



# Local norms describing the role of the state and the private provision of training

Andreas Kuhn<sup>a,b,c,\*,\*\*</sup>, Jürg Schweri<sup>a,b</sup>, Stefan C. Wolter<sup>b,d,e,c</sup>

<sup>a</sup> Swiss Federal University for Vocational Education and Training, Switzerland

<sup>b</sup> University of Bern, Switzerland

<sup>c</sup> IZA, Germany

<sup>d</sup> Swiss Coordination Centre for Research in Education, Switzerland

<sup>e</sup> CESifo, Germany

## ARTICLE INFO

### JEL classification:

D22  
D63  
H41  
I22  
J24

### Keywords:

Public goods  
Private provision of training  
Social norms  
Normative attitudes towards the role of the state  
Apprenticeship training

## ABSTRACT

Apprenticeship systems are essentially based on the voluntary participation of firms that provide, and usually also finance, training positions, often incurring considerable net training costs. One potential, yet under-researched explanation for this behavior is that firms act in accordance with the norms and expectations they face in the local labor market in which they operate. In this paper, we focus on the Swiss apprenticeship system and ask whether local norms towards the private, rather than the public, provision of training influence firms' decisions to offer apprenticeship positions. In line with this hypothesis, we find that the training incidence is higher in communities characterized by a stronger norm towards the private provision of training, which we measure using local results from two national-level plebiscites that explicitly dealt with the role of the state in the context of the apprenticeship system. This finding turns out to be robust to a series of alternative specifications and robustness checks.

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

## 1. Introduction

It has since long been argued that social norms may have the power to enforce and sustain the private provision of socially desirable goods. And indeed, ample experimental evidence on the effects of social norms on contributions to public goods has accumulated to date that appears to be broadly consistent with this idea.<sup>1</sup> Nonetheless, it has proven notoriously difficult to come

\* Correspondence to: Swiss Federal University for Vocational Education and Training, Kirchlindachstrasse 79, 3052 Zollikofen, Switzerland.

E-mail address: [andreas.kuhn@ehb.swiss](mailto:andreas.kuhn@ehb.swiss) (A. Kuhn).

\*\* We thank Manuel Aepli, Uschi Backes-Gellner, Giuliano Bonoli, Jörg Markowitsch, Sandra McNally and Ludger Woessmann, three anonymous referees, as well as conference participants in Barcelona, Berlin, Bern, Leuven, London, Manchester, Munich and Washington for many helpful comments and suggestions. We also thank Mirjam Strupler for compiling the survey data 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; Marcel Hageman for great research assistance; and Sally Gschwend-Fisher for proofreading the manuscript. Financial support by the Swiss State Secretariat for Education, Research and Innovation (SERI), Switzerland is gratefully acknowledged (contract number 1315000817).

<sup>1</sup> The experimental literature on this issue is too voluminous to give it due credit here (cf. Chaudhuri, 2011; Ledyard, 1995). Nonetheless, two results from this literature are worth mentioning. First, contributions to the public good tend to be larger in environments where the same individuals interact repeatedly with one another, as is the case in many real-world situations. Second, the possibility of punishing defectors increases average contributions (e.g. Fehr and Fischbacher, 2004) and, perhaps more importantly to us, internalized norms may have similar "power" in sustaining cooperation in public-good situations (Choi and Ahn, 2003; Dugar, 2010; Rege, 2004; Samek and Sheremeta, 2014).

<https://doi.org/10.1016/j.ejpeco.2022.102226>

Received 5 October 2021; Received in revised form 20 February 2022; Accepted 1 April 2022

Available online 9 April 2022

0176-2680/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

up with relevant real-world examples where such goods are provided privately, let alone examples where social norms have been important in setting up and/or sustaining a private solution on a larger scale.<sup>2</sup>

In this paper, we argue that apprenticeship systems, primarily in place in several European countries, provide an interesting and compelling real-world case for studying the impact of social norms on the private provision of education and training, an issue of considerable academic and practical relevance.<sup>3</sup> A first key feature of these systems is that mostly privately run companies provide, and usually also finance, these training positions on a fully voluntary basis. The other important feature of the system is that the human capital acquired during apprenticeship training is mostly occupation-specific and thus transferable across different employers. Therefore, non-training firms may also profit from the investments made by the training firms. Moreover, the society at large also benefits from the system because of all the possible positive externalities from providing youngsters with training which is in demand on the labor market. While empirical research on the subject has shown that the training firms not only incur costs, but also reap various benefits associated with training apprentices, thereby explaining part of the appeal to them, we believe that these conventional explanations are not sufficient to fully comprehend these often complex and historically grown systems.

Among the countries where these systems are prevalent, Switzerland is particularly well suited for studying the effect of social norms on firms' provision of apprenticeship positions (additional information on the institutional setup is given in Section 2 below). First, Switzerland has an exceptionally high share of youth participating in firm-based apprenticeships. About sixty percent of the most recent cohorts of adolescents enter firm-based apprenticeship training at the upper-secondary level, rendering apprenticeship training indeed a key pillar of the Swiss educational system. A second, not well-known feature calling for explanation is that, even within Switzerland, firms' training incidence differs substantially between regions. Training incidence increases from the western to the eastern part of the country, and the lowest share of training firms is found in the French language region located in western Switzerland, but the regional variation is high even within the French and German language regions (this variation in the regional training incidence is documented empirically in Section 5.1 below). This feature is hard to explain referring only to arguments based on the costs and benefits of apprenticeship training because both the Swiss system of vocational education and training and the labor market in general are primarily regulated at the national level. Moreover, wage levels have an ambiguous effect on firms' training incidence because they affect both costs and benefits from training apprentices, and thus the corresponding variation in wage levels probably cannot explain the regional pattern in the training probability.

Against this background, we hypothesize that a stronger local norm favoring the private – rather than the public – provision of public goods, or other socially desirable goods, increases firms' training incidence in Switzerland.<sup>4</sup> In that context, we believe that there are two main conceivable mechanisms through which social norms could influence individual behavior (e.g. Legros and Cislighi, 2020; Gross and Vostroknutov, 2022). First, external pressure may enforce compliance with the norm, and this could be especially relevant if acting in accordance with the norm imposes costs or only uncertain benefits. Consumers or other employers might enforce the norm by sanctioning non-training firms. For example, consumers might consciously prefer a store that trains apprentices over another that does not.<sup>5</sup> Second, and especially over longer time frames, individuals may internalize prevailing norms, e.g. through socialization. In the case where social norms have been internalized, there is no or less need for external enforcement of the norm, even if there are costs associated with acting in concordance with the norm. Given that the apprenticeship system has long historical roots (e.g. Berner and Gonon, 2016; Wettstein et al., 2017), we believe that at least a partial norm internalization is plausible in our setup. Anecdotal evidence appears to be in line with this argument (though we are aware that other explanations may be consistent with this behavior as well). For example, many 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, obviously being proud of their successful apprentices, or by placing a vignette signaling their training status on their entrance door or their shop window (the vignette is shown in appendix figure A.1).

To identify the effect of social norms on firms' training behavior, we further take advantage of the fact that Swiss citizens' are regularly asked to express their preferences on federal laws and amendments to the Swiss constitution in the voting booth. Specifically, two popular plebiscites in 1986 and 2003 asked for an amendment to the constitution stipulating a stronger involvement of the state in the provision of training positions. 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 positions. We merge the voting results with firm-level information on the provision of apprenticeship positions. The data on firms' willingness to provide apprenticeships comes from three national surveys on firms' training costs and benefits that were carried out in 2000, 2004, and 2009, respectively. These data contain detailed information about firms' training behavior along with some important firm-level characteristics.

<sup>2</sup> Cowen (1992) provides some real-world examples of privately provided public goods, and social norms have been shown to be relevant in the context of charitable giving (e.g. Shang and Croson, 2009) or energy consumption (e.g. Allcott, 2011). Another example that has received considerable attention from economic historians are turnpike road systems (e.g. Klein, 1990).

<sup>3</sup> Sadowski (2001) uses a similar conceptualization of vocational education and training which is, moreover, close to descriptions often used in comparative political science (e.g. Busemeyer et al., 2011). The notion of "community governance" (Bowles and Gintis, 2002) is another potentially useful conception of these systems.

<sup>4</sup> The concept of "civic virtue" (e.g. Algan and Cahuc, 2009) is very close to what we have in mind here. We prefer the conceptualization as a local social norm because it explicitly refers to the underlying mechanisms and because it fits neatly with our measurement based on local voting results (see Section 3 below).

<sup>5</sup> However, subtle mechanisms of expressing (dis)approval are presumably much more likely than explicit statements. Indeed, behavioral research shows that simply reminding people that there is a norm related to some behavior suffices; this indicates that social norms may work at a very subliminal level (see, for example, Agerström et al., 2016; Riyanto and Zhang, 2013; Pruckner and Sausgruber, 2013).

Our study also contributes to a growing body of evidence documenting the various effects of social norms on individual and corporate behavior, above and beyond their potential impact on the private provision of public goods. At the individual level, for example, social norms have been shown to influence such diverse individual behavior as fertility (Fernández and Fogli, 2006) and female labor supply (Fernández et al., 2004), tipping (Azar, 2004), investments into “sin stocks” (Hong and Kacperczyk, 2009), or political participation (Alesina and Giuliano, 2011).<sup>6</sup> Several studies using Swiss data have used voting results to measure attitudes or norms, documenting the effect of work attitudes on job search behavior (Stutzer and Lalive, 2004; Eugster et al., 2011) or from local gender norms on female well-being (Lalive and Stutzer, 2010). Evidence on the importance of social norms on corporate behavior is considerably rarer and focuses mainly on corporate social responsibility (e.g. Schmitz and Schrader, 2015). One exception is a 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. Another is the analysis by Janssen et al. (2016), who argue that the gender pay gap within firms is larger when local norms towards gender equality are weaker.

The remainder of this paper is organized as follows. The next section presents some important background information on the institutional setting, focusing on the key features of the Swiss apprenticeship system. Section 3 discusses the different data sources used in our empirical analysis, primarily 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 estimation 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. Section 6 concludes.

## 2. Institutional background

This section provides some background information on the Swiss educational system and the institutional setting of the Swiss apprenticeship system.<sup>7</sup>

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

The Swiss educational system is first and foremost characterized by its 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, 2019). The remainder mainly chooses further general education (taught at a “Gymnasium”) that prepares for, and grants access to, university studies. VET is thus by far the most often chosen educational track at the upper secondary level in Switzerland. It is fully integrated into Switzerland’s formal educational system, and there are several possibilities for entering into tertiary education with a completed apprenticeship (see appendix figure A.2).

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 four years, depending on the occupation learned.<sup>8</sup> During their training, apprentices spend most of their time in their training firms, 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). The employer pays the apprentices’ wages, but their wages are considerably lower than those of fully trained workers in the same occupation, even taking their lower productivity into account, which implies that apprentices share the costs of training with their employers.

### 2.2. 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 some public employers are, at least implicitly, expected to train apprentices (e.g. hospitals training nurses). Furthermore, there is no explicit regulation of wages paid to apprentices. Thus, the number of apprenticeship positions, both within specific training occupations and in the aggregate, is largely determined by the interaction of employers’ supply of and adolescents’ demand for the corresponding training positions.

Moreover, various formal associations and informal cooperation agreements 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 [organizations of the world of work]”) are not only responsible for the (further) development of the training curricula, they also prepare and implement the final practical examinations, which are decisive for completing the programs, on behalf of the cantons. Moreover, 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 skills on the labor market). Training activity by Swiss firms is thus embedded in a system of community governance that arguably supports adherence to established norms (Bowles and Gintis, 2002).

<sup>6</sup> Several studies have explored the reverse channel as well, i.e. the impact of the economic environment on one’s attitudes or preferences (e.g. Giuliano and Spilimbergo, 2014; La Ferrara et al., 2012), suggesting that reverse causality may be a relevant issue (an issue we will therefore take up again in Section 5.3).

<sup>7</sup> More information about the Swiss VET system, its historical roots, 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.

<sup>8</sup> Among these, about 91% enter a dual apprenticeship which combines practical training in a firm with vocational school. The remainder attends full-time school-based VET program (which is possible for selected occupations only).

### *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 costs accruing from apprenticeship training (within the firm) are almost fully borne by the firms actually providing the training positions and by the apprentices. Estimations based on survey data suggest that employers incur some CHF 2.7 billion of direct training costs per year and spend almost as much on apprentices' wages. In total, firms spend yearly nearly 1% of GDP on apprenticeship training.<sup>9</sup> In contrast, however, vocational schooling is almost fully funded publicly (by both the federal government and the cantons). The costs of vocational schooling amount to about 2.5 billion per year, according to official statistics (SERI, 2019).

### *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, quite in contrast, empirical evidence suggests that a substantial part of the human capital acquired through apprenticeship training in 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 in the labor market may explain the high fraction of training firms in Switzerland.<sup>10</sup>

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

In addition, there are several institutional features in place that explicitly aim to ensure that mobility across employers is possible for apprentices after the completion of their training (for example, there are centralized examinations at the end of the apprenticeship and the federal administration provides an external certification of the competencies acquired during the apprenticeship; 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 otherwise needs to be done by a trained worker) to a lower cost than when performed by a fully trained worker (see Aepli and Kuhn, 2021, on the potential substitution between training apprentices and hiring trained workers by Swiss employers).

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) 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).<sup>11</sup>

At the same time, however, many training firms incur substantial net costs from training apprentices.<sup>12</sup> Moreover, even if a training firm covers its costs by the end of training, it has no guarantee of this when it hires a new apprentice 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 or from the relatively high number of premature dropouts from apprenticeship training (e.g. Forsblom et al., 2016).

## **3. Data and key variables**

### *3.1. Firm-level survey data*

Our first data source are three consecutive surveys that were specifically designed to elicit detailed information about the costs and benefits of apprenticeship training from the point of view of the employers. The surveys were administered in the years 2000, 2004, and 2009, respectively (see Strupler and Wolter, 2012, for details and 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

<sup>9</sup> These estimates are based on the same firm-level survey data that we use in this paper (see Section 3 below for details).

<sup>10</sup> Consistent with this, comparisons between Switzerland and Germany (e.g. Muehleemann et al., 2010) and between Switzerland and Austria (Moretti et al., 2019) argue that corresponding differences in labor market regulation partially explain the observed differences in the net benefits to employers.

<sup>11</sup> It has further been shown that training firms may save recruiting costs that they would otherwise have to spend if they (are able to) retain apprentices who have completed their training (e.g. Blatter et al., 2012, 2016). Similarly, apprenticeship training may also serve as a (costly) screening device for employers (Mohrenweiser et al., 2020) or a (costly as well) signaling device (Backes-Gellner and Tuor, 2010).

<sup>12</sup> Net benefits are typically negative for the more technical and the more demanding apprenticeships (e.g. a mechanical engineer ("Polymechaniker"), a highly-skilled mechanic involved in the manufacturing of machinery, tools, and prototypes, among other things).

representative of almost the entire population of firms in Switzerland.<sup>13</sup> While it is possible for the same firm to appear more than once in the combined data because it might have 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 standard errors clustered at the municipality level throughout, however, we take this feature of the data into account with regard to statistical inference.

Because all three surveys cover both training and non-training firms, and because we know whether 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 employers' assessment of their motives for offering apprenticeship positions (see [Muehleman and Wolter, 2014](#), for an overview). The richness of information available in the data allows us to implement an empirical test on norm internalization by the employers (see Section 5.4 below).

A final 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 business census (see appendix B for additional details).

### 3.2. Community-level voting results and local norms describing the role of the state

As one of the main pillars of the direct-democratic political system of Switzerland, citizens are regularly asked to cast their votes 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 for using national-level voting results related to VET policy to measure individuals' normative attitudes towards the role of the state in this domain in a consistent way across all of Switzerland.

#### *Using disaggregated voting results to measure regional norms towards the role of the state: advantages and disadvantages*

Voting results are, first of all, a direct measure of voters' attitudes towards specific policy issues, and we believe that the use of the voting results has some distinct advantages compared to the use of attitudinal survey data. One important advantage of the use of voting results is that the outcome of a given vote has real consequences, and thus voters have a comparatively strong incentive to reveal their true preferences. In contrast, corresponding survey questions necessarily remain hypothetical, providing less incentive for 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, by 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 attitudes towards the role of the state in a convincing yet relatively straightforward way. Nonetheless, we acknowledge that there is expressive voting (e.g. [Hillman, 2010](#); [Schnellenbach and Schubert, 2015](#)) as well, which potentially complicates the interpretation of the voting results (we will come back to this issue in Section 5.2 below).

Moreover, we will interpret differences in the voting results across communities as reflecting primarily spatial differences in the local norm towards the private – rather than public – provision of training in the context of our study.<sup>14</sup> Indeed, aggregate-level voting results fulfill the two conditions noted by [Brennan et al. \(2013\)](#) for the existence of a (social) norm.<sup>15</sup> First, a significant fraction of individuals within a given community must hold a certain normative attitude towards a given subject. Second, people within a community must be aware that a shared norm exists in the community in which they live. The voting data that we use in our empirical analysis fit this definition closely – they directly measure the fraction of people sharing a given normative attitude on a specific subject. 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 to everyone else. However, it is also true that other factors beyond preferences related to the issue at hand, generate variation in the voting results, such as the economic environment in which a vote is held (again, we will come back to this issue in Section 5.2).

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. appendix B for details). A more subtle issue is the possibility that norms at a higher aggregation level may also be relevant because employers tend to associate with each other in various forms, as discussed in Section 2.2 (see also Section 4 below on how this could be relevant for our empirical analysis).

<sup>13</sup> 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, excluding the very smallest firms and employers from the primary sector (which were 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 \(2011\)](#).

<sup>14</sup> 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.

<sup>15</sup> Different definitions of social norms are abundant but are, for the most part, close or identical to the definition of [Brennan et al. \(2013\)](#) that we use in this paper. For example, [Fehr and Fischbacher \(2004\)](#) define social norms as “(...) standards of behavior that are based on widely shared beliefs how individual group members ought to behave in a given situation”.

**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
<i>(a) Votes about vocational education and training</i>					
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%
<i>(b) Other votes on the provision of public goods or demanding more public intervention</i>					
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	18.0%	42.2%
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 Administration. 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.

	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
$N_{j i}^{VET}$	-0.498*** (0.098) [-0.389]	-0.438*** (0.085) [-0.343]	-0.551*** (0.059) [-0.431]	-0.280*** (0.051) [-0.219]	-0.393*** (0.087) [-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. Approximate elasticities, evaluated at mean values, are given in brackets.

**Table 3**

Robustness checks.

	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
$N_{j i}^{VET}$	-0.342*** (0.089) [-0.267]	-0.227** (0.089) [-0.178]	-0.230** (0.092) [-0.180]	-0.322*** (0.123) [-0.213]	-0.503*** (0.115) [-0.397]	-0.449** (0.174) [-0.407]	-0.381*** (0.089) [-0.330]	-0.275*** (0.084) [-0.317]	-0.309** (0.122) [-0.289]	-1.344*** (0.334) [-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. Approximate elasticities, evaluated at mean values, are given in brackets.

A final 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. 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 excluded from voting because they lack Swiss citizenship.<sup>16</sup>

<sup>16</sup> 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 (we also checked that the interaction term between our measure of the norm towards the role of the state and voter turnout is insignificant; results not shown).

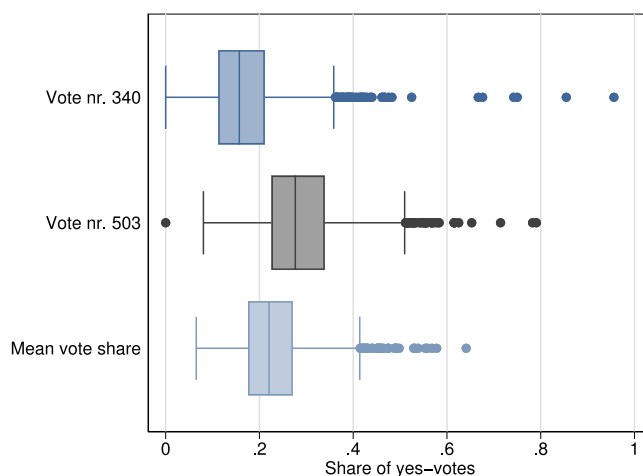


Fig. 1. Variation in the share of supporting votes, votes nr. 340 and nr. 503. Notes: The figure shows the distribution of the share of supporting votes for vote nr. 340 and vote nr. 503, as well as the mean across the two votes (see Table 1 for details).

#### Votes about the allocation of responsibilities within the VET system

Based on the above considerations, we therefore use municipality-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” [“Plebiscite for an adequate provision of vocational training”]), held on May 18, 2003; the second vote was also a popular initiative (“Initiative für eine gesicherte Berufsbildung und Umschulung” [“Plebiscite for a guaranteed vocational training and retraining”]), held on September 28, 1986. Both initiatives aimed to increase the public involvement regarding the provision of vocational education and training, and both initiatives were 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 in the Swiss apprenticeship system: both votes demanded that more school-based apprenticeships should be created, to be financed by both employers and the taxpayers.<sup>17</sup>

Panel (a) of Table 1 lists a few key figures for the two votes. Both initiatives were ultimately clearly rejected, 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 for its support, the 2003 vote captured about 31.6% of all votes cast. The last column shows that the overall turnout was 34.8% and 49.6%, respectively, which appears relatively low. However, note that the average turnout across all national-level votes from the 1980s through to and including the 2000s was about 43% only.<sup>18</sup> Given that the two votes made very similar demands, one may hypothesize that the higher turnout in the later vote is at least partially related to voters anticipating that the result could be less unequivocal. Similar to the voting results, there is large variation in turnout across the different municipalities, i.e. the mean turnout across the two votes from panel (a) of Table 1 varies between 18.2% and 86.2% (as mentioned in footnote 16, note that we include turnout as a control variable in most specifications).

As illustrated in Fig. 1, moreover, there was considerable variation in the share of votes in favor of each of the two initiatives across different municipalities. Municipality-level vote shares from the 1986 vote (the 2003 vote) vary between 0% and about 95% (between 0% and 79%). Not surprisingly, mean vote shares (i.e. municipality-level vote shares aggregated across the two votes) 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% (appendix figure A.3 shows the close correlation between the two voting results).

#### Votes about the role of the state beyond educational policy

Panel (b) of Table 1 lists three additional plebiscites that dealt with the provision of public goods or the role of the state more generally (i.e. these votes were 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 was accepted with a close majority of the votes (51.8%) in its favor. Each of the three votes clearly demanded more responsibilities for the state and in each case the vote took place relatively close in time to the firm-level surveys. 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, i.e. we can use these additional voting results to construct alternative parameterizations of the local norm describing the role of

<sup>17</sup> The full text of the two initiatives is available online (in German, French or Italian, but not in English) at: <https://www.bk.admin.ch/ch/d/pore/vi/vis150t.html> (vote nr. 340 from September 1986) and <https://www.bk.admin.ch/ch/d/pore/vi/vis284t.html> (vote nr. 503 from May 2003).

<sup>18</sup> See <https://www.bfs.admin.ch/bfs/de/home/statistiken/politik/abstimmungen/stimmbeteiligung.html>.

the state (see Section 5.2 below). Again, there is considerable spatial variation in the mean vote share across municipalities, with values ranging from a minimum of about 6% to a maximum of 66% (cf. appendix figures A.4 and A.5).

### 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 some regional-level characteristics. These variables will be used as control variables in the empirical analysis below, at different levels of regional aggregation (either at the municipality level or at the level of local labor markets). 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 municipality. 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 (descriptives are shown in appendix table A.1).

## 4. Estimation framework

Our empirical analysis consists of a series of linear regressions, mostly fit by OLS. In the main part of the analysis, we primarily focus on eliminating concerns related to unobserved heterogeneity. That is, our main goal is to include the essential confounding variables at both the firm- and the regional-level (in Section 5.3 below, we will discuss some additional analyses which focus on the potential bias due to simultaneity). More specifically, our baseline models all take on the following form:

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

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).<sup>19</sup> The regressor of primary interest is given by  $N_{j[i]}^{\text{VET}}$ , which denotes to the local share of votes supporting more activity of the state in the provision of training in community  $j$  in which firm  $i$  is located, and thus reflects normative attitudes towards the private provision of vocational education and training (as discussed in detail in Section 3 above). Parameter  $\beta$  is the main target of our empirical analysis because it quantifies the partial effect of social norms on individual firms' training behavior, at least under appropriate conditions. Because we hypothesize that stronger norms towards the private 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  $\beta$  to be negative.

In order to be able to identify  $\beta$ , our parameter of primary interest, several assumptions must hold. First and foremost, we must be able to control for all of the relevant confounding variables. Given that the dependent variable is measured at the firm level and the main regressor at the municipal 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 the number or structure of firms across communities, and these regional features could be predictive of whether an individual employer trains apprentices or not, independent of any regional differences in normative attitudes towards the role of the state. All additional variables are therefore used as controls for potential confounders when estimating  $\beta$ , and are thus of no (or only minor) direct interest. Eq. (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  $\phi_{l[i]}$ , 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.<sup>20</sup>

In addition, we must also assume that the local norm towards the role of the state may affect firms' training behavior, but at the same time rule out an effect running in the other direction – an assumption that may be violated in our context (for example, experiencing a successful apprenticeship training might instill or reinforce a positive attitude towards the role of private companies in the training of young people; indeed, such feedback effects may in fact underlie the process of norm internalization). We will try to approach this difficult problem using two different sets of instruments, as discussed in more detail in Section 5.3 below.

Moreover, we must also assume that there is no measurement error in using the observed voting results as a measure of the latent norm towards the role of the state, an issue that we take up in Section 5.2, and that the effect of the norm works mainly at the communal level and not at higher regional levels. This could be relevant because we include cantonal dummies in most of our specifications, as mentioned above. These do not only control for unobserved confounders at this regional level, they implicitly also eliminate variation in the local norm which potentially affects employers' training decision. This should be taken into account when interpreting the corresponding estimates.

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). We therefore

<sup>19</sup> 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 using the linear probability model because of its straightforward interpretation, but we also show average marginal effects from a probit model in column 10 of Table 3.

<sup>20</sup> 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 Eq. (1) against  $t$ , but only the survey-year fixed effects  $\phi_{l[i]}$ .



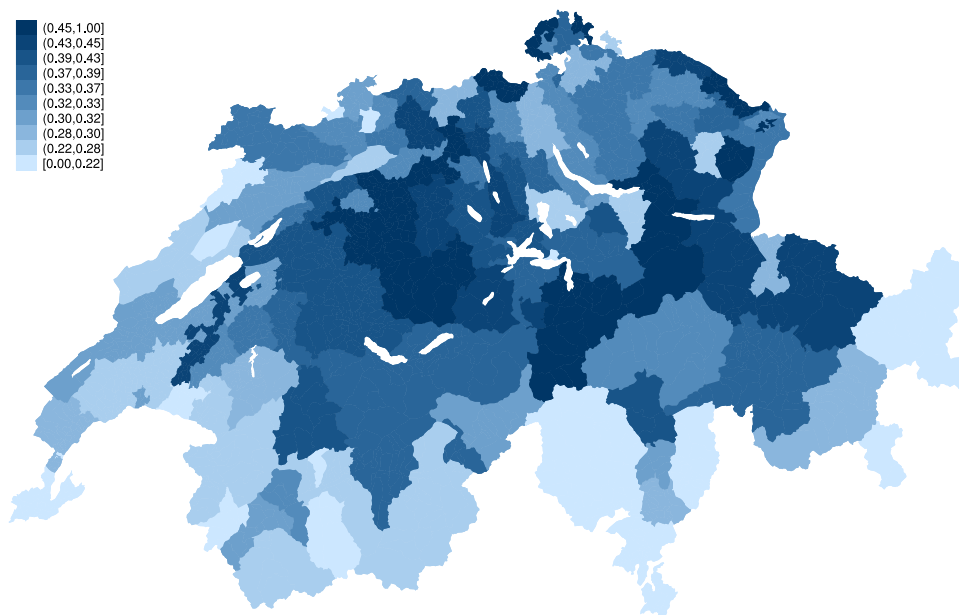


Fig. 2. Spatial variation in the incidence of training. Notes: The figure shows the spatial distribution of training incidence (i.e. the local mean of  $T_i$ ) across the 148 distinct districts of Switzerland. Darker shaded areas have a higher fraction of training firms.

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).

## 5. Results

### 5.1. Descriptive evidence

Starting with some descriptive evidence, Fig. 2 first illustrates how the training incidence among firms varies across regions within Switzerland. The figure documents that there is 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 even more remarkable that there appears to be systematic variation in the training incidence across regions. Specifically, the training incidence is considerably higher in the German language region of Switzerland than in both the French and Italian language regions (which are, respectively, located in the western and the southern parts of the country). It further appears that the regional training incidence among employers is higher in the more rural than in the more urban areas.

Analogously, Fig. 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 normative attitudes towards the role of the state as well, and 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 language areas of Switzerland than in the German language area, consistent with the findings of Eugster et al. (2011), for example.<sup>21</sup> Further, it appears that voters in the more urban regions have more favorable attitudes towards the role of state than those in the more rural areas. Overall, it thus appears that the pattern in Fig. 3 broadly mirrors that from Fig. 2.

Combined, the two figures therefore imply that we should find a pronounced 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 Fig. 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.

<sup>21</sup> This, in turn, opens up the possibility to estimate the eventual effect of such norms on employers' training behavior by focusing on firms located along the Swiss language border (Aepli et al., 2021).

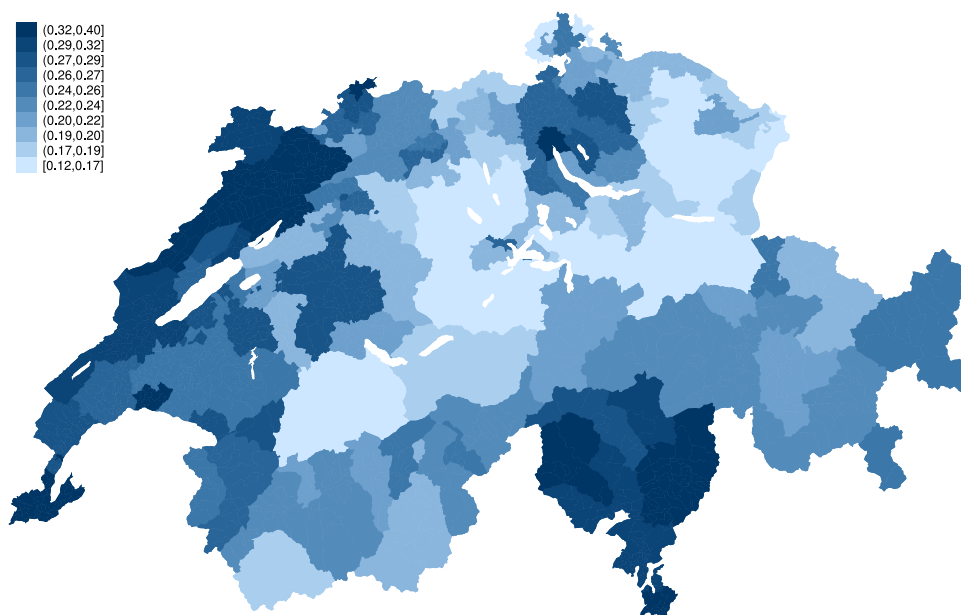


Fig. 3. Spatial variation in public attitudes towards the role of the state in apprenticeship training. Notes: The figure shows the spatial distribution of district-level voting results (i.e. the mean share of supporting votes from vote nr. 340 and vote nr. 503); see Table 1 for additional information concerning the three votes. Darker shaded areas are characterized by a weaker norm towards the private provision of training.

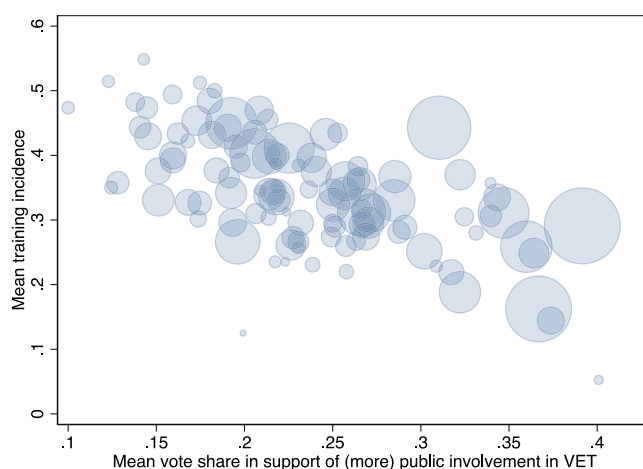


Fig. 4. The association between training incidence and public attitudes towards the role of the state. Notes: The figure plots the association between the local training incidence (on the y-axis) and public attitudes towards the role of the state within the VET system (on the x-axis). Both variables are aggregated up to the level of local labor markets, guaranteeing that the local training incidence is strictly larger than 0 and strictly 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 of firms.

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 is also robust to the inclusion of additional control variables.

## 5.2. Main 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.

The point estimate of  $\beta$  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  $\hat{\beta} = -0.497$ , confirming the pattern from Fig. 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 elasticity 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  $\beta$ . This is because the sampling frame included a different fraction of non-training firms in the different waves of the survey. The resulting point estimate is thus only slightly smaller, and it remains large and statistically significant ( $\hat{\beta} = -0.438$ , with a robust standard error of about 0.085).

We next add, in the third column, the size of firm, its sector of activity, and ownership (private versus public and nonprofit employers).<sup>22</sup> This specification yields a point estimate that is somewhat stronger, i.e. more negative, than the estimates from the preceding two columns, and it remains highly significant ( $\hat{\beta} = -0.551$ , with a robust standard error of about 0.059). The comparison with the preceding 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 located in communities with a weaker norm towards the private provision of training have characteristics 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  $\hat{\beta} = -0.280$ . As expected, the inclusion of the fixed effects lowers the point estimates substantially – by about 50%, compared to the preceding column. This confirms our expectation that there is large variation in the training incidence across cantons that 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  $\beta$  remains substantively large as well as statistically significant, with a robust  $t$ -value of about 5.49.

Finally, in the fifth and final columns of Table 2, we further add some 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 age distribution in a given region), yielding an estimate of  $\hat{\beta} = -0.393$  with an associated robust standard error of about 0.087.<sup>23</sup> Similar to the inclusion of the firm-level controls, adding community-level controls makes the estimated point estimate associated with 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 for employers to provide apprenticeship positions.<sup>24</sup>

Our first set of estimates thus shows that firms which are located in regions characterized by a stronger norm towards the private provision of training are significantly and substantively 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 some robustness checks.

A first check is to include additional or more detailed controls at the regional level. Thus, the specification in column 1 includes some 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 market). This yields an estimate of  $\beta$  that is only marginally smaller (i.e. less negative) than our baseline estimate ( $\hat{\beta} = -0.342$ , 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).<sup>25</sup> Similar to column 1, these two specifications yield estimates of  $\beta$  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  $\hat{\beta} = -0.227$  ( $\hat{\beta} = -0.230$ ), 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. Figs. 2 and 3). We thus restrict the estimation sample, in column 4, to those communities from the German language region of Switzerland (which reduces the sample size to 15,706 observations). The resulting point estimate of  $\hat{\beta} = -0.322$  is very close to our baseline estimate, however. Our result is thus not simply driven by simultaneous differences in both training behavior and normative attitudes towards the role of the state between the different language regions

<sup>22</sup> To save space, we do not show the full regression results but these are, of course, available upon request.

<sup>23</sup> 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). See appendix table A.1.

<sup>24</sup> 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.

<sup>25</sup> There are 148 (106) distinct districts (local labor markets), but only 26 cantons; see appendix table B.1. Obviously, the more disaggregated fixed effects will not only pick up much of the variation in employers' training behavior due to unobserved regional characteristics, but a substantial fraction of the variation in the local norm as well.

**Table 4**  
Alternative parameterizations of the regional norm describing the role of the state.

	Training firm (yes = 1)			
Mean	0.335	0.335	0.335	0.335
Standard deviation	0.472	0.472	0.472	0.472
$N_{j[i]}^{VET}$	-0.393*** (0.087) [-0.307]			-0.482*** (0.139) [-0.377]
$N_{j[i]}^{STATE}$		-0.341*** (0.097) [-0.337]		
$N_{j[i]}^{TOTAL}$			-0.447*** (0.100) [-0.405]	
Estimation method	OLS	OLS	OLS	2SLS
Instrument	-	-	-	$N_{j[i]}^{STATE}$
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
R-Squared	0.326	0.326	0.326	0.326
Partial R-squared (first stage)	-	-	-	0.388
F-value (first stage)	-	-	-	438.770
p-value (endogeneity)	-	-	-	0.377

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. Approximate elasticities, evaluated at mean values, are given in brackets.  $N_j^{VET}$ ,  $N_j^{STATE}$ , and  $N_j^{TOTAL}$  all measure local norms describing the role of the state, using different votes (see main text for details).

within Switzerland.<sup>26</sup> Another potential issue is that we have only few firm-level observations in 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 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,  $\hat{\beta} = -0.449$ , is not statistically different from our baseline estimate. Reassuringly, column 7 further shows that the point estimate hardly changes when we focus on private employers only ( $\hat{\beta} = -0.381$ , 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).

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 survey 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.

#### Alternative parameterizations of the local norm describing the role of the state and measurement error in the voting results

We next show that our main results are also robust to the use of alternative parameterizations of the key explanatory variable (see also appendix figure A.6 for a comparison of the variation in the different parameterizations of the local norm).

For the ease of comparison, the first column of Table 4 replicates our baseline specification from column (5) of Table 2 above. In the second column of Table 4, we use the mean share of supporting votes across the three plebiscites from panel (b) of Table 1, denoted by  $N_{j[i]}^{STATE}$ . These three popular votes also dealt with the appropriate role of the state, but outside the realm of the apprenticeship system or educational policy more generally. This alternative parameterization of the local norm yields a point estimate of  $\hat{\beta} = -0.341$ , almost indistinguishable from the baseline estimate of column 1, both in absolute and in relative terms (indeed, the two estimates are not statistically different from each other). Finally, in the third column we use the mean share of supporting votes across all five plebiscites from Table 1, denoted by  $N_{j[i]}^{TOTAL}$ . In this case, the resulting point estimate of  $\beta$  is slightly larger than, but still close to our baseline specification ( $\hat{\beta} = -0.447$ , with a robust standard error of 0.100). In relative terms, however, the effect is even larger than in the baseline case (approximate elasticity of about  $-0.405$  versus  $-0.307$ ).

<sup>26</sup> Similarly, using only the French language regions yields a point estimate of  $\hat{\beta} = -0.383$  (not shown in Table 3). With a robust  $t$ -value of about  $-2.4$ , this estimate remains statistically significantly different from zero. At the same time, it is not statistically different from that derived in the German language regions. Using only the Italian language regions, however, yields an imprecise and statistically insignificant, and even positive point estimate of  $\beta$ . Thus, consequently, excluding the Italian language regions yields a larger estimate of  $\hat{\beta} = -0.412$  (with a robust standard error of 0.093).

Another potential issue is that the observable voting results are presumably only a noisy measure of the local norm, in the sense that variation in the voting results is only partially driven by the latent norm towards the role of the state and that other factors have also an impact on the voting results, such as differences in budget constraints across municipalities. Formally, one may conceptualize the observed voting results as being a function of both the “true” but unobserved norm towards the role of the state,  $N_j^*$ , and a second term which includes issue-specific variation as well as more idiosyncratic variation in the voting results:

$$N_j^{\text{VET}} = N_j^* + \epsilon_j^{\text{VET}}, \text{ and} \quad (2a)$$

$$N_j^{\text{STATE}} = N_j^* + \epsilon_j^{\text{STATE}}, \text{ respectively.} \quad (2b)$$

For example,  $\epsilon_j^{\text{VET}}$  and  $\epsilon_j^{\text{STATE}}$  may represent variation in the voting results due to the economic environment in which the vote was held as well as more idiosyncratic variation due to, for example, expressive voting among part of the electorate (as mentioned in Section 3.2). Given that we have multiple votes at disposal for the measurement of the local norm, we may use our measure based on the alternative votes,  $N_j^{\text{STATE}}$ , as an instrument for our baseline measure of the local norm describing the role of the state,  $N_j^{\text{VET}}$ . This procedure will isolate the effect stemming from the unobservable local norm as long as the two idiosyncratic terms, i.e.  $\epsilon_j^{\text{VET}}$  and  $\epsilon_j^{\text{STATE}}$ , are each independent from the latent norm as well as independent from each other (e.g. Wooldridge, 2010). The resulting 2SLS point estimate equals  $\hat{\beta} = -0.482$ , as shown in the fourth column of Table 4 (the fact that the 2SLS point estimate is somewhat larger than the baseline estimate from column 1 is consistent with the structure assumed above).<sup>27</sup> Not surprisingly, there is a strong and statistically significant first-stage effect, as shown at the bottom of the table (i.e. there is a large partial R-Squared of 0.388 and a large F-statistic of about 438 associated with the first-stage regression; see also appendix figure A.7). At the same time, the standard error associated with the instrumental-variable estimate is about 60% larger than that from the OLS estimate in column 1 (0.139 versus 0.087). In comparison, the 2SLS estimate is thus very close to and not significantly different from the corresponding OLS estimate. Overall, this additional estimate suggests that there is no obvious bias associated with directly using the voting results from the two popular initiatives on apprenticeship training as a measure of the local norm towards the role of the state.

### 5.3. Tackling simultaneity bias

One remaining, and potentially relevant issue with the estimates based on Eq. (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 may not only be an effect of the local norm towards the private provision of training on the probability of training, but that a high training probability may at the same time strengthen individuals' belief that training is best provided privately. In this part of the empirical analysis, we thus try to correct for potential simultaneity bias resulting from such issues using instrumental variables (e.g. Angrist and Pischke, 2008; Wooldridge, 2010).

Specifically, we propose to use two distinct sets of instruments, the first one based on specific geographic features of a municipality, the second one on historical election results. Both sets of instruments rest on the presumption that social norms tend to have deep roots in human history and culture (e.g. Bazzi et al., 2020; Becker et al., 2020; Ostrom, 2000; Henrich, 2020; Cardullo et al., 2022). As we will show, both sets of instruments have a reasonably strong first-stage effect, i.e. they both predict variation in local norms towards the role of the state, conditional on the full set of controls. At the same time, however, both sets of instruments might suffer, to some degree, from a violation of the exclusion restriction; reflecting the inherent problem of finding truly exogenous instruments (e.g. Bound et al., 1995). We therefore view the resulting estimates more as an additional robustness check rather than as strictly preferable to the OLS estimates (nonetheless, however, these estimates can still inform us about the sign of the bias induced by simultaneity between local norms and employers' training behavior).

A first idea is to use geographical features of the various municipalities to instrument for local norms describing the role of the state. Specifically, we use the traveling distance from a given municipality to the national border as well as the median elevation above sea level of the area of a municipality (see appendix figure A.8 for a visualization of the two instrumental variables). Together, the two variables describe the geographic remoteness of any given municipality. We expect that norms towards a more active role of private actors are more pronounced in the more remote areas, where people historically were, and perhaps still are, more dependent on each other. The resulting 2SLS estimate of  $\hat{\beta} = -1.413$  is negative, statistically significant, and turns out to be considerably larger (in absolute value) than the corresponding OLS estimate. Also, the two instruments are reasonably strong in predicting the endogenous variable, as shown by the relevant first-stage F-statistic of 15.317.

In column 3 we use election results from the year 1947 as instruments. More precisely, we use the district-level results from the 1947 elections of the National Council (“Nationalrat”), which is part of the national parliament (the data contain the results for the seven most popular political parties in that year, plus one remainder category; appendix figure A.9 plots the spatial variation in the election results for four selected political parties). The idea here is that the strength of the different political parties also reflects regional norms towards the role of the state, since one of the main differences across political parties is how they judge the role of the state vis-à-vis private actors, across various policy domains. As evident from column 3 of Table 5, this specification also yields

<sup>27</sup> In addition, a principal-components analysis using all five votes shows that the first component explains about 68% of the total variation in the voting results (it is also the only component with an eigenvalue larger than 1). Moreover, we find that there is almost a perfect correlation between this first principal component and the mean vote share across all five votes of about 0.995. This is consistent with a structure where the result of each single vote  $v$  is conceptualized as  $N_j^v = N_j^* + \epsilon_j^v$ , analogous to Eqs. (2a) and (2b), respectively.

**Table 5**  
Tackling simultaneity bias.

	Training firm (yes = 1)			
Mean	0.335	0.335	0.335	0.335
Standard deviation	0.472	0.472	0.472	0.472
$N_{j t}^{VET}$	-0.393*** (0.087) [-0.307]	-1.413*** (0.371) [-1.105]	-0.780*** (0.293) [-0.610]	-1.076*** (0.239) [-0.841]
Estimation method	OLS	2SLS	2SLS	2SLS
Instruments	–	$R_{j t}$	$E_{kt t}$	$R_{j t} \cdot E_{kt t}$
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
R-Squared	0.326	0.320	0.325	0.323
Partial R-squared (first stage)	–	0.057	0.074	0.118
F-value (first stage)	–	15.317	12.778	13.914
p-value (endogeneity)	–	0.002	0.179	0.002

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. Approximate 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.  $R_{j|t}$  and  $E_{kt|t}$  denote, respectively, the two instruments based on the geography of a municipality and the district-level vote shares from the 1947 elections of the National Council (“Nationalrat”), for the seven most popular political parties in that year plus the remainder for all remaining parties (i.e. there are eight variables).

a negative 2SLS estimate of  $\hat{\beta} = -0.780$ , with a robust standard error of about 0.293. Again, this estimate is highly significant both in terms of size and statistical significance, but it is not statistically different from the baseline OLS estimate (as evident from the last row of Table 5); even though the first-stage effect of the instruments is again strong (associated F-statistic of about 12.8).

The final column of Table 5 uses both sets of instruments at the same time, yielding a point estimate of  $\hat{\beta} = -1.076$ , with a robust standard error of about 0.239. Not surprisingly, given that the two sets of instruments have about equal first-stage strength, the point estimate in this case lies between the estimates from the preceding two columns (the first-stage F-statistic of 13.9 is of similar size as in the preceding two columns). Compared to the baseline OLS estimate, this estimate is about 2.7 times larger, yet much less precisely estimated (nonetheless, though, the two point estimates are significantly different from each other).

Taken together, the additional estimates from Table 5 yield a consistent pattern of estimates, all of them suggesting a negative effect of normative attitudes towards the role of the state on the likelihood of offering apprenticeship positions. At the same time, as already mentioned, both sets of instruments are potentially subject to a violation of the exclusion restriction, i.e. one could argue that there is also a direct effect on the dependent variable. Moreover, the instrumental-variable estimates are also quantitatively similar to our baseline OLS estimates, if the statistical uncertainty of the estimates is also taken into account, and we therefore stick with OLS for the remainder of our empirical analysis.

#### 5.4. Evidence on norm internalization

A final issue worth exploring is whether local norms towards the role of the state affect employers' self-perception regarding the motives for (not) providing apprenticeship positions. In fact, we believe that one can argue that these variables in part reflect the internalization of the norm by employers. Specifically, training firms were directly asked in the survey 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 provision of training, whereas we expect to find no corresponding effect in the case of economic motives.<sup>28</sup>

Table 6 reports the corresponding estimates, using both the minimal and the full specification that we already used above. In the first four columns, the dependent variable reflects the importance of noneconomic and economic motives among training firms, respectively.<sup>29</sup> The first two columns look at the importance of noneconomic motives for employers' training decision. There is a

<sup>28</sup> A related idea is to look at the number of years an employer has been training apprentices. Again, the likelihood that a given firm has been training apprentices for a longer time is higher in those regions where the norm towards the private provision of training is stronger (see appendix table A.3).

<sup>29</sup> In a first step, we constructed a dummy variable indicating consent with any 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. See appendix table A.4 for the full list of survey items.

**Table 6**  
Employers' self-perception of their motives for training apprentices.

	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
$N_{j d}^{VET}$	-0.245*** (0.067) [-0.086]	-0.191* (0.102) [-0.067]	-0.198*** (0.045) [-0.109]	-0.064 (0.108) [-0.035]	-0.305*** (0.035) [-0.195]	-0.055 (0.060) [-0.035]
Training firms only	Yes	Yes	Yes	Yes	No	No
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. Approximate elasticities, evaluated at mean values, are given in brackets.

statistically 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 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 certainly consistent with awareness of the norm on the part of the employers, and it is also consistent with norm internalization (see appendix table A.5 for further evidence on norm internalization).

Column 3 shows that there is also a negative and significant effect of normative 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. Note that this result is not driven by an excessive increase in the associated standard error (the increase is similar to the one observed in the first two columns), but rather by the shrinkage in the corresponding point estimate.

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 preceding columns, the underlying survey items were answered by both training and non-training firms (using a slightly different set of questions, however; see appendix table A.6). The resulting estimates mirror the result from the two preceding 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 association between social norms towards the role of the state and the private provision of training – a topic not only of academic but also of considerable public interest. Specifically, we combine firm-level survey data with municipality-level voting results from different plebiscites 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 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 public goods are, *ceteris-paribus*, more likely to provide 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 normative 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 firms' provision of training positions and the local norm towards the private provision of training turns out to be very robust to a wide variety of robustness checks and alternative model specifications. The negative association between the local norm and the regional training incidence is robust to the inclusion of a wide variety of firm- and regional-level controls. Moreover, we find a quantitatively similar, yet somewhat larger effect of the norm on employers' training behavior when correcting for simultaneity bias using different instrumental variables. Taken together, the resulting estimates are robust and consistent across a wide range of different specifications and robustness checks. Finally, we also show that training firms are more likely to state that non-economic motives are relevant for their training decision if they are located in a region with a stronger local norm towards the private provision of training, further corroborating our main argument. We thus conclude that our findings point to the importance of norms describing the role of the state as an important explanatory factor with regard to firms'

training decisions. At a very general level, these findings suggest that the successful, yet also highly complex, Swiss apprenticeship system is deeply rooted in its broader social environment (cf. Algan and Cahuc, 2009; Alesina et al., 2015). Our analysis of firm-based apprenticeship training in Switzerland illustrates that social norms, by influencing firms' behavior, can help maintain local equilibria in which public goods are provided privately.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Data availability

The authors do not have permission to share data.

### Supplementary material

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.ejpoleco.2022.102226>.

### References

- Acemoglu, D., Pischke, J.-S., 2000. Certification of training and training outcomes. *Eur. Econ. Rev.* 44 (4), 917–927.
- Aeppli, M., Kuhn, A., 2021. Open labor markets and firms' substitution between training apprentices and hiring workers. *Labour Econ.* 70, 101979.
- Aeppli, M., Kuhn, A., Schweri, J., 2021. Culture, norms, and the provision of training by employers: Evidence from the swiss language border. *Labour Econ.* 73, 102057.
- Agell, J., 1999. On the benefits from rigid labour markets: norms, market failures, and social insurance. *Econom. J.* 109 (453), 143–164.
- Agerström, J., Carlsson, R., Nicklasson, L., Guntell, L., 2016. Using descriptive social norms to increase charitable giving: The power of local norms. *J. Econ. Psychol.* 52, 147–153.
- Alesina, A., Algan, Y., Cahuc, P., Giuliano, P., 2015. Family values and the regulation of labor. *J. Eur. Econom. Assoc.* 13 (4), 599–630.
- Alesina, A., Giuliano, P., 2011. Family ties and political participation. *J. Eur. Econom. Assoc.* 9 (5), 817–839.
- Algan, Y., Cahuc, P., 2009. Civic virtue and labor market institutions. *Am. Econ. J. Macroecon.* 1 (1), 111–145.
- Allcott, H., 2011. Social norms and energy conservation. *J. Public Econ.* 95 (9–10), 1082–1095.
- Angrist, J.D., 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. *J. Econ. Behav. Organ.* 54 (1), 49–64.
- Backes-Gellner, U., Tuor, S.N., 2010. Avoiding labor shortages by employer signaling: on the importance of good work climate and labor relations. *Ind. Labor Relat. Rev.* 63 (2), 271–286.
- Bassanini, A., Brunello, G., Caroli, E., 2017. Not in my community: Social pressure and the geography of dismissals. *J. Labor Econ.* 35 (2), 429–483.
- Bazzi, S., Fiszbein, M., Gebresilasie, M., 2020. Frontier culture: The roots and persistence of rugged individualism in the United States. *Econometrica* 88 (6), 2329–2368.
- Becker, A., Enke, B., Falk, A., 2020. Ancient origins of the global variation in economic preferences. *AEA Pap. Proc.* 110, 319–323.
- Berner, E., Gonon, P., 2016. History of Vocational Education and Training in Europe: Cases, Concepts and Challenges. In: *Studies in Vocational and Continuing Education*, vol. 14, Peter Lang.
- Bertrand, M., Mullainathan, S., 2001. Do people mean what they say? Implications for subjective survey data. *Amer. Econ. Rev.* 91 (2), 67–72.
- Blatter, M., Muehleman, S., Schenker, S., 2012. The costs of hiring skilled workers. *Eur. Econ. Rev.* 56 (1), 20–35.
- Blatter, M., Muehleman, S., Schenker, S., Wolter, S.C., 2016. Hiring costs for skilled workers and the supply of firm-provided training. *Oxford Econ. Pap.* 68 (1), 238–257.
- Bound, J., Jaeger, D.A., Baker, R.M., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *J. Amer. Statist. Assoc.* 90 (430), 443–450.
- Bowles, S., 2016. *The Moral Economy. Why Good Incentives are No Substitute for Good Citizens*. Yale University Press.
- Bowles, S., Gintis, H., 2002. Social capital and community governance. *Econom. J.* 112 (483), F419–F436.
- Brennan, G., Eriksson, L., Goodin, R.E., Southwood, N., 2013. *Explaining Norms*. Oxford University Press.
- Busemeyer, M.R., Cattaneo, M.A., Wolter, S.C., 2011. Individual policy preferences for vocational versus academic education: Microlevel evidence for the case of Switzerland. *J. Eur. Soc. Policy* 21 (3), 253–273.
- Busemeyer, M.R., Trampusch, C., 2011. *The Political Economy of Collective Skill Formation*. Oxford University Press.
- Cameron, A.C., Miller, D.L., 2015. A practitioner's guide to cluster-robust inference. *J. Hum. Resour.* 50 (2), 317–372.
- Cardullo, G., Conti, M., Ricci, A., Scicchitano, S., Sulis, G., 2022. On the Emergence of Cooperative Industrial and Labor Relations. Working Paper 2022/01, Centro Ricerche Economiche Nord Sud.
- Chaudhuri, A., 2011. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Exp. Econ.* 14 (1), 47–83.
- Choi, J.-K., Ahn, T., 2013. Strategic reward and altruistic punishment support cooperation in a public goods game experiment. *J. Econ. Psychol.* 35, 17–30.
- Cowen, T., 1992. *Public Goods and Market Failures: A Critical Examination*. Transaction Publishers.
- Dugar, S., 2010. Nonmonetary sanctions and rewards in an experimental coordination game. *J. Econ. Behav. Organ.* 73 (3), 377–386.
- Eugster, B., Lalive, R., Steinhauer, A., Zweimüller, J., 2011. The demand for social insurance: Does culture matter? *Econom. J.* 121 (556), F413–F448.
- Eugster, B., Lalive, R., Steinhauer, A., Zweimüller, J., 2017. Culture, work attitudes, and job search: Evidence from the swiss language border. *J. Eur. Econom. Assoc.* 15 (5), 1056–1100.
- Fehr, E., Fischbacher, U., 2004. Social norms and human cooperation. *Trends Cogn. Sci.* 8 (4), 185–190.
- Fernández, R., Fogli, A., 2006. Fertility: The role of culture and family experience. *J. Eur. Econom. Assoc.* 4 (2–3), 552–561.
- Fernández, R., Fogli, A., Olivetti, C., 2004. Mothers and sons: Preference formation and female labor force dynamics. *Q. J. Econ.* 119 (4), 1249–1299.
- Forsblom, L., Negrini, L., Gurtner, J.-L., Schumann, S., 2016. Dropouts in swiss vocational education and the effect of training companies' trainee selection methods. *J. Vocat. Educ. Train.* 68 (4), 399–415.
- Giuliano, P., Spillimbergo, A., 2014. Growing up in a recession. *Rev. Econom. Stud.* 81 (2), 787–817.
- Gross, J., Vostroknutov, A., 2022. Why do people follow social norms? *Curr. Opin. Psychol.* 44, 1–6.
- Henrich, J., 2020. *The Weirdest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*. Allen Lane.



- Hillman, A.L., 2010. Expressive behavior in economics and politics. *Eur. J. Political Econ.* 26 (4), 403–418.
- Hong, H., Kacperczyk, M., 2009. The price of sin: The effects of social norms on markets. *J. Financ. Econ.* 93 (1), 15–36.
- Janssen, S., Sartore, S.T., Backes-Gellner, U., 2016. Discriminatory social attitudes and varying gender pay gaps within firms. *Ind. Labor Relat. Rev.* 69 (1), 253–279.
- Klein, D.B., 1990. The voluntary provision of public goods? The turnpike companies of early America. *Econ. Inq.* 28 (4), 788–812.
- La Ferrara, E., Chong, A., Duryea, S., 2012. Soap operas and fertility: Evidence from Brazil. *Am. Econ. J. Appl. Econ.* 4 (4), 1–31.
- Lalive, R., Stutzer, A., 2010. Approval of equal rights and gender differences in well-being. *J. Popul. Econ.* 23 (3), 933–962.
- Ledyard, J.O., 1995. Public goods: A survey of experimental research. In: Kagel, J., Toth, A. (Eds.), *Handbook of Experimental Economics*. Princeton University Press, pp. 111–193, chapter 2.
- Legros, S., Cislighi, B., 2020. Mapping the social-norms literature: An overview of reviews. *Perspect. Psychol. Sci.* 15 (1), 62–80.
- Mohrenweiser, J., Wydra-Somagio, G., Zwick, T., 2020. Information advantages of training employers despite credible training certificates. *Oxford Econ. Pap.* 72 (3), 651–671.
- Moretti, L., Mayerl, M., Muehleemann, S., Schlögl, P., Wolter, S., 2019. So similar and yet so different: A comparative analysis of a firm's cost and benefits of apprenticeship training in Austria and Switzerland. *Evidence-Based HRM* 7 (2), 229–246.
- Muehleemann, S., Pfeifer, H., Walden, G., Wenzelmann, F., Wolter, S.C., 2010. The financing of apprenticeship training in the light of labor market regulations. *Labour Econ.* 17 (5), 799–809.
- Muehleemann, S., Wolter, S.C., 2011. Firm-sponsored training and poaching externalities in regional labor markets. *Reg. Sci. Urban Econ.* 41 (6), 560–570.
- Muehleemann, S., Wolter, S.C., 2014. Return on investment of apprenticeship systems for enterprises: Evidence from cost-benefit analyses. *IZA J. Labor Policy* 3 (1), 25.
- Mueller, B., Schweri, J., 2015. How specific is apprenticeship training? Evidence from inter-firm and occupational mobility after graduation. *Oxford Econ. Pap.* 67 (4), 1057–1077.
- Ostrom, E., 2000. Collective action and the evolution of social norms. *J. Econ. Perspect.* 14 (3), 137–158.
- Potterat, J., 2011. Methodenbericht. In: *Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe im Jahr 2009 (KNBB09)*. Stichprobenplan, Gewichtung und Schätzverfahren. Bundesamt für Statistik (BFS), Neuchâtel.
- Pruckner, G.J., Sausgruber, R., 2013. Honesty on the streets: A field study on newspaper purchasing. *J. Eur. Econom. Assoc.* 11 (3), 661–679.
- Rege, M., 2004. Social norms and private provision of public goods. *J. Public Econ. Theory* 6 (1), 65–77.
- Riyanto, Y.E., Zhang, J., 2013. The impact of social comparison of ability on pro-social behaviour. *J. Soc.-Econ.* 47, 37–46.
- Sadowski, D., 2001. *Berufsbildung: Ein Allmendegut*. In: Badelt, C. (Ed.), *Aphorismen für den Sozio-ökonomischen Fortschritt - Gewidmet Werner Clement Zum 60. Geburtstag*. Peter Lang, Frankfurt/Main, pp. 181–188.
- Samek, A.S., Sheremeta, R.M., 2014. Recognizing contributors: an experiment on public goods. *Exp. Econ.* 17 (4), 673–690.
- Schmitz, J., Schrader, J., 2015. Corporate social responsibility: A microeconomic review of the literature. *J. Econ. Surv.* 29 (1), 27–45.
- Schnellenbach, J., Schubert, C., 2015. Behavioral political economy: A survey. *Eur. J. Political Econ.* 40, 395–417.
- SERI, 2019. Vocational and Professional Education and Training in Switzerland. Facts and Figures 2019. State Secretary for Education, Research and Innovation SERI.
- Shang, J., Croson, R., 2009. A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. *Econom. J.* 119 (540), 1422–1439.
- Strupler, M., Wolter, S.C., 2012. *Die Duale Lehre: Eine Erfolgsgeschichte – Auch für die Betriebe*. Rüegger Verlag, Glarus.
- Stutzer, A., Lalive, R., 2004. The role of social work norms in job searching and subjective well-being. *J. Eur. Econom. Assoc.* 2 (4), 696–719.
- Wettstein, E., Schmid, E., Gonon, P., 2017. *Swiss Vocational and Professional Education and Training (VPET). Forms, System, Stakeholders*. hep Verlag, Bern.
- Wolter, S.C., Ryan, P., 2011. Apprenticeship. *Handb. Econ. Educ.* 3, 521–576.
- Wooldridge, J.M., 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.