% (cf. table 1), potentially inducing a bias due to selective participation (though one may argue that those not willing to participate do not care about the outcome of the vote). Perhaps more importantly, however, many individuals are not allowed to vote because

Votes about the allocation of responsibilities within the VET system

Based on the above considerations, we therefore use community-level results from several national-level votes in our empirical analysis. Most importantly, there were two votes that directly touched the issue of private versus public provision of vocational education and training and that were temporally close to the collection of the survey data. The first vote was a popular initiative ("Initiative für ein ausreichendes Berufsbildungsangebot"), held on May 18, 2003; the second vote was also a popular initiative ("Initiative für eine gesicherte Berufsbildung und Umschulung"), held on September 28, 1986. Both initiatives aimed at an increase in the public involvement regarding the provision of vocational education and training, and both initiatives were finally rejected by a majority of the votes. From a substantive point of view, note that both initiatives demanded a shift away from private towards more public responsibility within the Swiss apprenticeship system.

Table 1

Panel (a) of table 1 lists a few key figures for the two votes. Both initiatives were clearly rejected in the end, with only a minority of all valid votes in support of the respective initiative: the 1986 vote gained only 18.3% of all valid votes in it's support, the 2003 vote captured about 31.6% of all votes cast.

Figure 1

As illustrated in figure 1, however, there was considerable variation in the share of votes in favor of each of the two initiatives across different communities. Community-level vote shares from vote nr. 340 (nr. 503) vary between 0% and about 95% (between 0% and 79%). Not surprisingly, mean vote shares are somewhat less spread out, but there is still a large amount of variation, with mean vote shares varying between a low of about 6.5% to a maximum of about 64%.

Panel (b) in the middle of table 1 lists three additional votes on educational policy; two of these votes also dealth directly with the regulation of VET, while the third one dealt with

they lack Swiss citizenship. To take these two issues into account, we will include the mean turnout across the two votes as well as the fraction of foreigners within a community as control variables in most of the regressions presented below.

higher education at the university level.[13] Note that these three votes took place in the more distant past than the two votes from panel (a), thus they present kind of a "historical" measure of attitudes towards the role of the state within the context of VET and educational policy. In the empirical analyisis below we will use the mean vote share among these three votes to check the internal validity of our main attitudinal measure as well as an instrument for current attitudes towards public responsibility in VET. As above, there is considerable variation in the voting results, both for the single votes and the mean share across the three votes. Specifically, the mean share of supporting votes across the three votes varies between about 28% and 76% (as illustrated graphically in appendix figure A.1).

Votes about the role of the state beyond educational policy

Finally, panel (c) of table 1 lists three additional votes that dealt with the provision of public goods or the role of the state more generally (i.e. these votes are concerned with issues outside the realm of educational policy). Specifically, the table includes the results from two votes on public health insurance and one vote which asked for the introduction of a female quota within the federal administration. While two of these votes were clearly rejected, the vote on the introduction of a mandatory health insurance (vote nr. 415) was accepted with a close majority of the votes (51.8%) in its favor. Each of the three votes demanded more responsibilities for the state. Consequently, we will use the results from these additional votes to construct a measure of attitudes towards the role of the state in the non-educational context (see also section **??** below). Again, there is considerable variation in the mean vote share across communities, with values ranging from a minimum of about 6% to a maximum of 66% (cf. appendix figure A.1).

3.3 Community characteristics from the Swiss census and the Swiss business census

In addition, we use selected data from the Swiss census ("Volkszählung") and the Swiss business census ("Betriebszählung") to construct a couple of regional-level characteristics. These variables will be used as control variables in the empirical analysis below, at different levels of regional aggregation (e.g. at the community level or at the level of local labor markets).

[13] We also include the third vote because it was more contentious than the other two votes, thereby inducing additional variation in the mean vote share.

More specifically, we use data (mainly) from the 2000 Swiss census to construct a variety of control variables that describe the composition of the population living within a given community. We further use data from the Swiss business census, mainly from the year 2008, to construct complementary measures describing the structure of economic activity. Additional details regarding these variables are given in section 5 below.

3.4 Spatial structure of the data

In the main part of the empirical analysis, our basic unit of observation is always the individual firm. For each firm i in the pooled survey data we know the postal code of the firm, extracted from the information used to contact the firms in the course of the survey, and we can use this information to identifiy the community j where a specific firm i is located in.[14]

This regional information with regard to the physical location of a firm in turn is key for our empirical analysis because it allows us to merge the firm-level survey data with aggregate-level information derived from the voting results and with variables constructed from either the census data or the business census data.[15] Further, given that we know the political community a firm is located in, it is easy to derive additional spatial information. For example, in the context of apprenticeship training, it is relevant to control for institutional differences across the cantons because educationaly policy is, to a large extent, under the supervision of the cantons, as discussed in section 2 above.

As a final remark, it is worth pointing out that Switzerland is a very small-sized country as a whole. Subnational entities, such as communities, are therefore small in size as well, both with regard to area and number of inhabitants (cf. appendix table A.1). This makes us confident that the spatial units for which we observe variation in the voting results are actually small enough such that local norms can plausibly take effect on individual behavior; measuring social norms at a too broad level should attenuate any effect of the norm on employer behavior towards zero (to be sure that the aggregation level has no impact on our results we estimate

[14]Postal codes can be mapped to community numbers, even though there is no one-to-one correspondence between postal codes and communities (there is a table of correspondence provided by the Federal Statistical Office). We map postal codes to community numbers because most additional data, such as the voting results, are only available at the community-level (e.g. our key regressor). As shown in appendix table A.1, there are more postal codes than communities (mainly because there are several postal codes within the larger cities).

[15]An important issue that we have to take into account is that the structure of the communities constantly changes over time (most importantly, the total number of communities has significantly decreased over time as more and more communities have merged with each other).

our baseline specification using different aggregation levels for the key regressor; see section 5.2 below).

4 Econometric framework

Our empirical analysis will basically proceed in three consecutive steps. In a first step we start with some regular OLS regressions in which social norms are treated as exogenous, but where we try to control for as many potential confounders as possible. In a second step we try to tackle the imminent issue of simultaneity between firms' training behavior and public attitudes towards the role of the state using various instruments. In a third and final step we we will provide a series of ancillary analyses with the aim of strengthening the credibility of our main estimates.

4.1 Baseline OLS estimates

Our baseline OLS regression models will take the following form:

$$T_i = \alpha + \beta N_{j[i]}^{\text{VET}} + \gamma F_i + \delta C_{j[i]} + \psi_{r[i]} + \phi_{t[i]} + \epsilon_i, \tag{1}$$

with the dependent variable T_i being a binary variable indicating whether firm i offers apprenticeship training or not (i.e. T_i equals 1 if firm i trains any apprentices, and 0 otherwise).[16] For the most part of the empirical analyisis T_i will be the dependent variable, but we will also have a look at some alternative outcomes in section 5.5 below (such as the number of apprentices). The regressor of primary interest is given by $N_{j[i]}^{\text{VET}}$, which denotes to the local share of votes supporting the private provision of training in community j in which firm i is located, and thus reflects public attitudes towards the private provision of vocational education and training (as discussed in detail in section 3 above). Parameter β is the main target of our empirical analysis because it quantifies the impact of social norms on individual firms' training behavior, at least under appropriate conditions. Because we hypothesize that stronger norms towards the private

[16] One may object that a nonlinear probability model would be more suitable for the data at hand (because of the binary nature of the dependent variable). We prefer to use the linear probability model because of its straightforward interpretation and because the comparison across OLS and instrumental-variable estimates is much easier. Nonetheless, we show average marginal effects from a probit model in column 10 of table 5 below.

provision of training are associated with firms being more likely to be involved in the training of apprentices, and because lower values on $N_{j[i]}^{\text{VET}}$ indicate a stronger norm towards the private provision of training, we expect β to be negative.

Given that the dependent variable is measured at the firm level and the main regressor at the community level, the most obvious confounding variables are either at the firm level or at the regional level. For example, there may be regional differences in firm-level characteristics across communities with different levels of civic virtue that are themselves predictive of whether an individual employer trains apprentices or not. All additional variables are therefore used as controls for potential confounders when estimating β, and are thus of no (or only minor) direct interest. Equation (1) distinguishes between F_i and $C_{j[i]}$ which denote, respectively, the inclusion of additional firm- and community-level controls. In most of our regression specifications, we will also include regional fixed-effects and survey-year fixed effects, denoted by $\psi_{r[i]}$ and $\phi_{t[i]}$, respectively. The regional fixed effects are potentially important because regional subentities in Switzerland have considerable impact on educational policy and thus, potentially, also on the probability that a given employer provides apprenticeship positions. Survey-year fixed effects in turn could be important if there are differences in the sampling frame and/or response behavior across the three different years of the survey.[17]

Finally note that our main regressor, $N_{j[i]}^{\text{VET}}$, varies at a higher level of regional variation than the dependent variable, potentially biasing conventional standard errors that ignore this specific feature of the data (e.g. Cameron and Miller, 2015; Moulton, 1986). We therefore report standard errors that are clustered at the community-level throughout the empirical analysis. Clustering at the community level also takes into account that we may observe the same firm in more than one wave of the survey (as discussed in section 3 above).

4.2 Tackling reverse causality

One remaining issue with the estimates based on equation (1) is that they do not take the potential simultaneity of local norms and employers' training behavior into account. That is, one might argue that there might not only be an effect of the local norm towards the private

[17]We basically treat our data as one large cross-section of firms, and we only use the survey-year fixed effects to allow for differences in the baseline probability of training across survey years. For that reason, we do not index the whole equation (1) against t, but only the survey-year fixed effects $\phi_{t[i]}$.

provision of training on the probability of training, but that there is a reverse effect as well (i.e. one might argue that, say, a high training probability strengthens individuals' belief that training is best provided privately). In the second step of our empirical analysis, we thus try to correct for potential simultaneity bias using different instrumental variables.

Our main idea is to use additional voting results as instrument(s) for our measure of civic virtue. More specifically, voting results that have no direct relation to VET, but that also relate to the issue of how responsibilities should be split between private and public actors, are likely to be correlated with corresponding attitudes in the context of VET and may therefore be used as an instrument. A similar idea is to use results from past votes on the distribution of responsibilities between private actors and the state within VET and the educational context more generally as an instrument for current attitudes towards the role of the state. Thus our first-stage regressions look as follows:

$$N_{j[i]}^{\text{VET}} = \pi_0 + \pi_1 N_{j[i]}^{\text{STATE}} + \pi_2 F_i + \pi_3 C_{j[i]} + \psi_{r[i]} + \phi_{t[i]} + \varepsilon_i, \quad \text{or} \tag{2a}$$

$$N_{j[i]}^{\text{VET}} = \pi_0 + \pi_1 N_{j[i]}^{\text{HIST}} + \pi_2 F_i + \pi_3 C_{j[i]} + \psi_{r[i]} + \phi_{t[i]} + \varepsilon_i, \tag{2b}$$

with $N_{j[i]}^{\text{VET}}$ denoting, as above, the strength of public attitudes towards the private provision of VET in a given community j. In a first specification we instrument $N_{j[i]}^{\text{VET}}$ with either $N_{j[i]}^{\text{HIST}}$ and/or with $N_{j[i]}^{\text{STATE}}$, representing public attitudes towards the role of the state outside educational policy and past attitudes towards the role of the state in the realm of VET, respectively. If $N_{j[i]}^{\text{STATE}}$ ($N_{j[i]}^{\text{HIST}}$, respectively) is a valid instrument for $N_{j[i]}^{\text{VET}}$, then π_1 must be significantly different from zero, which can easily be checked once the corresponding first-stage regression has been estimated.

We must also assume that the instrument does not suffer from the same problem as the endogenous variable. That is, we must assume that there is no reverse effect running from the local training incidence on neither $N_{j[i]}^{\text{STATE}}$ nor on $N_{j[i]}^{\text{HIST}}$. While we fear that $N_{j[i]}^{\text{STATE}}$ is somewhat susceptible to this problem, $N_{j[i]}^{\text{HIST}}$ is arguably much less so because it is measured many years before the endogenous variable. Finally, we must also assume that the instrument has only an indirect effect on firms' training behavior through its effect on $N_{j[i]}^{\text{VET}}$. While it is unlikely that this assumption holds unconditionally, for either of the two instruments, we believe that one can argue that this assumption holds once we condition on observable as well as on unobervable

factors (by including cantonal fixed effects).

5 Results

We next present our findings. Starting with some descriptive evidence, we move on to the estimates from both OLS and instrumental-variable estimates. After a short illustration of the quantitative implications of our main estimates, we then present a series of additional results and test a couple of ancillary hypotheses.

5.1 Descriptive evidence

Figure 2 illustrates how the training incidence among firms varies across regions within Switzerland. There is, quite obviously, considerable spatial variation in the training incidence across the different regions within Switzerland (note that, for the purpose of illustration, the figure plots data aggregated up to the level of districts). It is also remarkable that there appears to be systematic variation in the training incidence across regions. Specifically, the training incidence appears to be higher in the German-speaking part of Switzerland than in both the French- and Italian-speaking regions (which are, respectively, located in the Western and the Southern part of the country). It further appears that the regional training incidence among employers is higher in the more rural than in the urban areas.

Figure 2

Analogously, figure 3 shows how the norm towards the private, rather than the public, provision of vocational education and training varies across the different districts. This figure shows that there is pronounced spatial variation in public attitudes towards the role of the state as well. Even more interestingly, note that the variation in attitudes also follows a systematic spatial pattern. Specifically, the support for more public involvement in the provision of apprenticeship training is much stronger in the both the French- and Italian-speaking part of Switzerland than in the German-speaking regions, consistent with the findings of Eugster *et al.* (2011), for example.[18] Further, it appears that voters in the more urban regions have

[18]This opens up the possibility to estimate the effect of public attitudes towards the role of the state on employers' training behavior along the language border using a spatial regression discontinuity design (cf. Aepli *et al.*, 2018).

more favorable attitudes towards the role of state than those in the more rural areas. Overall, it thus appears that the pattern in figure 3 mirrors the one from figure 2 (with reverse sign, however).

Figure 3

Combined, the two figures apparently imply that we should find a pronouced association between the local norm towards the public provision of vocational education and training and the observed training incidence among firms. This is confirmed by figure 4, which plots the regional incidence of apprenticeship training (shown on the y-axis) and public attitudes towards the role of the state (shown on the x–axis). The figure shows that there is an obvious negative correlation between the local incidence of training and the mean vote share in favor of more public involvement in the provision of apprenticeship training. Thus, as expected, the probability of a firm offering apprenticeship positions is higher in those communities characterized by a stronger norm towards the private provision of training. Moreover, the association between the two variables turns out to be unambiguous, virtually linear and surprisingly strong, with an estimated correlation coefficient of about -0.61, based on data weighted by the number of firms within a region in the pooled sample.

Figure 4

Thus, in line with our main hypothesis, the raw data indeed suggest that part of the observed variation in the training incidence across regions can be explained by corresponding variation in public attitudes towards the role of the state. In the following section we will test whether this association turns out robust to the inclusion of additional control variables.

5.2 OLS estimates

Baseline estimates

Table 2 presents our first set of estimates of the effect of public attitudes towards the role of the state on the training incidence at the firm level.

Table 2

The point estimate of β in the first column of panel (a) is from a simple regression of T_i on the communal vote share in favor of private provision of vocational education and training, $N_{j[i]}^{\text{VET}}$, as described in section 3 above. This simple specification yields a point estimate of $\widehat{\beta} = -0.497$, confirming the pattern from figure 4 that there is a strong negative association between the local norm towards the public provision of training and the observed incidence of apprenticeship training among firms. Further note that the point estimate is statistically highly significant, with a large robust t-value of about 5.1. Moreover, the point estimate implies quite a large elasticitiy of -0.389 (approximate elasticities, evaluated at mean values of the involved variables, are given in brackets in this and the following tables).

The inclusion of survey-year dummies picks up a lot of variation in firms' training behavior, as shown in column 2 (i.e. there is quite a large increase in the R-squared), but at the same time it does not heavily influence the point estimate of β. This is because the sampling frame included a different fraction of non-training firms in the different waves of the survey (i.e. because the response rate varied across the waves). The resulting point estimate is thus only slightly smaller, and it remains large and statistically significant ($\widehat{\beta} = -0.438$, with a robust standard error of about 0.085).

We next add, in the third column, the size of firm, their sector of activity as well as ownership (private versus public and nonprofit employers).[19] This specification yields a point estimate that is somewhat stronger (i.e. more negative) than the estimates from the preceeding two columns, and it remains highly significant ($\widehat{\beta} = -0.551$, with a robust standard error of about 0.059). The comparison with the preceeding columns shows that the firm-level variables, taken together, are highly predictive of a firm's training behavior (as indicated by the large increase in the R-squared, from 0.127 to 0.326). Yet it appears that firms residing in communities with a weaker norm towards the private provision of training have in fact characterized that make them, a-priori, more likely to train apprentices than those in regions with a stronger norm. For that reason, the inclusion of these controls makes the effect of the local norm towards the private provision of training even stronger.

Next, column 4 further adds a full set of cantonal fixed effects, yielding a point estimate of $\widehat{\beta} = -0.280$. As expected, the inclusion of the fixed effects lowers the point estimates

[19] To save space, we do not show the full regression results but they are, of course, available upon request.

substantively – by about 50%, compared to the preceeding column. This confirms our expectation that there is large variation in the training incidence across cantons which is potentially due to institutional factors (e.g. regulations of general education at the upper-secondary level). Nonetheless, even in this demanding specification, the point estimate of β remains substantively large as well as statistically significant, with a robust t–value of about 5.49.

Finally, in the fifth and final column of table 2, we further add a couple of community-level controls (e.g. the size and the type of the community, i.e. whether a community is an agglomeration or rural community, or the the age distribution in a given region), yielding an estimate of $\widehat{\beta} = -0.393$ with an associated robust standard error of about 0.087.[20] Similar to the inclusion of the firm-level controls, adding community-level controls makes the estimated point estimate of the local norm stronger, i.e. more negative, suggesting that those communities with a weaker norm towards the private provision of VET have features that make it more likely that employers provide apprenticeship positions.[21]

Our first set of estimates thus shows that firms which are located in regions characterized by a strong norm towards the private provision of training are significantly and substantively (see section 5.4 below for an illustration of the size of the estimated effect) more likely to provide apprenticeship positions than comparable firms in locations with a weaker norm. We next provide several additional checks to further probe the robustness of this result.

Robustness checks

Treating the specification from column of table 2 as our benchmark, table 3 presents a couple of robustness checks.

Table 3

A first check is to include additional or more detailed controls at the regional level. Thus the specification in column 1 includes a couple of additional, regional-level controls (such as the log number of firms within a local labor market or the average size of a firm in a local labor

[20]The full list of controls is as follows: log population size of the community in the year 2000, the change in log population size (i.e. growth) between 1970 and 2000, the share of foreigners (i.e. inhabitants without Swiss citizenship), the change in the share of foreigner between 1970 and 2000, the mean age in the year 2000 in the local population, the share of individuals aged below 18 (above 65), the type of community, the area of a community, and the mean turnout in the two votes (i.e. vote nr. 340 and 503).

[21]Appendix table A.2 further shows that the negative effect of public attitudes towards the role of the state on employers' training behavior exists for different aggregation levels with regard to the local norm.

market). This yields an estimate of β that is only marginally smaller (i.e. less negative) than our baseline estimate ($\widehat{\beta} = -0.342$, with a robust standard error of about 0.089). A similar check is to include regional fixed effects at a finer level of aggregation. This is done in columns 2 and 3 which, respectively, include a full set of fixed effects at the level of districts and local labor markets (instead of cantonal fixed effects).[22] Similar to column 1, these two specifications yield estimates of β that are considerably smaller (in absolute terms) than our baseline estimate – yet they remain large, both statistically and substantively. Controlling for fixed effects at the level of districts (local labor markets) yields an estimate of $\widehat{\beta} = -0.227$ ($\widehat{\beta} = -0.230$), with a robust standard error of 0.089 (0.092).

A next check is to see whether the result is simply driven by the obvious difference in the training incidence between the different language regions within Switzerland (cf. figures 2 and 3). We thus restrict the estimation sample, in column 4, to those communities from the German-speaking part of Switzerland only (which reduces the sample size to 15,706 observations). The resulting point estimate of $\widehat{\beta} = -0.322$, however, is very close to our baseline estimate. Thus our result is not simply driven by corresponding differences in training behavior and attitudes towards the role of the state between the different language regions within Switzerland.[23] Another potential issue is that we have only few firm-level observations is some communities (while having full information regarding the main regressor). However, using only observations from regions with at least ten different employers per region also yields a point estimate of similar size as our baseline estimate, suggesting that this is actually not an important issue in our context. At the same time, it is somewhat less clear whether social norms can be effective in larger communities. We thus focus on observations located in the larger regions (i.e. regions with more than 10,000 inhabitants) only in column 6, again finding that the resulting point estimate, $\widehat{\beta} = -0.449$, is not very different from our baseline estimate. Column 7 further, and reassuringly, shows that the point estimate does hardly change when we focus on private

[22]There are 148 (106) distinct districts (local labor markets), but only 26 cantons; see appendix table A.1. Obviously, the more disaggregated fixed effects will not only pick up much of the variation in employers' traiing behavior due to unobserved regional characteristics, but a substantial fraction of the variation in the local norm as well.

[23]Similarily, using only the French-speaking regions yields a point estimate of $\widehat{\beta} = -0.383$ (not shown in table 3). With a robust t-value of about -2.4, this estimate is statistically significantly different from zero. At the same time, the point estimate is not statistically different from the one derived using the German-speaking regions. Using only the Italian-speaking regions, however, yields an insignificant, but even positive point estimate of β. Thus, consequently, excluding the Italian-speaking regions yields an even higher estimate of $\widehat{\beta} = -0.412$ (with a robust standard error of 0.093).

employers only ($\widehat{\beta} = -0.381$, with a robust standard error of 0.089). Next, column 8 shows that the point estimate remains negative and significant when we focus on smaller employers (employers with less than 50 employees) only.

The final two columns present robustness checks with respect to more technical issues. The first check, shown in column 9, uses the sampling weights provided along with the data; the final column estimates the model by probit. Again, our result is robust against these checks, as we find a very similar point estimate when we use the sampling weights that come along with the survey data. Finally, the average marginal effect from a probit model (equal to -0.344) is also very close to the marginal effect from our baseline OLS estimate.

5.3 Instrumental-variable estimates

Next, table 4 presents a set of instrumental-variable estimates. For the ease of comparison, the first column of table 4 simply replicates the OLS estimates from column 5 of table 2 above. The remaining columns present instrumental-variables estimates using different instruments and/or different estimation methods (as detailed in the bottom of the table).[24]

Table 4

The second column shows 2SLS estimates of our baseline specification, instrumenting our main regressor $N_{j[i]}^{\text{VET}}$ using $N_{j[i]}^{\text{STATE}}$, the mean share of supporting votes from the three votes on the division of responsibilities between private and public actors outside the realm of educational policy (as discussed in section 4). This yields an 2SLS estimate of $\widehat{\beta} = -0.503$, which is close in size to our baseline OLS estimate (as expected, however, it also turns out to be estimated with less precision; blurring the difference to the corresponding OLS estimate). In fact, the 2SLS estimate from column 2 is not statistically different from our baseline OLS estimate, which is indicated by a formal test on the equivalence between OLS and 2SLS estimates (the p-value associated with the corresponding regression test proposed by Wooldridge (2010) is shown at the bottom of table 4). In column 3 we use $N_{j[i]}^{\text{HIST}}$ as instrument for $N_{j[i]}^{\text{VET}}$, which yields an

[24]First-stage F-values are shown at the bottom of table 4, first-stage and reduced form estimates are shown in appendix tables A.3 and A.4, respectively. Moreover, appendix table A.5 shows additional estimates that use voting results from "placebo votes" (i.e. votes that we exepct to be unrelated to firms' training behavior). These additional estimates show that both OLS and 2SLS estimates turn out to be statistically insignificant when using the placebo votes either directly as regressor instead of $N_{j[i]}^{\text{VET}}$, or as an instrument for $N_{j[i]}^{\text{VET}}$.

estimate of $\widehat{\beta} = -0.427$. While this estimate is not statistically significant anymore, note that this is mainly due to the large increase in the associated standard error (the estimate is very close to being significant, however, with a robust t–value of about -1.61). At the same time, the confidence interval associated with this point estimate overlaps with the one from column 2, suggesting that the two instruments yield point estimates of the effect of public attitudes on firms' training behavior which are largely consistent with each other. This is confirmed by the large p-value (about 0.893) associated with the corresponding formal test of the equivalence between the 2SLS and the OLS estimate. A similar, though again statistically significant, estimate results when we use the historical elections results $E_{j[i]}^{HIST}$ as instruments for $N_{j[i]}^{\text{VET}}$, as shown in column 4.

In column 5 we use the historical voting results and the historical election results as instruments at the same time, yielding again a statistically significant and negative point estimate of $\widehat{\beta} = -0.567$ (with a robust standard error of 0.208). Again, the resulting point estimate is close to our baseline OLS estimate and a formal test on the equivalence between the two estimates does not reject the null hypothesis (the p-value associated with the test is 0.376). The final column of table 4 shows the estimate resulting from using the full set of instruments simultaneously, estimated. Again, this yield a statistically significant point estimate of $\widehat{\beta} = -0.503$, with a robust standard error of about 0.132. Moreover, this estimate is close in size to the estimate from column 2 (which in turn is very close to the baseline OLS estimate, as already discussed). In fact, it is not statistically different from the baseline OLS estimate shown in the first column.[25]

Taken together, the different instrumental-variable estimates yield a coherent pattern of estimates suggesting a negative effect of normative attitudes on the likelihood of offering apprenticeship positions. Moreover, almost all instrumental-variable estimates are quantitatively very close to our baseline OLS estimates, and we thus stick with OLS for the remainder of our empirical analysis. If anything, the comparison between with the instrumental-variable estimates shows that OLS will tend to underestimate the effect of the local norm towards the private provision of training – which implies that the reported estimates are perhaps slightly

[25]Because we use clustered standard errors, we have also estimated the two specifications from columns 5 and 6 of table 4 using GMM instead of 2SLS. This yields point estimates very close to those reported in table 4, however. They are therefore not reported in the table.

conservative.

5.4 Quantitative implications

Our main estimates from tables 2 and 4 are not only statistically significant, they also imply that variation in the social norm towards the private provision of public goods has a sizeable economic impact on firms' training behavior. This is probably best illustrated by some simple back-of-the-envelope calculations. One natural starting point are the observed differences in training incidence across larger regions within Switzerland, such as across cantons. This is illustrated graphically in panel (a) of figure 5 (which essentially reproduces the pattern from figure 4 above).

Figure 5

To illustrate the quantitative implications of the estimate of β from our baseline specification, we simply predict T_i using the estimated coefficients from our baseline specification of equation (1) and replacing the actual value of $N_{j[i]}$ with the maximum value observed from the distribution of $N_{j[i]}$ across all communities. We then aggregate these predictions within each canton and plot, in panel (b) of figure 5, the hypothetical change in the cantonal training probability against the effectively observed incidence of training.

The figure shows that the estimated effect of $N_{j[i]}$ on the regional training probability is quantitatively important. For example, the canton of Geneva is the canton with both the lowest average training incidence across its communities as well as the weakest norm towards the private provision of VET (in panel (a) of figure 5, Geneva is thus located below right, indicated by "GE"). Shifting attitudes towards the role of the state to the level most critital of the state observed in the data would imply a huge incrase in the incidence of training, however. In the case of Geneva, the training incidence would increase by about 12 percentage points (starting from a low observed training incidence of 15 percent, this implies a relative increase in the training incidence of about 80%). In contrast, there is not much of an effect for the canton of Glarus (labeled "GL" in figure 5), which already has the highest training incidence and a very strong norm towards the private provision of training.

5.5 Additional results

We next provide a couple of additional results.

Regional differences in the supply of apprentices

A first set of results, shown in columns 1 to 3, shows that our baseline estimates are robust to the inclusion of additional variables that explicitly aim to control for regional differences in the supply of apprentices.[26]

Table 5

In the first column of table 5, we add three variables measuring the distance to the nearest high school ("Gymnasium"), to the nearest vocational school, and the nearest full-time vocational school as additional controls for the supply of apprentices. In the second column we add, on top of these three variables, a set of variables representing the share of individuals with a given level of educational attainment. These variables may serve as additional supply controls because parents tend to pass on their educational preferences to their children.[27] We also add the ratio of [xxx] as an additional supply control in this specification. These two additional specifications yield an estimate of β close to the specification without the supply-side controls ($\widehat{\beta} = -0.379$ and $\widehat{\beta} = -0.322$, respectively). This suggests that our main estimates are not driven by differences in the supply of apprentices, conditional on the other controls.

Alternative data

The next two columns of table 5 show estimates with a similar set of controls as our baseline specification, but that are based on an alternative source of data. Instead of using the pooled survey data, we here merge firm-level data from the business census from the year 2008 to the set of regional variables. We select basically the same subset of firms as those sampled in the survey data (i.e. only employers from the non-agricultural sector and excluding micro

[26] We use a wording here that is consistent with the existence of a market for apprenticeship positions, where there is a demand for apprentices by firms and a corresponding supply by adolescents.

[27] Note that the variables are constructed using the census data from the year 2000, and thus these variables do not refer to the same year(s) as the firm-level data. At the same time, however, these variables presumably also reflect differences in firms' demand for VET. In that case it would be better not to include these variables as controls (e.g. Angrist and Pischke, 2008). For that reason, we prefer the specifications that do not include these additional controls, and thus our baseline estimates do not include these variables as controls.

enterprises), and we show both estimates without and with the inclusion of control variables.[28] As expected, the resulting estimates are consistent with our baseline estimate using the firm-level survey data. While not surprising, given that the survey data have originally been drawn from the business census (tough not exactly from the data we use here), it is nonetheless a reassuring finding that the estimates are comparable in size.

Alternative outcomes

The richness of the survey data further allows us to construct a variety of alternative outcome variables. Therefore, in the remaining columns of table 5 we show estimates for several alternative outcomes (in most cases, however, alternatice outcomes are available for training firms only). First, column 5 uses the absolute number of apprentices (including zero apprentices) as dependent variable, yielding an insignificant point estimate ($\widehat{\beta} = 0.321$, with a robust standard error of 1.854). Consistent with this, using only the subset of training firms also yields an insignificant estimate for the number of apprentices (cf. column 6 of table 5). We think that this finding tends to confirm our main result because it implies that the social norm affects whether an employer trains or not – given that decision, the number of apprentices is then primarily influenced by other factors, such as the size the of the employer, for example.

Moreover, the remaining two columns of table 5 show that the local norm does neither have a significant effect on the overall training costs nor on the net benefits from training apprentices accruing to the training firm.[29] Again, it is reassuring to find that there is no effect of public attitudes towards the role of the state on neither the overall training costs nor on the net benefits from training. Finding no effect of the local norm on neither costs nor on net benefits is consistent with the fact that employers from different regions within Switzerland are essentially faced with the same institutional context regarding VET (as discussed in section 2 above). The zero effect on costs and net benefits in turn also suggests that our main result can not be explained away by regional differences in either of these factors.

[28] In the survey, there was an additional step that excluded firms from the sampling frame which stated that they were unable to train apprentices (see Potterat, 2011, for details). Because we can not reproduce this specific step here, the two populations are not exactly identical.

[29] Using the log of training costs or the log of net benefits yields the same qualitative finding (i.e. a positive but insignificant estimate for log costs, and a negative and insignificant estimate for log net benefits).

5.6 Testing subsidiary hypotheses

In the final results section we present a series of additional results that test subsidiary hypotheses in an effort to support the credibility of our main results from section 5.

Interaction with the expected short-run benefits of an apprenticeship

A first such test is based on the observation that some apprenticeships yield a positive average net benefit in the short-run, i.e. until the end of the training period, while others are associated with considerable net costs. Based on this, an ancillary hypothesis postulates that the local norm towards the private provision of training has a weaker partial effect on the incidence of training if there is an expected net benefit from training. That is, one may assume that it is less costly for an employer to comply with the norm if the costs of training are lower.

Table 6

The first column of table 6 thus includes the interaction term between $N_{j[i]}^{\text{VET}}$ and a dummy variable indicating whether an employer is active in a sector where training apprentices is, on average, associated with positive net benefits.[30] We find that the interaction term yields a positive and significant coefficient estimate of about 0.127, implying that the marginal effect of $N_{j[i]}^{\text{VET}}$ on the probability of training is weaker for those employers who train apprentices that are less costly from an ex-ante point of view. We show estimates that use a slightly different construction of the expected net benefit from training in column 2 of table 6 (using a finer level of aggregation with regards to the net benefits from training). This alternative specification yields also a positive, and statistically significant, coefficient estimate of about 0.228.

Interactions with employer characteristics

We further expect that some employers are more likely to be influenced in their decisions by local norms than others. Specifically, we speculate that large(r) employers are presumably more sensitive to local attitudes towards the role of the state within the context of VET because they are more visible and because they are faced with more elevated expectations than the smaller

[30]More specifically, we determine the net benefits from training (i.e. the monetarized benefits from training minus the costs of training) within a given economic sector and then assume that firms active within the same sector can expect to realize the same net benefits.

employers. As shown in columns 3 and 4 of table 6, we indeed find that larger employers react more strongly to changes in the local norm than smaller firms (i.e. the estimated coefficient on the interaction term is large and statistically significant in both cases).[31]

Another dimension that could be relevant is foreign ownership of an enterprise or if a firm mainly serves foreign demand. A priori, we expect to find that employers in foreign ownership and firms that mainly serve foreign demand to be less sensitive to local norms than other employers. We thus include the interaction between $N_{j[i]}^{\text{VET}}$ and a dummy indicating that an employer is in foreign ownership in column 5 of table 6, and the interaction term between $N_{j[i]}^{\text{VET}}$ and a dummy indicating that a given firm mainly supplies foreign demand in column 6. In these two cases, however, the point estimate associated with the interaction term is insignificant (even though the main effect has the expected sign in both specifications).

Employers' self-perception of the motives for providing apprenticeship positions

A final issue worth exploring is whether public attitudes towards the role of the state affect employers' self-perception regarding the motives for (not) providing apprenticeship positions. In fact, one can argue that these variables in part reflect the internalization of the norm by employers. Specifically, in the survey training firms were directly asked about the importance of various motives for providing apprenticeship positions from their own point of view, some of them reflecting economic considerations (i.e. they may state that "training apprentices is important to remain competitive" or that "training apprentices is essential for keeping innovative"), others being of less or no obvious economic significance (for example, employers may state that "training apprentices is a community task" or that "training is part of the corporate identity"). We expect that employers are more likely to state that they care about noneconomic motives if they are located in a region with a strong norm towards the private provison of training, whereas we expect to find no corresponding effect in the case of economic motives.

Table 7

Table 7 reports the corresponding estimates, using both the minimal and the full specification that we have already used before. In the first four columns, the dependent variable reflects

[31] We define employers with more than 10 (more than 50) employees as large firms in column 3 (column 4) of table 6, while noting that about 84% of all firms had less than 10 employees in the year 2008.

the importance of noneconomic and economic motives, respectively.[32] The first two columns look at the importance of noneconomic motives for employers' training decision. There is a negative and significant effect of the local norm towards the private provision of training on the likelihood that an employer states that noneconomic motives are important in his/her decision to train apprentices. Remarkably, the negative effect is robust to the inclusion of the full set of control variables used in the baseline regressions above. Thus employers located in regions with a stronger norm towards the private provision of training are more likely than similar employers in regions with a weak(er) norm to state that apparently noneconomic motives are relevant for their decision to train apprentices. This evidence is consistent with awareness of the norm on the part of the employers, and perhaps even with norm internalization.

Column 3 shows that there is also a negative and significant effect of public attitudes on the likelihood that employers state that economic motives are important for their decision to train apprentices. However, and in contrast to the importance of noneconomic motives, this effect completely vanishes once we include additional control variables, as shown in column 4 of table 7. Note that this result is not driven an excessive increase in the associated standard error (the increase is similar to the one observed in the first two columns). Rather, it is the shrinkage in the corresponding point estimate that is responsible for this finding.

Columns 5 and 6 also look at the effect of the local norm on the importance of economic motives but, in contrast to the two preceeding columns, the underlying survey items were answered by both training and non-training firms (using a slightly different set of questions, however). The resulting estimates mirror the the result from the two preceeding columns: there is a negative association between the local norm and the self-assessed importance of economic motives for training, but this effect is driven towards zero when additional controls are included in the regression.

6 Conclusions

In this paper, we use a unique combination of different data sources to estimate the impact of social norms with regard to the role of the state on the private provision of training – a

[32] In a first step, we have constructed a dummy variable indicating consent for every single survey item. The dependent variable in columns 1 and 2 (3 and 4) simply measures the fraction of items an employer has classified as important for his/her training decision within the set of noneconomic (economic) motives.

topic not only of academic but also of considerable public interest. We combine firm-level survey data with community-level voting results from different votes that dealt with the issue of public versus private involvement in the provision of public goods. We use the voting results to measure local public attitudes towards the role of the state in general and towards the private – rather than the public – provision of vocational education and training in particular. In the first part of our empirical analysis, we present several pieces of evidence which suggests that the voting results are a internally consistent and valid measure of the local norm towards the private provision of VET.

In line with the vast, though mainly experimental, evidence on the effect of social norms on the private provision of public goods, we hypothesize that firms which are located in regions with a stronger norm towards the private provision of training are, ceteris-paribus, more likely to provide such training positions; either because they have internalized the norm and/or because the norm is enforced in their community. In line with our hypothesis, we find that there is a significant and surprisingly strong correlation between public attitudes towards the role of the state and the incidence of training among employers within a given region. Employers located in regions with a strong norm towards the private provision of training are much more likely to provide apprenticeship positions than similar employers located in regions where the corresponding norm is weaker.

The negative association between the local norm towards the private provision of VET turns out to be very robust to a wide variety of robustness checks and alternative model specifications. Most importantly, perhaps, the negative association between the local norm and the regional training incidence is robust to the inclusion of a variety of firm- and regional-level controls, such as the number of employees or cantonal fixed effects. Moreover, we find a quantitatively similar-sized effect of the norm on employers' training behavior when correcting for simultaneity bias using different instrumental variables. Taken together, the resulting estimates are surprisingly robust and consistent across different specifications. We thus conclude that our findings point to the importance of attitudes towards the role of the state as an important explanatory factor with regard to firms' training decisions.

References

Acemoglu, D. and Pischke, J.-S. (2000). Certification of training and training outcomes. *European Economic Review*, **44**(4), 917–927.

Aepli, M., Kuhn, A., and Schweri, J. (2018). Culture Matters: How Social Norms Affect Firms' Training Decisions. Mimeo, Swiss Federal Institute for Vocational Education and Training.

Agell, J. (1999). On the benefits from rigid labour markets: norms, market failures, and social insurance. *Economic Journal*, **109**(453), 143–164.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Azar, O. H. (2004). What sustains social norms and how they evolve? The case of tipping. *Journal of Economic Behavior & Organization*, **54**(1), 49–64.

Bandiera, O., Barankay, I., and Rasul, I. (2005). Social preferences and the response to incentives: Evidence from personnel data. *Quarterly Journal of Economics*, **120**(3), 917–962.

Bassanini, A., Brunello, G., and Caroli, E. (2017). Not in my community: Social pressure and the geography of dismissals. *Journal of Labor Economics*, **35**(2), 429–483.

Benz, M. and Meier, S. (2008). Do people behave in experiments as in the field? Evidence from donations. *Experimental Economics*, **11**(3), 268–281.

Bertrand, M. and Mullainathan, S. (2001). Do People Mean What They Say? Implications for Subjective Survey Data. *American Economic Review*, **91**(2), 67–72.

Blatter, M., Muehlemann, S., and Schenker, S. (2012). The costs of hiring skilled workers. *European Economic Review*, **56**(1), 20–35.

Bowles, S. (2016). *The moral economy. Why good incentives are no substitute for good citizens*. Yale University Press.

Brennan, G., Eriksson, L., Goodin, R. E., and Southwood, N. (2013). *Explaining norms*. Oxford University Press.

Busemeyer, M. R. and Trampusch, C. (2011). *The political economy of collective skill formation*. Oxford University Press.

Cameron, A. C. and Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, **50**(2), 317–372.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, **14**(1), 47–83.

Eugster, B., Lalive, R., Steinhauer, A., and Zweimüller, J. (2011). The demand for social insurance: Does culture matter? *Economic Journal*, **121**(556), F413–F448.

Eugster, B., Lalive, R., Steinhauer, A., and Zweimüller, J. (2017). Culture, Work Attitudes, and Job Search: Evidence from the Swiss Language Border. *Journal of the European Economic Association*, **15**(5), 1056–1100.

Fehr, E. and Fischbacher, U. (2004). Social norms and human cooperation. *Trends in Cognitive Sciences*, **8**(4), 185–190.

Frey, B. S. and Meier, S. (2004). Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment. *American Economic Review*, **94**(5), 1717–1722.

Hong, H. and Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, **93**(1), 15–36.

Janssen, S., Sartore, S. T., and Backes-Gellner, U. (2016). Discriminatory social attitudes and varying gender pay gaps within firms. *Industrial & Labor Relations Review*, **69**(1), 253–279.

Lalive, R. and Stutzer, A. (2010). Approval of equal rights and gender differences in well-being. *Journal of Population Economics*, **23**(3), 933–962.

Laury, S. K. and Taylor, L. O. (2008). Altruism spillovers: Are behaviors in context-free experiments predictive of altruism toward a naturally occurring public good? *Journal of Economic Behavior & Organization*, **65**(1), 9–29.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. Kagel and A. Toth, editors, *Handbook of Experimental Economics*, chapter 2, pages 111–193. Princeton University Press.

Levitt, S. D. and List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, **21**(2), 153–174.

Mohrenweiser, J., Wydra-Somaggio, G., and Zwick, T. (2017). Information Advantages of Training Employers Despite Credible Training Certificates. Technical report, University of Zurich, Department of Business Administration (IBW).

Moretti, L., Mayerl, M., Muehlemann, S., Schlögl, P., and Wolter, S. (2017). So similar and yet so different: A comparative analysis of a firm's cost and benefits of apprenticeship training in Austria and Switzerland. *IZA Discussion Paper No. 11081*.

Moulton, B. R. (1986). Random group effects and the precision of regression estimates. *Journal of Econometrics*, **32**(3), 385–397.

Muehlemann, S. and Wolter, S. C. (2011). Firm-sponsored training and poaching externalities in regional labor markets. *Regional Science and Urban Economics*, **41**(6), 560–570.

Muehlemann, S. and Wolter, S. C. (2014). Return on investment of apprenticeship systems for enterprises: Evidence from cost-benefit analyses. *IZA Journal of Labor Policy*, **3**(1), 25.

Muehlemann, S., Pfeifer, H., Walden, G., Wenzelmann, F., and Wolter, S. C. (2010). The financing of apprenticeship training in the light of labor market regulations. *Labour Economics*, **17**(5), 799–809.

Muehlemann, S., Ryan, P., and Wolter, S. C. (2013). Monopsony power, pay structure, and training. *Industrial & Labor Relations Review*, **66**(5), 1097–1114.

Mueller, B. and Schweri, J. (2015). How specific is apprenticeship training? Evidence from interfirm and occupational mobility after graduation. *Oxford Economic Papers*, **67**(4), 1057–1077.

OECD (2008). Education at a glance. Paris: OECD.

Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives*, **14**(3), 137–158.

Potterat, J. (2011). Methodenbericht. Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe im Jahr 2009 (KNBB09). Stichprobenplan, Gewichtung und Schätzverfahren. Neuchâtel: Bundesamt für Statistik (BFS).

SERI (2014). Vocational and Professional Education and Training in Switzerland. Facts and Figures 2014. State Secretary for Education, Research and Innovation SERI.

Shang, J. and Croson, R. (2009). A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. *Economic Journal*, **119**(540), 1422–1439.

Strupler, M. and Wolter, S. C. (2012). *Die duale Lehre: Eine Erfolgsgeschichte – auch für die Betriebe*. Rüegger Verlag: Glarus.

Stutzer, A. and Lalive, R. (2004). The role of social work norms in job searching and subjective well-being. *Journal of the European Economic Association*, **2**(4), 696–719.

Wettstein, E., Schmid, E., and Gonon, P. (2017). *Swiss Vocational and Professional Education and Training (VPET). Forms, System, Stakeholders*. Bern: hep Verlag.

Wolter, S. C. and Ryan, P. (2011). Apprenticeship. *Handbook of the Economics of Education*, **3**, 521–576.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Table 1: List of votes used to measure norms towards the role of the state

Nr.	Date	Title/description	Result	Share of supporting votes	Turnout
(a) Votes about vocational education and training					
503	18.05.2003	Popular initiative for a "sufficient supply of vocational education and training"	Rejected	31.6%	49.6%
340	28.09.1986	Popular initiative for a "secured vocational education and training and retraining"	Rejected	18.4%	34.8%
(b) "Historical" votes about vocational education and training and other educational policies					
292	03.12.1978	Federal law on vocational education and training	Accepted	56.0%	43.2%
286	07.10.1977	Federal law on funding universities and research	Rejected	43.3%	48.9%
207	24.05.1964	Federal law on vocational education and training	Accepted	68.8%	37.0%
(c) Other votes on the provision of public goods or demanding more public intervention					
528	11.03.2007	Popular initiative for a "unitary public health insurance"	Rejected	28.8%	45.9%
461	12.03.2000	Popular initiative for a "fair representation of women in the Federal Administration"	Rejected	18.00%	42.20%
415	04.12.1994	Federal law concerning health insurance	Accepted	51.8%	44.0%

Notes: The vote number corresponds to the official numbering of the votes used by the Swiss Federal Administation. The share of supporting votes equals the fraction of all valid votes cast that were in favor of the vote, while turnout describes the fraction of eligible voters taking part in the vote.

Table 2: Baseline estimates (OLS estimates)

	Training firm (yes = 1), T_i				
Mean	0.335	0.335	0.335	0.335	0.335
Standard deviation	0.472	0.472	0.472	0.472	0.472
$N_{j[i]}^{\text{VET}}$	-0.498^{***}	-0.438^{***}	-0.551^{***}	-0.280^{***}	-0.393^{***}
	(0.098)	(0.085)	(0.059)	(0.051)	(0.087)
	[−0.389]	[−0.343]	[−0.431]	[−0.219]	[−0.307]
Survey-year dummies	No	Yes	Yes	Yes	Yes
Firm-level controls	No	No	Yes	Yes	Yes
Cantonal dummies	No	No	No	Yes	Yes
Community-level controls	No	No	No	No	Yes
Number of observations	21,339	21,339	21,339	21,339	21,339
R-Squared	0.007	0.127	0.316	0.323	0.326

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by communities. Approxmiate elasticities, evaluated at mean values, are given in brackets.

Table 3: Robustness checks (OLS estimates)

	Training firm (yes = 1), T_i									
Mean	0.335	0.335	0.335	0.367	0.345	0.331	0.302	0.225	0.276	0.335
Standard deviation	0.472	0.472	0.472	0.482	0.475	0.471	0.459	0.418	0.447	0.472
$N_{j[i]}^{\text{VET}}$	-0.342^{***}	-0.227^{**}	-0.230^{**}	-0.322^{***}	-0.503^{***}	-0.449^{**}	-0.381^{***}	-0.275^{***}	-0.309^{**}	-1.344^{***}
	(0.089)	(0.089)	(0.092)	(0.123)	(0.115)	(0.174)	(0.089)	(0.084)	(0.122)	(0.334)
	[−0.267]	[−0.178]	[−0.180]	[−0.213]	[−0.397]	[−0.407]	[−0.330]	[−0.317]	[−0.289]	[−0.344]
Robustness check	Regional controls			Subsample of regions			Subsample of firms		Statistical issues	
	LLM controls	District FEs	LLM FEs	German	$n_j \geq 10$	Large	Private	Smaller	Weights	Probit
Survey-year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cantonal dummies	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Community-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	21,339	21,339	21,339	15,706	17,260	11,432	17,930	16,279	21,339	21,339
(Pseudo) R-Squared	0.327	0.330	0.331	0.334	0.334	0.351	0.328	0.196	0.138	0.283

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by communities. Approxmiate elasticities, evaluated at mean values, are given in brackets.

Table 4: Instrumental-variable estimates (2SLS estimates)

	Training firm (yes = 1), T_i					
Mean	0.335	0.335	0.335	0.335	0.335	0.335
Standard deviation	0.472	0.472	0.472	0.472	0.472	0.472
$N_{j[i]}^{\text{VET}}$	-0.393^{***}	-0.503^{***}	-0.427	-0.823^{***}	-0.567^{***}	-0.503^{***}
	(0.087)	(0.138)	(0.265)	(0.294)	(0.208)	(0.132)
	$[-0.307]$	$[-0.393]$	$[-0.334]$	$[-0.644]$	$[-0.443]$	$[-0.393]$
Estimation method	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Instrument(s)	–	$N_{j[i]}^{\text{STATE}}$	$N_{j[i]}^{\text{HIST}}$	$E_{j[i]}^{\text{HIST}}$	$N_{j[i]}^{\text{HIST}} + E_{j[i]}^{\text{HIST}}$	All together
Survey-year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Cantonal dummies	Yes	Yes	Yes	Yes	Yes	Yes
Community-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	21,339	21,339	21,331	21,279	21272	21,272
R-Squared	0.326	0.326	0.326	0.325	0.326	0.326
F-value (first-stage)	–	425.964	89.348	12.833	29.109	88.503
p-value (endogeneity)	–	0.300	0.893	0.136	0.376	0.281

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by communities. Approxmiate elasticities, evaluated at mean values, are given in brackets. The p-value shown in the last row is associated with a formal test on the equivalence between the OLS estimate from column 1 and the corresponding instrumental-variable estimate.

Table 5: Additional results (OLS estimates)

	All firms				Training firms only			
	Training firm (yes = 1), T_i				#apprentices		costs	net benefits
Mean	0.335	0.335	0.311	0.311	1.741	5.198	98.582	6.322
Standard deviation	0.472	0.472	0.463	0.463	19.032	32.612	32.777	39.748
$N_{j[i]}^{\text{VET}}$	-0.379^{***}	-0.322^{***}	-0.573^{***}	-0.205^{***}	0.321	4.546	5.530	-19.988
	(0.087)	(0.084)	(0.071)	(0.036)	(1.854)	(5.325)	(9.784)	(12.792)
	[−0.296]	[−0.252]	[−0.486]	[−0.174]	[0.048]	[0.221]	[0.014]	[−0.797]
Additional feature	Supply-side controls		Alternative data		Alternative outcomes			
Survey-year dummies	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Cantonal dummies	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Community-level controls	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Number of observations	21,339	21,339	208,857	208,857	21,339	7,147	7,147	7,147
p-value (F-statistic)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
R-Squared	0.326	0.327	0.010	0.207	0.028	0.041	0.186	0.086

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by communities. Approxmiate elasticities, evaluated at mean values, are given in brackets.

Table 6: Heterogeneous effects (OLS estimates)

	Training firm (yes = 1)					
Mean	0.335	0.335	0.335	0.335	0.302	0.304
Standard deviation	0.472	0.472	0.472	0.472	0.459	0.460
$N_{j[i]}^{\text{VET}}$	-0.487^{***}	-0.585^{***}	-0.318^{***}	-0.258^{***}	-0.262^{***}	-0.272^{***}
	(0.105)	(0.110)	(0.088)	(0.094)	(0.082)	(0.084)
$N_{j[i]}^{\text{VET}} \times 1$(net benefits > 0)	0.127^{*}	0.228^{***}				
	(0.074)	(0.078)				
$N_{j[i]}^{\text{VET}} \times 1$(large firm)			-0.385^{**}	-0.301^{***}		
			(0.151)	(0.102)		
$N_{j[i]}^{\text{VET}} \times 1$(foreign ownership)					-0.082	
					(0.107)	
$N_{j[i]}^{\text{VET}} \times 1$(foreign demand)						-0.041
						(0.162)
Survey-year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Cantonal dummies	Yes	Yes	Yes	Yes	Yes	Yes
Community-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	21,339	21,337	21,339	21,339	15,619	15,488
R-Squared	0.326	0.327	0.327	0.327	0.379	0.378

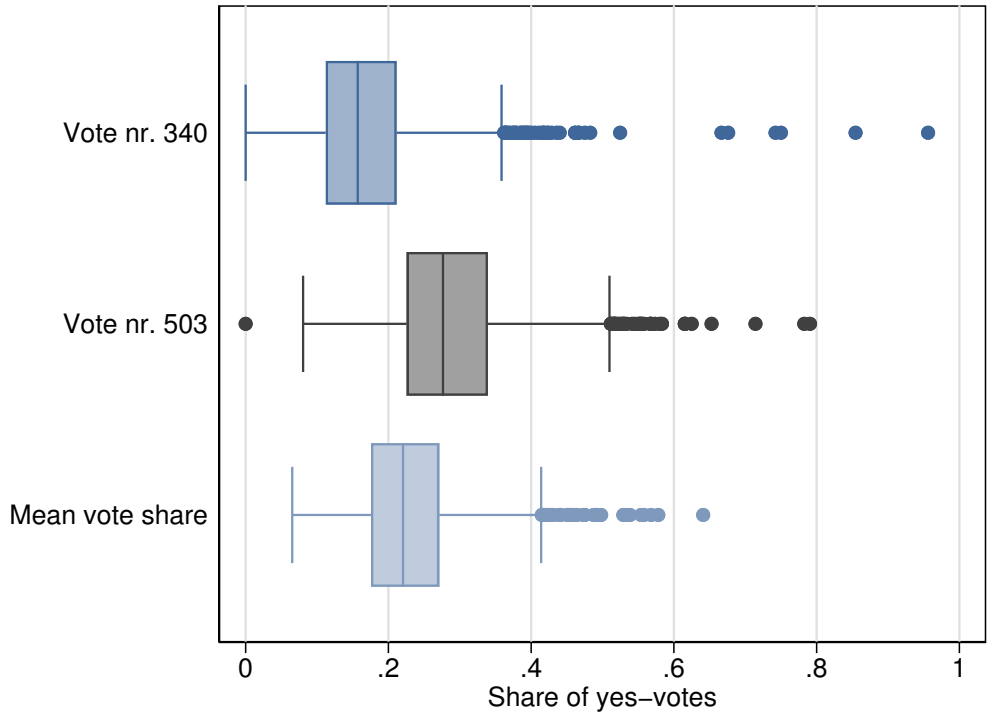
Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by communities.

Table 7: Employers' self-perception of their motives for training apprentices (OLS estimates)

	Noneconomic motives		Economic motives		Economic motives (all firms)	
Mean	0.720	0.720	0.459	0.459	0.409	0.409
Standard deviation	0.242	0.242	0.289	0.289	0.289	0.289
$N_{j[i]}^{\text{VET}}$	-0.245^{***}	-0.191^{*}	-0.198^{***}	-0.064	-0.305^{***}	-0.055
	(0.067)	(0.102)	(0.045)	(0.108)	(0.035)	(0.060)
	[−0.086]	[−0.067]	[−0.109]	[−0.035]	[−0.195]	[−0.035]
Survey-year dummies	No	Yes	No	Yes	No	Yes
Firm-level controls	No	Yes	No	Yes	No	Yes
Cantonal dummies	No	Yes	No	Yes	No	Yes
Community-level controls	No	Yes	No	Yes	No	Yes
Number of observations	7147	7147	7147	7147	18,322	18,322
R-Squared	0.006	0.065	0.003	0.060	0.007	0.049

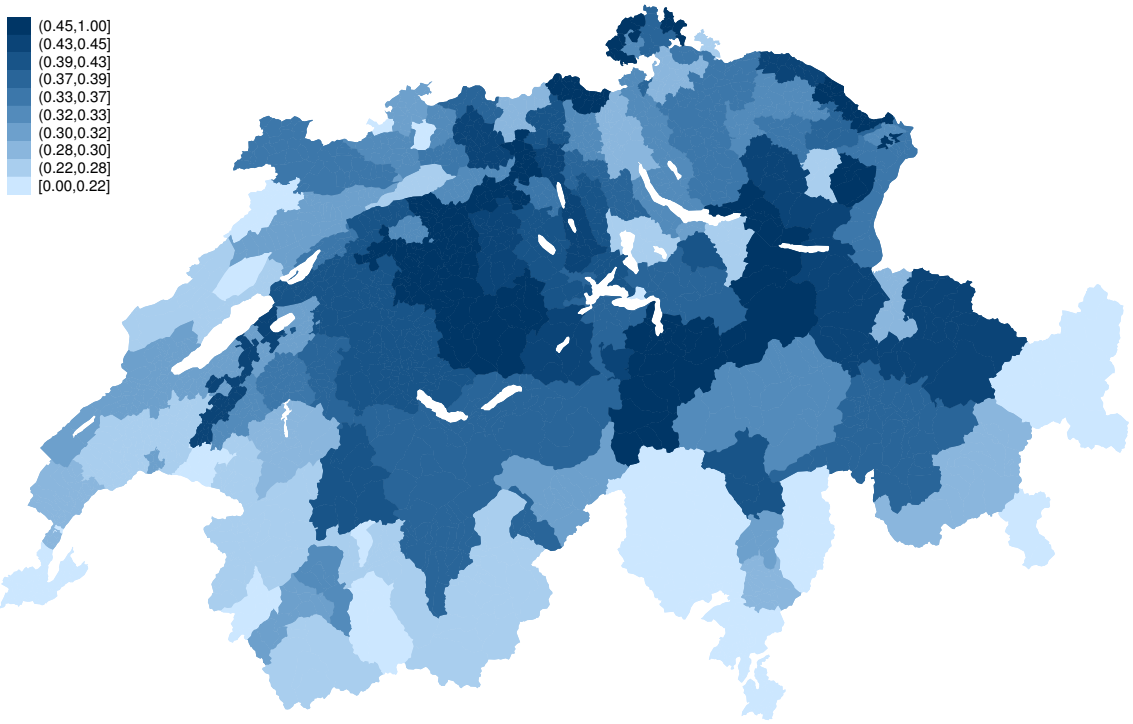
Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by communities. Approxmiate elasticities, evaluated at mean values, are given in brackets.

Figure 1: Variation in the share of supporting votes, votes nr. 340 and nr. 503



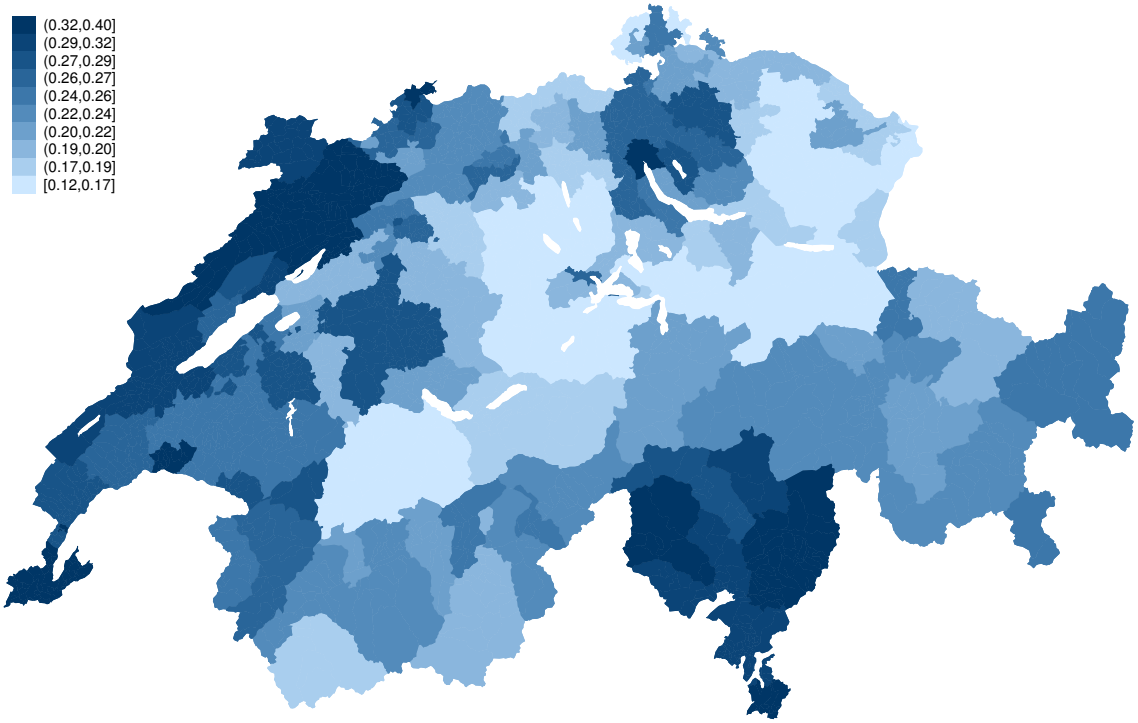
Notes: The figure shows the distribution of the share of supporting votes for vote nr. 340 and vote nr. 503, as well as the mean across the two votes (see table 1 for details).

Figure 2: Spatial variation in the incidence of training



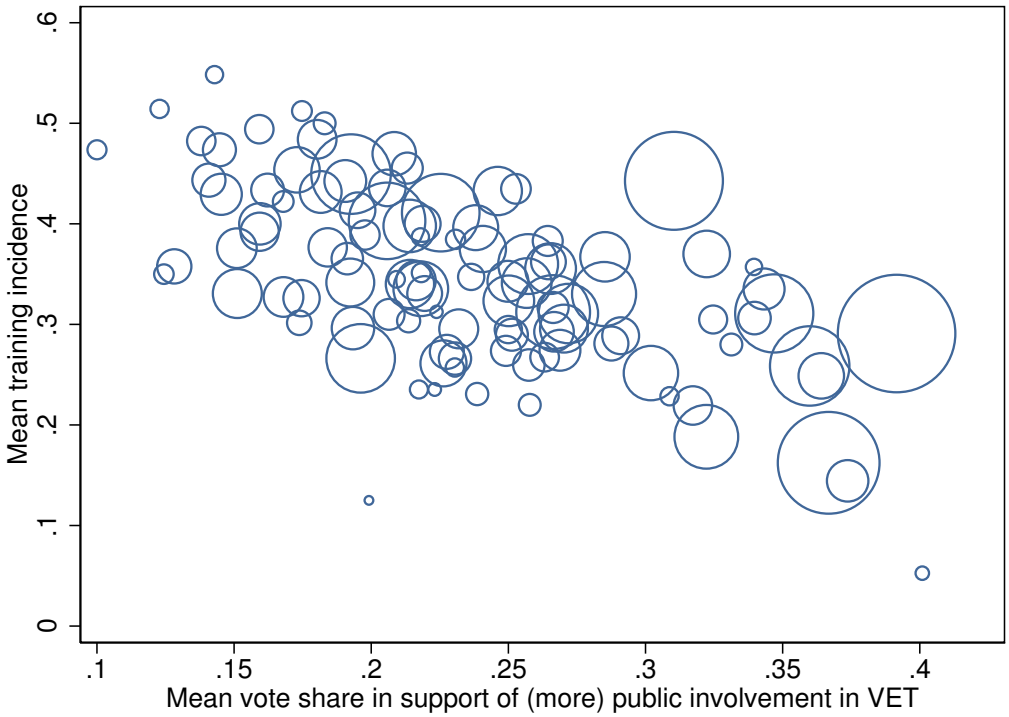
Notes: The figure shows the spatial distribution of training incidence (i.e. the local mean of T_i) across the 148 distinct districts of Switzerland.

Figure 3: Spatial variation in public attitudes towards the role of the state in VET



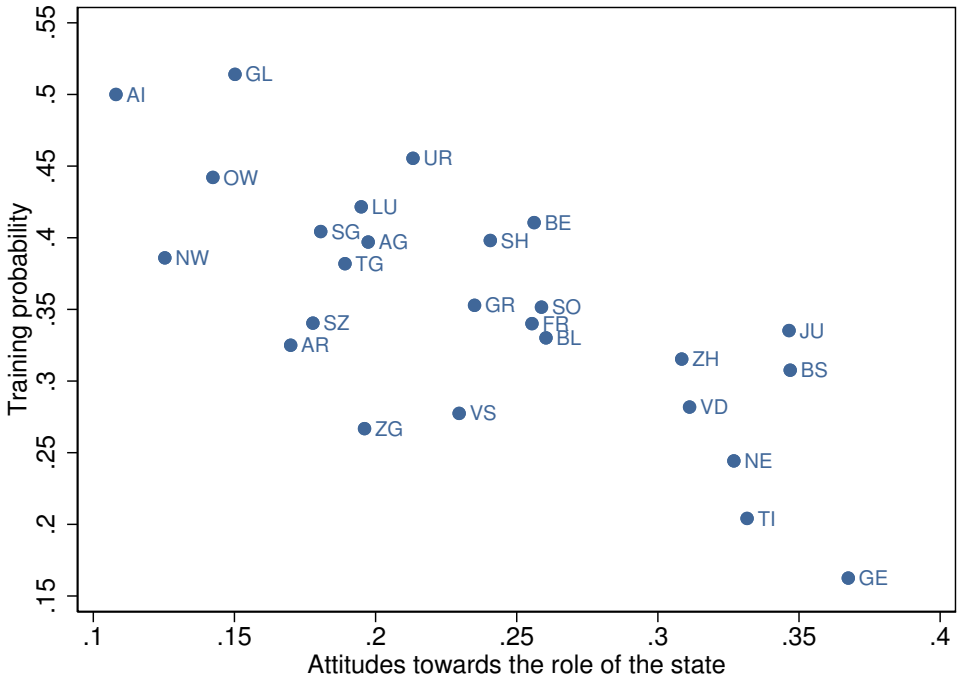
Notes: The figure shows the spatial distribution of district-level voting results (i.e. the mean share of supporting votes from vote nr. 207, 286, and 292); see table 1 for additional information concerning the three votes.

Figure 4: The association between training incidence and public attitudes towards the role of the state

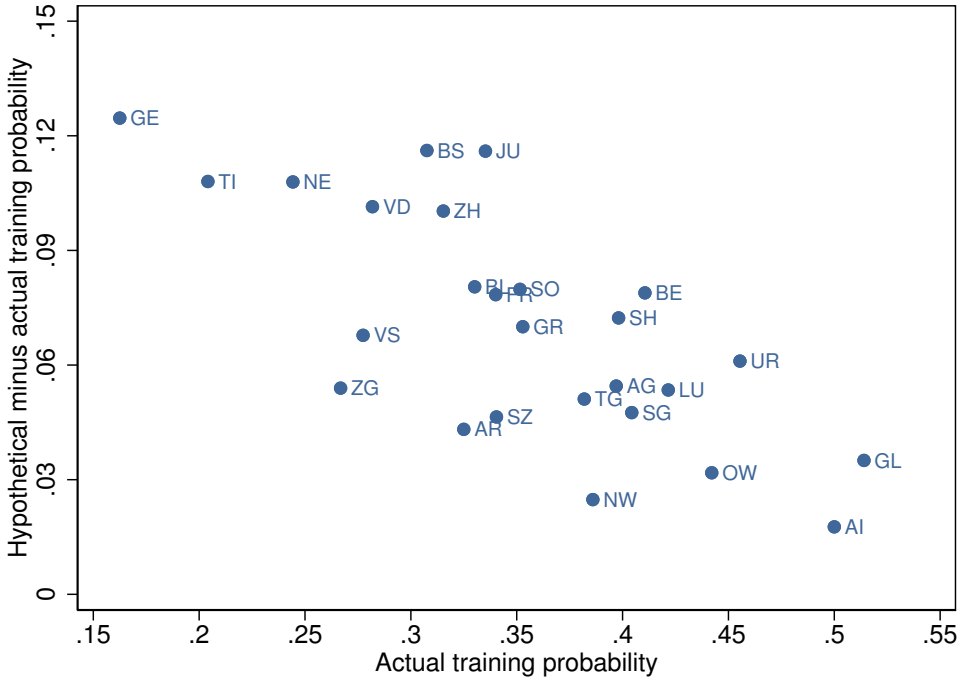


Notes: The figure plots the association between the local training incidence (on the y–axis) and public attitudes towards the role of the state within the VET system (on the x–axis). Both variables are aggregated up to the level of local labor markets, guaranteeing that the local training incidence is strictly larger than 0 and stricly smaller than 1 for each region. The size of the circles is proportional to the size of (i.e. the number of firms in) the local labor markets in the pooled sample.

Figure 5: Illustrating the quantitative effect of civic virtue on the regional incidence of training among employers



(a)



(b)

Notes: Panel (a) plots cantonal-level training probabilities against mean attitudes towards the role of the state in VET. Panel (b) illustrates the quantitative implications of the estimate of β from column 5 of table 2. The y–axis shows the predicted increase in the cantonal training probability from a hypothetical shift of $N_{j[i]}^{\text{VET}}$ to the maximum value observed in each community.

A Additional Tables and Figures

Table A.1: Spatial structure of the sample data

Spatial unit	Unique units	# of inhabitants	# of firms	Area, in km^2
Postal code	3,102	2,656	146	13.31
Community	2,352	3,502	192	17.55
District	148	55,660	3,052	278.95
Local labor market	106	77,714	4,261	389.48
Canton	26	316,833	17,372	1,587.89
Sum		8,237,666	451,663	41,285.00

Notes: The table shows the number of unique units (as of the year 2014) for different levels of regional aggregation, along with the average number of inhabitants, the average number of firms (as of 2008), and the mean area (in hectares) per spatial unit.

Table A.2: Attitudes towards the role of the state measured at different aggregation levels

	Training firm (yes = 1)					
Mean	0.335	0.335	0.335	0.335	0.335	0.335
Standard deviation	0.472	0.472	0.472	0.472	0.472	0.472
$N_{r[i]}^{\text{VET}}$	$-0.498^{\star\star\star}$	$-0.393^{\star\star\star}$	$-0.742^{\star\star\star}$	$-0.551^{\star\star\star}$	$-0.750^{\star\star\star}$	$-0.565^{\star\star\star}$
	(0.098)	(0.087)	(0.147)	(0.097)	(0.152)	(0.102)
	[−0.389]	[−0.307]	[−0.572]	[−0.425]	[−0.578]	[−0.436]
Aggregation level of $N_{r[i]}^{\text{VET}}$	Postal codes ($R = 1,984$)		Districts ($R = 148$)		Local labor markets ($R = 106$)	
Survey-year dummies	No	Yes	No	Yes	No	Yes
Firm-level controls	No	Yes	No	Yes	No	Yes
Cantonal dummies	No	Yes	No	Yes	No	Yes
Community-level controls	No	Yes	No	Yes	No	Yes
Number of observations	21,339	21,339	21,339	21,339	21,339	21,339
R-Squared	0.007	0.326	0.013	0.326	0.013	0.326

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors, clustered by r, are given in parentheses. $N_{r[i]}^{\text{VET}}$ denotes public attitudes towards the role of the state in region r, where the level of spatial aggregation varies across columns.

Table A.3: First-stage estimates

	$N_{j[i]}^{\text{VET}}$				
Mean	0.262	0.262	0.262	0.262	0.262
Standard deviation	0.081	0.081	0.081	0.081	0.081
$N_{j[i]}^{\text{STATE}}$	0.678***				0.619***
	(0.033)				(0.031)
$N_{j[i]}^{\text{HIST}}$		0.368***		0.341***	0.135***
		(0.039)		(0.037)	(0.030)
FDP^{1947}			0.001*	0.001**	0.000
			(0.000)	(0.000)	(0.000)
CVP^{1947}			0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)
SP^{1947}			0.001**	0.001*	0.001*
			(0.000)	(0.000)	(0.000)
SVP^{1947}			−0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)
LPS^{1947}			−0.002***	−0.001***	−0.001***
			(0.001)	(0.000)	(0.000)
Mitte^{1947}			0.000	−0.001	−0.000
			(0.001)	(0.001)	(0.000)
PDA^{1947}			0.003***	0.002***	0.001***
			(0.001)	(0.001)	(0.000)
Survey-year dummies	Yes	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes	Yes
Cantonal dummies	Yes	Yes	Yes	Yes	Yes
Community-level controls	Yes	Yes	Yes	Yes	Yes
Number of observations	21,339	21,331	21,279	21,272	21,272
R-Squared	0.888	0.834	0.832	0.844	0.894
F-value (instruments)	425.964	89.348	12.833	29.109	88.503

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by communities.

Table A.4: Reduced-form estimates

	Training firm (yes = 1)				
Mean	0.335	0.335	0.335	0.335	0.335
Standard deviation	0.472	0.472	0.472	0.472	0.472
$N_{j[i]}^{\text{STATE}}$	-0.341^{***}				-0.290^{***}
	(0.097)				(0.103)
$N_{j[i]}^{\text{HIST}}$		−0.157		−0.102	−0.006
		(0.101)		(0.109)	(0.111)
FDP^{1947}			0.000	0.000	0.001
			(0.001)	(0.001)	(0.001)
CVP^{1947}			0.001	0.001	0.001
			(0.001)	(0.001)	(0.001)
SP^{1947}			0.000	0.000	0.000
			(0.001)	(0.001)	(0.001)
SVP^{1947}			$0.002^{\star}$	$0.002^{\star}$	0.002
			(0.001)	(0.001)	(0.001)
LPS^{1947}			0.001	0.001	0.001
			(0.002)	(0.002)	(0.002)
Mitte^{1947}			0.000	0.000	0.000
			(0.002)	(0.002)	(0.002)
PDA^{1947}			−0.001	−0.001	0.000
			(0.002)	(0.002)	(0.002)
Survey-year dummies	Yes	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes	Yes
Cantonal dummies	Yes	Yes	Yes	Yes	Yes
Community-level controls	Yes	Yes	Yes	Yes	Yes
Number of observations	21,339	21,331	21,279	21,272	21,272
R-Squared	0.326	0.325	0.326	0.326	0.326
F-value (instruments)	12.252	2.410	1.945	1.952	2.844

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by communities.

Table A.5: Placebo votes

	Training firm (yes = 1)			
Mean	0.335	0.335	0.335	0.335
Standard deviation	0.472	0.472	0.472	0.472
$N_{j[i]}^{\text{VET}}$	-0.393^{***}	-0.503^{***}		0.328
	(0.087)	(0.138)		(0.298)
	[−0.307]	[−0.393]		[0.257]
$N_{j[i]}^{\text{Placebo}}$			−0.110	
			(0.097)	
			[−0.194]	
Estimation method	OLS	2SLS	OLS	2SLS
Instrument	–	$N_{j[i]}^{\text{STATE}}$	–	$N_{j[i]}^{\text{Placebo}}$
Survey-year dummies	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes
Cantonal dummies	Yes	Yes	Yes	Yes
Community-level controls	Yes	Yes	Yes	Yes
Number of observations	21,339	21,339	21,339	21,339
p-value (F-statistic)	0.000	0.000	0.000	0.000
R-Squared	0.326	0.326	0.325	0.323

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by communities. Approxmiate elasticities, evaluated at mean values, are given in brackets. $N_{j[i]}^{\text{Placebo}}$ is the mean share of supporting votes from two votes on different subjects: the first on the rehabilitation of the Gotthard road tunnel (vote from February 28, 2016), the second on the protection of children from paedophiles (vote from May 18, 2014).

Table A.6: The correlation between voting results and indicators of voluntary work

	"Being active"		"Being a member"		"Making donations"		"Trusting"	
Mean	1.113	1.113	2.038	2.038	2.279	2.279	6.451	6.451
Standard deviation	1.305	1.305	1.671	1.671	1.931	1.931	2.171	2.171
$N_{j[i]}^{\text{VET}}$	-1.319^{***}	-1.043^{***}	-2.524^{***}	-1.314^{***}	-2.418^{***}	-1.280^{***}	-1.560^{***}	-0.828^{*}
	(0.219)	(0.298)	(0.337)	(0.342)	(0.457)	(0.382)	(0.456)	(0.469)
	[−0.296]	[−0.234]	[−0.309]	[−0.161]	[−0.265]	[−0.140]	[−0.060]	[−0.032]
Individual-level controls	No	Yes	No	Yes	No	Yes	No	Yes
District-level dummies	No	Yes	No	Yes	No	Yes	No	Yes
Number of observations	13,820	13,820	13,820	13,820	13,820	13,820	13,587	13,587
R-Squared	0.007	0.069	0.016	0.129	0.011	0.140	0.004	0.053

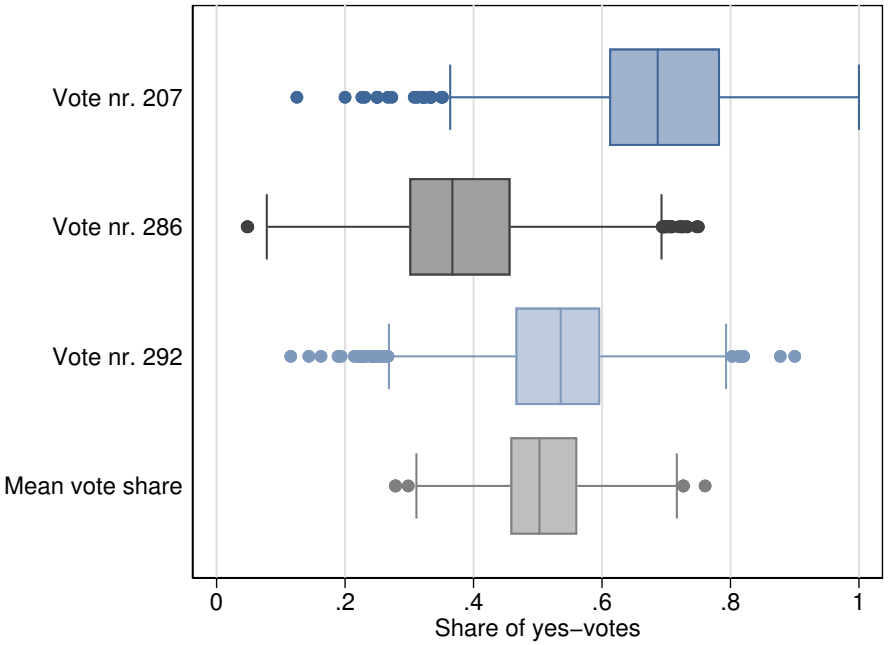
Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by community. Approxmiate elasticities, evaluated at mean values, are given in brackets.

Table A.7: Employers' long-term engagement in training (ordered probit estimates)

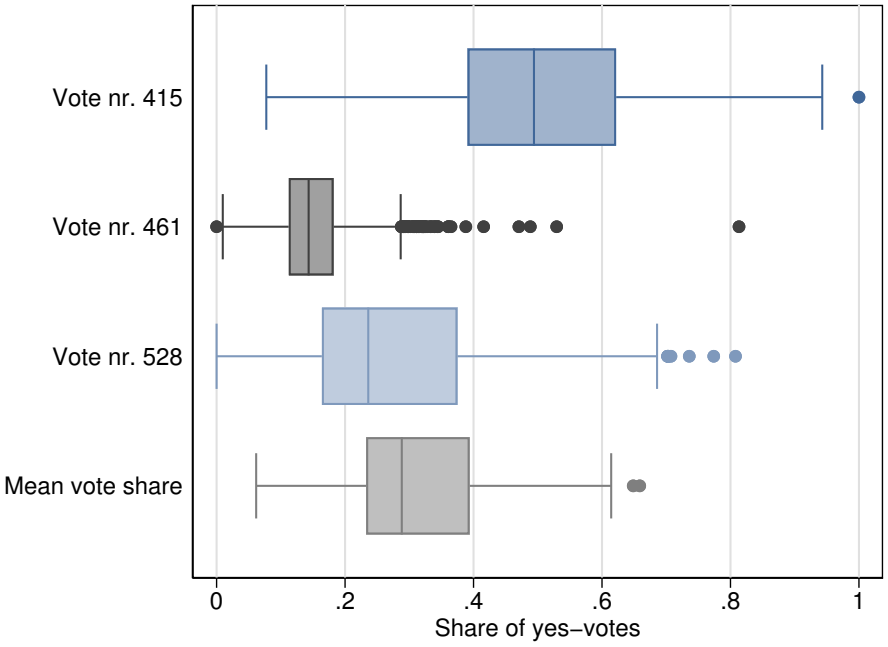
	Number of years	
Mean	6.788	6.788
Standard deviation	2.071	2.071
$N_{j[i]}^{\text{VET}}$	-0.663^{***}	-0.861^{*}
	(0.180)	(0.432)
Survey-year dummies	Yes	Yes
Firm-level controls	Yes	Yes
Cantonal dummies	Yes	Yes
Community-level controls	Yes	Yes
Number of observations	7,075	7,075
Pseudo R-Squared	0.001	0.030

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by communities.

Figure A.1: Variation in the voting results across communities



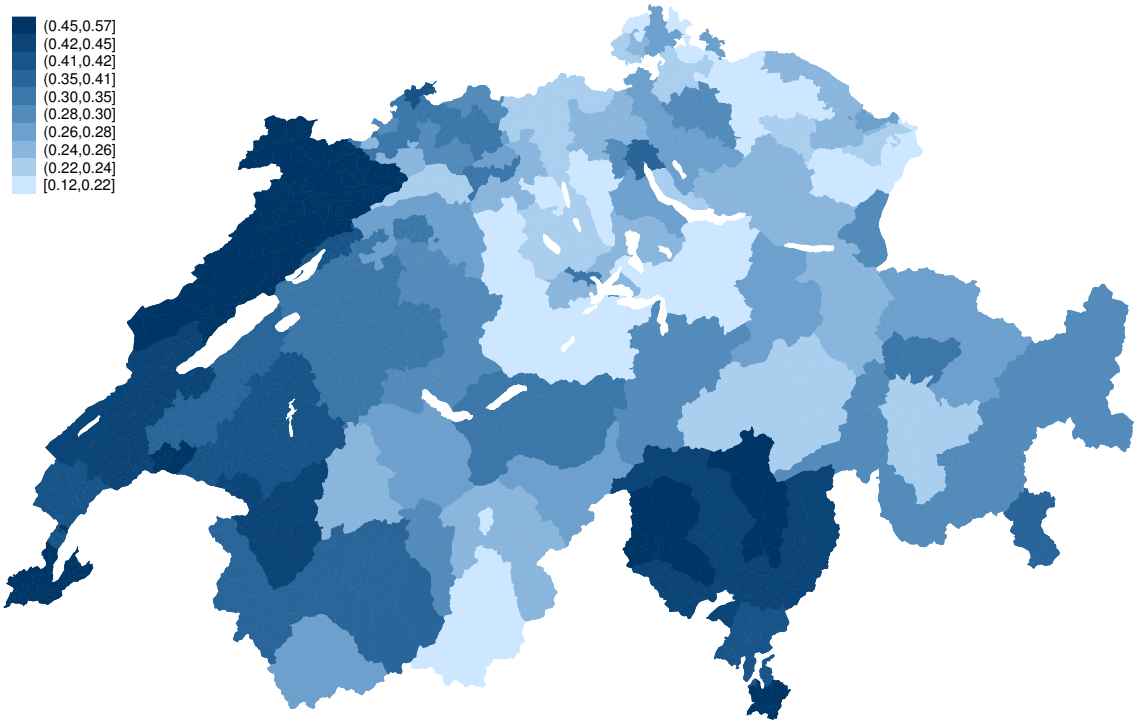
(a) Past votes on the role of the state in the provision/financing of VET



(b) Votes on the more general role of the state

Notes: The figure shows the amount of cross-sectional variation across communities in the voting results, for each vote separately as well as for the corresponding mean vote shares. See table 1 for additional details concerning the votes.

Figure A.2: Spatial variation in public attitudes towards the role of the state outside the educational context



Notes: The figure shows the spatial distribution of district-level voting results that measure attitudes towards the role of the state (mean share of supporting votes across votes nr. 415, 461, and 528). See table 1 for additional details concerning the votes.