Industrial Policy at Work: Evidence from Romania's Income Tax Break for Workers in IT*

Isabela Manelici† and Smaranda Pantea‡

June 22, 2020

Abstract

We study the firm and sector-level effects of an industrial policy designed to support the development of the IT sector in Romania. In 2001, Romania introduced an unexpected personal income tax break to programmers with eligible bachelor's degrees and who work on software development for firms in eligible IT sector codes. In 2013, policy-makers suddenly expanded the scope of the original tax break to cover more bachelor's degrees and sector codes in IT. We first use firm-level data and difference-in-difference designs around each policy episode to show that treated firms experience strong and long-lasting growth. We then employ sector-level data and a synthetic control design to show that after the introduction of this policy in 2001, the IT sector grew faster in Romania than in otherwise similar countries. Finally, downstream sectors relying more on IT services also grew faster in Romania after 2001. Our results suggest that this policy has been effective in promoting the development of the IT sector, a sector typically seen as key to the transition to a knowledge economy.

Keywords: Industrial policy, Firm growth, Economic development, Information technology, Labor income taxation, Central and Eastern Europe, Downstream effects

JEL Codes: O25, O14, O38, L86, H24, D22, L25, O52, D57

---

*We thank Treb Allen, Alan Auerbach, Marc Bogdanowicz, Sebastian Buhai, Teresa Fort, Eva Hagsten, Douglas Irwin, Nathan Lane, Ethan Lewis, Dan Matei, Krisztina Orban, Nina Pavcnik, Andrés Rodríguez-Clare, Jose P. Vasquez, Emmanuel Saez, Benjamin Schoefer, Mine Senses, Danny Yagan, and Paolo Zacchia for very useful feedback on preliminary versions of the paper. We also appreciate the helpful suggestions received during seminar and conference presentations. We thank the National Center for Financial Information of Romania for granting us access to firm-level data, and Dan Matei and Ştefan Nanu (from the Ministry of Public Finance of Romania) for their decisive support. We thank Andrei Caprau, Mihai Codreanu, Micole De Vera, and Madeline Duhon for excellent research assistance. Last, we thank Ana-Maria Cazacu, Mihai Copaciu, and Irina Mihai (from the National Bank of Romania) for facilitating access to historical sector-level statistics. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The views expressed herein are those of the authors and do not necessarily reflect the views of either of the institutions to which the authors are affiliated.

†Department of Economics (Princeton University). Email: manelici@princeton.edu

‡Directorate of Policy, Analysis, and Research in Public Economics (Ministry of Public Finance of Romania) and the Departments of Entrepreneurship and of the World Economy (University of Economics, Prague). Email: smaranda.pantea@vse.cz

1 Introduction

"Nations have and will continue to shape their economies through industrial policy. Nevertheless, the empirical literature on these interventions is thin, dwarfed by the attention industrial policies receive from policymakers across the world" [according to a contemporaneous review of this literature by Lane, 2020]. In this article, we contribute with plausibly causal evidence on the effectiveness of an industrial policy introduced by Romania in 2001. This policy targeted the development of its information technology (IT) sector, a sector seen as key to the transition to a knowledge economy.

Since 2001, this policy provides a full personal income tax break to employees with an eligible bachelor's degree specialization who work directly on software development and generate revenues from this activity for a firm in the eligible "Software consultancy and supply" sector. In 2013, an amendment to the 2001 law greatly expanded the pool of eligible firms and workers by adding several newly eligible sector codes for the firms and bachelor's degree specializations for the workers. This intricate set of rules – on the worker, the firm, and the activity performed by the worker in the firm – implies that the tax break rewards specific matches between workers and firms. The expectation was that lowering the tax burden on these matches would lead to an increase in their prevalence. Moreover, this restrictive set of conditions ensures that the tax break actually subsidizes software development (as opposed to other misreported activities). Finally, for workers to benefit from this tax break, their employer has to prepare the necessary paperwork and apply for their income tax break. This requirement of an explicit "buy-in" from the employer motivates the expectation of a shared economic incidence of this tax incentive.

In the first part of the paper, we use firm-level data and difference-in-differences (DiD) designs to study the effects of the introduction of the tax break in 2001 and its 2013 reform. For the 2001 analysis, we use firm-level data from Amadeus. In the 2001 DiD strategy, a firm is considered treated if it belongs to the sector that became targeted by the law ("Software consultancy and supply" sector, NACE Rev 1 code 722). We compare the outcomes of firms in the eligible IT sector to the outcomes of firms in ineligible high-tech knowledge-intensive (HTKI) service sectors. For identification, we rely on the unexpected nature of the passing of the law and the lack of preexisting differential trends in outcomes between treated and comparison firms. Our estimates show that after 2001 firms in the eligible IT sector embark on a differential upward trend relative to firms in ineligible HTKI sectors. By 2005, firms in the eligible sector have, on average, 24% higher operating revenues than in 2000, 13% more workers, and 12% more assets than firms in comparison sectors. These DiD estimates measure the "intent-to-treat" effects of the policy on firms in the eligible IT sector. We use two alternative sets of comparison sectors to show that results are not an artifact of the baseline choice of comparison sectors.

Next, we examine the 2013 amendment to the income tax break law, which greatly expanded the set of eligible firms and workers. We base the 2013 analysis on more comprehensive administrative data that includes information on the number of income tax exempt employees. This allows us to estimate the effects of this reform on firms whose programmers actually benefit from the now more widely available income tax break. Therefore, we refine the definition of treatment from one based on the sector of the firm to one based on the extent of workforce exemption. We classify a firm as treated if it jumps from under 5% of workforce exemption before 2013 to over 20% of workforce exemption after 2013. The reference group contains other firms in HTKI service sectors (eligible and ineligible) that remain at 5%

of workforce exemption throughout the entire sample period.

The identification relies on the unexpected and generous expansion in firm and worker eligibility that occurred in 2013 – the most plausible driver of the sudden jump in the firm-level share of tax-exempt workers – and the lack of preexisting differential trends between treated and comparison firms. We find that firms treated by the 2013 reform experience large and long-lasting increases in size. In 2015, these firms have 20% more revenues than in 2012, 10% more employees, and 17% more assets (all relative to the comparison firms). These baseline estimates are robust to (i) using three alternative comparison group of firms in other ICT service sectors, in non-ICT HTKI service sectors, and only in the eligible IT sectors; (ii) running the baseline regression on a dataset from Amadeus that starts in 2008; (iii) varying the threshold choice from 20 to 10, 15, 25 and 30%; and (iv) defining treatment based only on the sector of the firm, as in the 2001 DiD exercise.

These firm-level results corroborate the hypothesis of a shared economic incidence of the tax break between workers and firms. Lacking worker-level data, we do not attempt to estimate how the tax incentive is split between workers and firms.[1] To interpret the magnitudes of the estimates from the two analyses, we first assume a 25% take-up between 2001 and 2005 for firms in the eligible sector. We can then translate the 13% DiD ("intent-to-treat") 2005 estimate on employment into a 52% treatment-on-the-treated estimate. As the median firm in the eligible sector has three workers in 2000, a 52% increase in employment between 2001 and 2005 is sensible. This 52% estimate is larger than the corresponding 11% estimate from the 2013 analysis. This is most likely due to the fact that the "early adopters" (i.e., firms whose workers become exempt from the income tax just after 2001) are positively selected relative to the "late adopters." That said, the very purpose of the 2013 reform was to expand the scope of the tax break to include firms and workers unable to benefit from the tax break beforehand.

In the second part of the paper, we switch to a sector-level cross-country study of the impacts of the 2001 introduction of the tax break. Using data from Eurostat and the World Bank and the synthetic control method (SCM), we evaluate the effects of this policy on the relative growth of the IT sector. This analysis is complementary to the firm-level analysis in two ways. First, it captures not only the within-firm intensive margin of growth of the IT sector but also its extensive margin (i.e., the entry of new firms into the IT sector). Second, because we benchmark the relative growth of the IT sector in Romania to that in a synthetic Romania based on comparable Central and Eastern European (CEE) countries, we control for potentially confounding sector-specific productivity or demand shocks.

The SCM estimates indicate that, in 2015, the gross revenues (employment) in the IT sector of Romania is 6.52 (1.83) times larger than the gross revenues (employment) in 2000. This value reflects the exceptional growth of the IT sector in Romania – plausibly owed to the 2001 policy – as it is relative to the growth of gross revenues (employment) in all other sectors in Romania and relative to the same difference in growth rates in synthetic Romania. Given this extra double-differencing – which controls for broader trends in the rest of the Romanian economy and similar economies – the SCM estimates of the growth of the IT sector in Romania are smaller than its actual growth between 2000 and 2015 (14-fold in gross revenues and six-fold in employment). Placebo tests suggest that a similar growth cannot be replicated in countries that did not implement this policy.

[1] Using linked employer-employee data from Sweden, Saez et al. [2019] find that a payroll tax reduction for young workers has had a full incidence on firms, who saw increases in their sales, number of workers, and capital (as in this paper).

Next, we provide evidence on the inter-sector effects of the tax break, again using SCM on Eurostat, Comtrade, and World Bank data. The improvements in the prices, quality, and variety of IT services – which are likely to have occurred alongside the expansion of the IT sector – are expected to benefit more those downstream sectors which have a stronger reliance on IT service inputs. We, therefore, ask whether sectors that relied more on IT services before the tax break expanded more than sectors with less of such a reliance (in Romania, relative to synthetic Romania). We group sectors into high- and low-intensity of use of IT services, based on the share of the IT sector in their total input expenditures (according to the input-output table of Romania for the year 2000).

These SCM results suggest that, after 2001, the high-intensity sectors of Romania grew more than its low-intensity sectors and more than in synthetic Romania (e.g., 0.75 times more in terms of gross revenues and 0.61 times more in terms of employment). Moreover, high-intensity sectors also improved their export performance more, which suggests a shift in Romania's comparative advantage. These results serve two purposes. First, they represent an additional indirect check on the effectiveness of the policy of interest, particularly due to the delayed onset of the downstream SCM effects and their smaller magnitude (both relative to the direct SCM effects). Second, while not a definitive test, the faster growth of downstream sectors relying more on IT service inputs is a necessary condition for the IT sector to have generated inter-sector externalities (one of the theoretical conditions that justify industrial policies favoring a certain sector). The SCM results survive a battery of robustness checks.

Finally, we propose back-of-the-envelope cost estimates of this policy to the government. In 2015, the foregone fiscal revenues are equivalent to 4.5 to 7.4% of the total gross wage bill in the IT sector and 2.7 and 4.5% of its total value added. We conclude that the income tax break is not only a signal of the commitment of the government to the development of the IT sector but is also a sizable incentive.

Most directly, this article contributes to a long-standing academic debate on industrial policy. Academics are typically skeptical about the effectiveness of industrial policies due to possible capture by "sunset sectors" or lobbying firms, and difficulties to design and implement successful industrial policies (particularly so in countries with weak institutions). Despite this skepticism, the set of papers that provide well-identified reduced-form evidence on the effectiveness of industrial policy is rather small and recent [see the review of Lane, 2020]. Görg et al. [2008], Aghion et al. [2015], Juhász [2018], Criscuolo et al. [2018], Cai and Harrison [2019] are among these recent papers providing credible evidence, through only a minority pertain to emerging economies [a notable shortcoming of this literature Rodrik, 2008]. We bring to this debate evidence on the effectiveness of a Romanian industrial policy with a unique design[2] and targeting a sector widely regarded as systemically important.

The natural follow-up question is whether an effective industrial policy is also efficient. While answering this question lies outside the scope of this paper, this tax break appears to meet the theoretical criteria for welfare-improving industrial policy put forward by the literature. First, it encourages software development, a "new" activity for the domestic economy in 2001 [Rodrik, 2004], and one that is knowledge-intensive [Aghion et al., 2011, European Commission, 2017, Cherif and Hasanov, 2019]. Second, given its skill endowment in STEM inherited from communism, Romania most likely had a latent comparative advantage in this activity and only lacked a policy signal to tilt resources towards it

[2] The most commonly studied industrial policies involve tax breaks or credits for capital and R&D [Fowkes et al., 2015, Boeing, 2016, Cai and Harrison, 2019], grants [Görg et al., 2008, Criscuolo et al., 2018], or trade tariffs [Aghion et al., 2015]

[Rodrik, 1996, Harrison and Rodríguez-Clare, 2010, Itskhoki and Moll, 2018]. Third, we find that the growth of the IT sector has supported the growth of IT-using sectors – a necessary condition for the sector to generate inter-sector externalities.[3] Last, this policy benefits a sector that is competitive, and in that sector, all firms and workers meeting the eligibility criteria [Aghion et al., 2015].

By studying a policy that targets the IT sector, we naturally relate to research on this specific sector. One strand of this literature establishes the wide-ranging effects of IT, of which the effects on productivity have garnered the most attention.[4] In particular, Van Ark et al. [2008] makes the case that the later emergence and smaller size of the IT sector in the European Union (compared to the United States) explains its slower productivity growth. This makes the policy we study especially relevant to countries that grapple with the drawbacks of an underdeveloped IT sector. Another strand of this literature studies the drivers of growth of the IT sector. To our knowledge, this is the first paper to provide plausibly causal estimates of the effectiveness of tax incentives to boost firm size in the IT sector.

Finally, given the specifics of the policy, we also relate to research on policies aimed at reducing non-wage labor costs. Most of this research evaluates the effects of reductions in non-wage labor costs for hard-to-employ workers (such as unemployed individuals, youth, parents returning to work, or people with disabilities) or wage subsidies that aim to support job-creation in general. The typical finding is one of positive effects on firm size.[5] For instance, Kangasharju [2007] studies the effects of wage subsidies for hard-to-place job seekers (mostly long-term unemployed) in Finland and finds a pooled DiD estimate of a 9% increase in employment. Our larger estimates for IT firms (e.g, of 11% in the 2013 analysis) are most likely explained by the higher value of the incentive for the IT sector; the labor costs associated with programmers are likely to be more significant to IT firms than those associated with hard-to-employ workers (typically hired in low-skilled support positions, such as cleaners, by firms in all sectors).

By studying a personal income tax break for a high-skill/high-wage occupation, we also relate to a smaller set of papers that study reductions in taxes on the wages of R&D workers. As most papers study the effects over a short-term horizon – during which the supply of researchers is most likely inelastic – their most common finding is an increase in researcher wages [Goolsbee, 1998, Lokshin and Mohnen, 2013].[6] Our positive effects on employment most likely stem from a more elastic supply of programmers (relative to researchers), the fact that this policy reduced the incentives of Romanian programmers to emigrate (thus increasing the number programmers available to firms in the country), and the fact that the 2013 amendment acted like a shock to the supply of programmers eligible for the income tax break.

The remainder of the article is organized as follows. Section 2 describes the two policy episodes of

[3] An important theoretical motive to deviate from policy neutrality requires that the targeted sector would later generate externalities [Succar, 1987, Greenwald and Stiglitz, 2006, Harrison and Rodríguez-Clare, 2010]. One such externality occurs between sectors, through the supply of specialized inputs used by (many or high-technology) downstream sectors. By studying not only the targeted sector, but also those affected through I-O linkages, we relate to both seminal [Hirschman, 1958, Pack and Westphal, 1986] and recent work on the I-O implications of industrial policy [Forslid and Midelfart, 2005, Du et al., 2014, Blonigen, 2016, Huremović and Vega-Redondo, 2016, Lane, 2017, Liu, 2019, Joya and Rougier, 2019].

[4] IT – or ICT, more broadly – has been found to improve productivity in IT-using firms [Jorgenson et al., 2008, Syverson, 2011, Bloom et al., 2012], increase wages in high-skill locations [Forman et al., 2012], reduce information frictions [Steinwender, 2018], foster service exports [Kneller and Timmis, 2016], and improve educational attainment [Beaudry et al., 2010] etc.

[5] For examples on positive effects on employment, see Kangasharju [2007], Girma et al. [2008], Lombardi et al. [2018], Saez et al. [2019], Banai et al. [2020]; on positive effects on profits, revenues, and investment, see Månsson and Quoreshi [2015], Saez et al. [2019], Banai et al. [2020]. For a review, see Eurofound [2017].

[6] For a review, see European Commission [2014]. In addition to wage effects, a subset of these studies also find positive effects on the employment of researchers [Guceri, 2018] or other high-skilled workers [Kaiser and Kuhn, 2016].

interest: the 2001 introduction of the income tax break to workers in IT and its 2013 reform. Section 3 presents our firm-level empirical strategy and findings. In Section 4, we bring sector-level cross-country evidence on both the direct and downstream effects of the 2001 introduction of the tax break. Section 5 provides back-of-the-envelope cost estimates of this tax break. Section 6 concludes.

2 Romania's Income Tax Break for Workers in IT

In 2001, high labor taxes were seen as a major constraint to the development of the IT sector in Romania. At the time, personal income taxes were progressive, with rates between 18 and 40%. Payroll taxes (social security contributions and insurance) were paid by both employers (up to 40%) and employees (17%) on gross salary. While the tax rates were the same across sectors, the progressive taxation combined with deductions based on household characteristics and gross salaries, led to varying ratios between net and gross salaries. In particular, the IT sector had a lower ratio (0.68) than that of the entire economy (0.75) or comparable sectors (over 0.70), or put differently, a higher burden of taxation. In addition, programmers are relatively better positioned to have their qualifications recognized abroad, due to lower language barriers and more standardized ways to appraise IT skills. Together, these reasons were seen as important drivers of the high emigration rates of programmers. High labor taxes also led to relatively high labor costs for firms, limiting their growth.

Since 2001, Romania's IT sector has experienced dramatic shifts. The sector has greatly expanded, both in absolute terms and as a share of GDP. It has also become more integrated into the global economy through flows of foreign direct investment (FDI) into the sector and the growing importance of foreign revenues. IT-related bachelor's degrees remain among the most popular degrees to this day. A tax break introduced in 2001 and expanded in 2013, is widely perceived as having triggered these shifts.

In 2001, Romania introduced a personal income tax break for programmers at the initiative of Mr. Varujan Pambuccian, a member of the Chamber of Deputies. He expected that this tax break would address the concerns of both workers and firms in the IT sector. Mr. Pambuccian was, and still is, the representative of the Armenian minority in Romania in the Chamber of Deputies. For this reason, his proposal was not part of the program of a major party and widely discussed during the electoral campaign. It also did not benefit from automatic support in the Chamber of Deputies. Eventually, however, Mr. Pambuccian managed to convince the Prime Minister and the government about the potential benefits of this exemption. His first proposal had been to reduce the top marginal income tax rate on the salaries of programmers from 40 to 8%. The Ministry of Finance assessed this proposal difficult to implement. Instead, the final version entailed a full tax break from the income tax for programmers. The tax exemption was adopted by Government Emergency Ordinance 94/2001 and only later approved as law by the Parliament. Given the implementation difficulties and criticism related to the first version, in addition to the initial adoption by Government Emergency Ordinance, it is plausible that the introduction of the tax exemption was an unexpected and positive shock for the IT sector.

To benefit from the tax break, workers had to simultaneously (i) have an eligible bachelor's degree (in either automation, computers, informatics, cybernetics, mathematics, or electronics); (ii) work for a firm in the "Software consultancy and supply" sector (NACE Rev 1 code 7220); (iii) work for the unit in charge of software development; (iv) have an eligible occupation title (e.g., "programmer" or "computer systems designer"); and finally (v) work for a firm that kept separate balance sheets recording revenues from software development and generated gross revenues of at least 10,000 U.S. dollars from this activity in the previous fiscal year (per exempted employee). An important feature of the policy is that, while the exemption applies to the income tax owed by workers, the firm is responsible for preparing the justifying documents, applying for the tax break, and archiving the documents for potential audits.

Four features of this tax break deserve emphasis. First, this break was particularly generous at its introduction. In 2001 the wages of programmers were among the highest in the country. Before the full income tax break, programmers faced a progressive personal income tax with rates between 18 and 40%. In 2005, all workers in the country saw a switch from progressive taxation to a flat rate of 16%. Between 2004 and 2005, the tax wedge for workers earning the average economy-wide salary decreased by two percentage points and for high wage earners by four.[7] Despite this change, the full exemption from the income tax of 16% and the trend of growing salaries in the IT sector meant that this break remained a sizable incentive for the IT sector. Second, the rules to benefit from the tax break were meant to ensure that the economic incidence of the tax break was shared between workers and firms. The policy rewarded specific types of matches between workers and firms engaged in software development activities. We expect that lowering the tax burden on these activities led to an increase in their prevalence. We document the equilibrium effects of this policy using firm- and sector-level data. Third, its strict accountability rules (explained above) ensured that exempted workers were actually developing software. Hence, the effects we estimate are plausibly real responses to the incentive and not a mere relabeling of activities.[8] Finally, this policy was designed to benefit all eligible workers and firms in the IT sector. Industrial policy works better when benefits are less concentrated [Aghion et al., 2011].

The first major amendment to the tax break law occurred in 2013. It was initially prompted by the need for Romania to transition from the NACE Rev 1.1 classification of sector codes to the NACE Rev 2 classification.[9] Similar to the initial introduction of the policy in 2001, the passing of this amendment and its final eligibility criteria were made uncertain by the fact that the negotiations spanned three different governments in a period of high political instability. The final form was first approved as Order 539/225/1479/2013, to be only later adopted as law. The amendment greatly expanded the scope of the income tax break by both increasing the number of eligible sector codes (for the firms)[10] and the number of eligible majors for the bachelor's degree (for the workers).[11]

[7]The first statistic characterizes single workers without children who earned 100% of the economy-wide average salary and whose tax wedge decreased from 28.1% to 26.1%, whereas the second characterizes single workers without children who earned 167% of the average salary and whose tax wedge decreased from 32.5% to 28.5%. Source: INSSE. The tax wedge is defined as the ratio between the amount of taxes paid and the corresponding total labor cost for the employer.

[8]Chen et al. [2018] study a Chinese policy that awards substantial corporate tax cuts to firms that increase R&D investment and find that 30% of the increase in R&D comes from the relabeling of administrative expenses.

[9]In 2004, the law was revised to reflect the transition from the NACE Rev 1 to the NACE Rev 1.1 classification. This revision had no economic effects because the newly-beneficiary NACE Rev 1.1 sectors 7221 and 7222 were the same as the formerly-beneficiary NACE Rev 1 sector 7220.

[10]In 2013, Romania had to transition from the NACE Rev 1 classification of sector codes to the NACE Rev 2 classification. Because the crosswalks between classifications are not bijective, some sector codes became newly eligible during the transition between classifications. See Table A2 in Appendix A.1 for details.

[11]While two additional amendments were introduced in the second half of 2015 and 2016 respectively, we do not study them due to data availability constraints and their more limited scope. See Appendix A.1 for details. Appendix A.2 summarizes other policies relevant to the Romanian IT sector and our approach to isolating the effects of the income tax break.

3 Firm-Level Analysis

3.1 2001 Income Tax Break

We begin by studying the impact of the 2001 income tax break on firms already active in 2001. The main advantage of this analysis is that it studies the IT sector over a period when, still in its infancy, it received the unexpected positive news of the introduction of a dedicated tax break to programmers.

Data. For the 2001 analysis, we rely on Amadeus to construct a panel of Romanian firms with three outcome variables: log operating revenues, log number of workers and log total assets. The panel starts in 1997 and ends in 2005. These variables are both likely to react to the new incentive and are among the few variables whose values are less frequently missing. To remove outliers, the sample is winsorized at the 1^{st} and 99^{th} percentiles of the distribution of each of the outcome variables and the change in operating revenues per worker. Moreover, we only keep firms which are active for at least 2000, 2001 and 2002. Last, in the baseline sample we only keep firm-year observations with non-missing values for all three main outcome variables: log revenue, log number of workers, and log total assets.[12]

Empirical Strategy. We estimate the firm-level effects of the introduction of the tax break on firms via a difference-in-differences (DiD) design. The first difference is taken between firm outcomes in a given year between 1997 and 2005 and the same firm outcomes in the year 2000 (the reference year). The second difference is taken between the contemporaneous outcomes of firms in the sector with NACE Rev 1 code 722 (the eligible sector) and the outcomes of firms in comparable sectors. Formally,

$$y_{ist} = \alpha_i + \sum_{k=1997}^{2005} \delta_k \mathbb{1}[t=k] + \sum_{k=1997}^{2005} \beta_{DiD,k} \mathbb{1}[t=k] Eligible\_sector_{is} + \varepsilon_{ist}, \tag{1}$$

where i stands for firm, s for the sector of firm i, t for the calendar year. α_i is the firm fixed effect. $\mathbb{1}[t=k]$ is a dummy taking value 1 if an observation pertains to calendar year k and is meant to capture common shocks to all firms that year. We use as outcome variables, y_{ist}, the firm i, year t, log operating revenue, log number of workers and log total assets. We cluster standard errors at the firm level.

The treated firms are those in the eligible IT sector (NACE Rev 1 sector 722, Software consultancy and supply); for those firms, $Eligible\_sector_{is} = 1$. The comparison group for our baseline results ($Eligible\_sector_{is} = 0$) includes firms in (ineligible) high-tech knowledge intensive (HTKI) sectors (as classified by Eurostat).[13] These comparison sectors resemble the eligible IT sector in their focus on

[12] Of the 27 European countries in Amadeus, Romania is in the top tier of countries with the highest coverage [Kalemli-Özcan et al., 2015]. For Romania, the data provider to Amadeus is the Chamber of Commerce and Industry, which reports on an annual basis on the account of more than 500,000 companies (joint stock companies, partnerships limited by shares, limited liability companies, state owned companies, co-operative companies). The variables used in our analysis (operating revenue, employment and total assets) are those with the most consistent coverage in Amadeus [Kalemli-Özcan et al., 2015].

[13] Specifically, the baseline comparison group contains firms in the following sectors: 2214 (Publishing of sound recordings), 721 (Hardware consultancy), 723 (Data processing), 724 (Database activities), 725 (Maintenance and repair of office, accounting, and computing machinery), 726 (Other computer related activities), 3002 (Manufacture of computers and other information processing equipment), 731 (Research and experimental development on natural sciences and engineering), 732 (Research and experimental development on social sciences and humanities), 7487 (Miscellaneous business activities n.e.c.), 921 (Motion picture and video activities), 922 (Radio and television activities), and 924 (News agency activities). We exclude sector 642 (Telecommunications) as it underwent a massive liberalization in 2002 in Romania.

high value-added services and reliance on high-skilled workers and technology. Moreover, Table B1 (Appendix B.1.1) shows that both the median and average size of firms in treated and comparison sectors are similar. For instance, both the median treated and comparison firms had revenues of 27,000 euros in 2000 and employed three workers.

To deliver credible estimates of the treatment effect of this policy, this strategy relies on two aspects of the research design. First, it relies on its unexpected introduction. As discussed in Section 2, this policy was introduced by Emergency Ordinance at the initiative of one independent policy-maker alone. The lack of wide debate over the policy and the decision to grant a full personal income tax break to programmers (as opposed to the initially-proposed partial break) turned the policy into an unanticipated shock to the IT sector. Second, identification hinges on the assumption that firms in the comparison group form a suitable counterfactual for firms in the IT sector, after accounting for time-invariant differences between firms and common year-specific shocks. The lack of differential pre-trends between treated and comparison firms is an important test for the validity of both assumptions.

As Amadeus does not include information on the extent to which the workers of a given firm have actually benefited from the income tax break, these DiD estimates do not measure the impact of the introduction of the exemption on firms that start having exempted workers, but instead measure the average effects on firms that are part of a sector with a newly available exemption for their workers involved in software development (i.e., "intent-to-treat" effects).

Baseline Results. Figure 1 plots the DiD estimates from the model in Equation (1). These estimates pertain to our baseline comparison group (i.e., firms in HTKI sectors). Reassuringly, across all outcome variables, we observe a lack of preexisting differential trends between treated and control firms, and between 1997 and 2000. After 2001, however, firms in the eligible sector experience significant improvements in all three measures of firm size, such that by 2005, firms in the IT sector have a 24% higher operating revenue, employ 13% more workers, and have 12% more assets than firms in comparable sectors (relative to 2000). Table 1 provides more details. The estimates of the pooled DiD coefficients (measuring the average increase in an outcome from the 1997-2000 period to the 2001-2005 period) are 31% for operating revenue, 10% for the number of workers, and 23% for total assets.

Robustness Checks. The strong and lasting effect of the policy on firm size survives different comparison groups. We first use an alternative comparison group which contains firms in the subset of HTKI sectors that are the closest to the treated IT sector 722, namely R&D sectors (as focused on knowledge creation as sector 722) and IT sectors that became newly eligible for the income tax break after the 2013 reform (due to their perceived similarity to the already tax-exempt sector 722).[14] The second alternative comparison group is comprised of HTKI sectors that *are not* ICT service sectors.[15] This second group allows us to rule out concerns of spillovers from the eligible software development sector to other ineligible yet related sectors (such as the hardware consultancy sector).

Columns (4)-(6) and (7)-(9) in Table 1 report the DiD estimates for the two alternative comparison

[14] Sectors 721 (Hardware consultancy), 724 (Database activities), 726 (Other computer related activities), and 3002 (Manufacture of computers and other information processing equipment) became eligible for the income tax break in 2013. In addition to these sectors, the first comparison group also contains sectors 731 (Research and experimental development on natural sciences and engineering) and 732 (Research and experimental development on social sciences and humanities).

[15] The second alternative comparison group contains NACE Rev 1.1 sector codes 2214, 731, 732, 7487, 921, 922, and 924.

groups. Contrasting these estimates with those using the baseline comparison group suggests that our results are driven by the policy itself and not by the choice of the comparison group. First, we note that the sign and significance of the DiD coefficients remains unchanged. Second, given that the 95% confidence intervals overlap between specifications for the same variable, we do not find compelling evidence of contamination between the treated IT sector and similar sectors.

3.2 2013 Reform to the Income Tax Break Law

We now move on to study the impact of the 2013 reform to the conditions of eligibility for the income tax break for workers in IT. The main advantage of this analysis is that it is performed on administrative data recording the actual firm-level exemption rate from the income tax of its workers. This allows us to estimate the effect of the actual exemption as opposed to the "intent-to-treat" effect on firms in the eligible sector, irrespective of the actual exemption status of their workers.

Data. The firm-level analysis of the impact of the 2013 reform is based on administrative datasets collected by the National Agency of Fiscal Administration of Romania. The first dataset contains firms' balance sheets, which give us information on yearly revenue and total assets. We add firm-level information coming from two compulsory fiscal forms that record the income taxes paid by workers. In particular, these forms track the firm-level number of workers exempted from paying any personal income tax. The resulting dataset starts in 2011 and ends in 2015.[16]

We first remove from the sample firms that benefited from major State Aid programs (797/2012 and 332/2014) during the period studied. Second, we limit the sample to firm-observations with a non-missing number of income tax exempted workers. Third, we only keep firms which are active at least in 2012, 2013 and 2014. Last, in the baseline sample, we only keep firm-year observations with non-missing values for all three outcome variables: log revenue, log number of workers and log total assets.

Empirical Strategy. In order to estimate the firm-level effect of the 2013 extension of the tax break to new firm activities and bachelor's degree majors, we estimate the following DiD specification:

$$y_{ist} = \alpha_i + \lambda_{st} + \sum_{k=2011}^{2015} \delta_k \mathbb{1}[t = k] + \sum_{k=2011}^{2015} \beta_{DiD,k} \mathbb{1}[t = k] Exempted_{isk} + \varepsilon_{ist}, \tag{2}$$

where, with the exception of λ_{st}, $Exempted_{isk}$ and the new analysis period, the specification in Equation (2) is identical to the specification in Equation (1). λ_{st} is the sector-by-year fixed effect.

$Exempted_{isk}$ is a dummy equal to 1 whenever in year k firm i in sector s has more than 20% of its workforce exempted from the income tax. We keep in the baseline sample only firms who in 2011 and 2012, had less than 5% of workers exempted from the income tax. To improve the interpretation of the estimates, we exclude firms who never reach the 20% threshold after 2013, while at the same time surpassing the 5% threshold at least once between 2011 and 2015. Hence, all firms in the baseline

[16] Firms fill in and submit the D112 and D205 forms, retain the owed income taxes from the wages of their workers, and transfer these taxes to the tax authority (all on behalf of their workers). The 2013 analysis relies heavily on these forms. As these forms were first introduced in 2011, this analysis can only start that year. The original dataset includes 2016 as well, but we are excluding 2016 from the analysis because the outcomes in that year are likely to be affected by the 2015 amendment to the income tax break law (see Section 2 and Appendix A.1 for details).

sample have $Exempted_{isk} = 0$ for $k < 2013$ and only treated firms have $Exempted_{isk} = 1$ for $k \geq 2013$.

We choose 5% – instead of 0% – because there are other categories of workers who are eligible for this tax exemption (in particular those with disabilities).[17] We keep only firms with at most 5% workforce exemption before 2013 to mimic the tax conditions of IT firms before the initial introduction of the policy in 2001. The 20% threshold after 2013 is chosen to avoid inadvertently measuring the effect of other income tax exemptions that are unrelated to the 2013 reform, and to ensure that the exemption from the income tax has a non-trivial effect on firm labor costs. This choice is meant to capture both firms whose sector code became suddenly eligible in 2013 and firms (with an eligible NACE Rev 1.1 sector code before 2013) for whom a significantly larger share of workers became eligible in 2013.[18] 20% is also a typical percentage of exemption in the eligible sector.

In addition to these sample restrictions, in the baseline exercise, we only keep firms in high-tech knowledge-intensive (HTKI) service sectors (as classified by Eurostat).[19] Appendix B.1.2 contains descriptive statistics on our baseline sample. Table B2 compares the firm size and demographic characteristics in 2011 for the three groups of firms: firms in ineligible sectors, firms in eligible sectors with less than 5% of employees exempted from the income tax throughout the entire sample period, and firms with less than 5% of exempted employees in 2011 and 2012, and which jumped to over 20% of exempted employees after 2013. Firms in ineligible sectors are, on average, the largest; firms in eligible sectors for which a large share of workers became exempted after 2013 have the second-highest average size; while firms in eligible sectors that remain under 5% exemption rate have the smallest average size. In all our specifications, we control for time-invariant differences in size using firm fixed effects.

Table B3 shows that the percentage of firms in eligible sectors with at least one worker exempted from the income tax has increased from 36% in 2011 to 45% in 2015. If we require firms to have more than 20% of their workers exempted from the income tax, the share of such firms increases from 28% in 2011 to 36% in 2015.[20] Table B4 documents the predictors of the firm-level share of workers who are exempt from the income tax. Firms that are foreign-owned, larger, and/or located in counties with a higher overall share of workforce exemption are more likely to have a higher share of workforce exemption from the income tax. These correlations may reflect the ability of these firms to meet the restrictive conditions of the exemption (for instance, by having access to a larger local pool of eligible programmers) or to assemble the documentation necessary to solicit the tax break.

Table 2 studies the "first-stage" effects of the 2013 reform on the firm-level share of workforce exemption from the income tax. We find that in 2013 and after, firms in eligible sectors (according to the new definition of eligibility post-2013) experience a marked increase in their share of workforce exemption relative to 2012. In 2015, firms in eligible sectors have a share of exempted workers that

---

[17] See Appendix A.2 for details on these other categories and the reasons why they do not jeopardize our empirical strategy.

[18] We cannot separate these two scenarios for two reasons. First, we do not have worker-level information. Second, we only observe firms' sector codes in 2016, already in the NACE Rev 2 classification. The NACE Rev 2 codes that became eligible in 2013 contain NACE Rev 1.1 codes that were both eligible and ineligible before 2013 (see Table A2, Appendix A.1).

[19] Specifically, the following NACE Rev 2 sectors are classified as HTKI service sectors: 582, 59, 60, 61, 62, 63, and 72.

[20] This relatively low level of take-up of the tax exemption is likely due to: (i) the restrictive set of conditions that must be jointly met by firms and programmers in order for programmers to qualify for the tax break; (ii) difficulties in hiring eligible programmers in a tight labor market; (iii) a lack of knowledge of the administrative procedures to apply for the break; or (iv) the high perceived cost of preparing the required documentation. While we cannot distinguish between these scenarios, the first two are the most plausible. The restrictive nature of the criteria to qualify for the break and the need to implement the NACE Rev 2 classification are the main motivations of the 2013 amendment.

is 4% higher than in 2012. Also, in 2015 there are 7% more firms in eligible sectors whose share of workforce exemption is larger than 20% (again relative to 2012). We find no evidence of trends in these measures of workforce exemption between 2011 and 2012. These findings suggest that the 2013 reform was effective in its goal to broaden the access of firms and workers to the tax break.

Before proceeding to the results, we want to highlight two key differences between the DiD specification in Equation (2) and the one in Equation (1). First, while in the previous specification firms were deemed treated after 2001 when they were part of the eligible sector, firms are now deemed treated if they are part of the eligible sectors *and* if they start with less than 5% of workers exempted from the income tax pre-2013 and suddenly exceed the 20% threshold of workforce exemption after 2013. Second, we now provide estimates that characterize a subset of the firms in the eligible IT sectors, containing firms that were most likely ineligible for the tax exemption before the 2013 amendment (either due to their sector code or due to their workers' bachelor's degrees).

This new definition of treatment has several advantages. While in the Amadeus data used for the 2001 exercise we did not observe how many of a firm's employees actually benefited from the tax break (if any), the administrative data we use in this exercise tracks this number and allows us to focus on firms whose workforce became treated to a sizable degree. In addition, defining treatment as firm-specific allows us to control for sector-by-year FEs, in addition to firm FEs. This new set of FEs control for potential sector-specific demand and/or technology shocks contemporaneous to the 2013 reform.

This new definition also raises concerns over the extent to which a jump after 2013 in the share of tax exempted employees is endogenous to firm characteristics. Lacking worker-level data and data on the NACE Rev 1.1 sector code of firms pre-2013, we cannot unequivocally address these concerns. Notwithstanding, we rely on the unexpected timing and generous nature of the 2013 reform (see Section 2). We also build on general information about the state of the IT sector around 2013. At that time, the sector faced a notably scarce labor supply relative to demand, which forced IT firms to hire programmers who were ineligible for the tax break. It is therefore very plausible that if a firm suddenly jumped from under 5 to over 20% of workforce exemptions after 2013, this jump was caused by the expansion of the lists of eligible sector codes and bachelor's degree majors. Conversely, firms who stayed under the 5% threshold throughout 2001 to 2015 either had a sector code that did not become eligible for the tax break or employed programmers whose bachelor's degree majors remained ineligible after 2013.

Moreover, all our specifications include firm fixed effects to control for time-invariant firm-specific unobservables. What we cannot rule out directly are firm-specific shocks contemporaneous with the 2013 policy shock (such as a change in management), which may be the actual driver of both the jump in the firm-level workforce exemption share and the estimated firm growth effects. That said, the jumps we study occurred after 2013, implying that they were most likely driven by the policy shock and not by large-scale changes in management. Last, the lack of preexisting differential trends with respect to firms in the comparison group also suggests that firms that experienced a sudden increase in the share of exempted workers were unlikely to be undergoing notable productive or organizational changes.

Last, we will show that our baseline estimates are robust to (i) using three alternative comparison group of firms in other ICT service sectors, in non-ICT HTKI service sectors, and only in the eligible IT sectors; (ii) running the baseline regression on a dataset from Amadeus that starts in 2008 and allows us to observe a longer pre-reform period; (iii) varying the threshold choice from 20 to 10, 15, 25 and 30%; and (iv) defining treatment based only on the sector of the firm.

Baseline Results. Figure 2 plots the DiD estimates from the model in Equation (2). We find that across all outcome variables the estimate of $\beta_{DiD,2011}$ is not statistically different from zero. This lack of anticipatory effects is in line with the unexpected nature of the expansion. Moreover, firms whose workforce became significantly exempted from the income tax after 2013 were not embarked on differential growth trends relative to firms whose exemption rate was left unaffected by the reform. After 2013 however, the treated firms experience a gradual growth along all three measures of firm size, such that in 2015, they have a 20% higher revenues, 10% more workers, and 17% more assets relative to 2012. Columns (1)-(3) in Table 3 provide more details on these estimations. The estimates of the pooled DiD coefficients are 22% for revenues, 11% for the number of workers, and 13% for total assets.

Heterogeneity Analysis. The DiD estimates so far refer to the average effect on treated firms. However, an important policy concern is that the tax break expansion benefited only specific groups of firms, for instance large foreign firms. We estimate the baseline specification on four groups of firms, defined based on their size and age in 2011: (i) employing strictly less than ten workers (micro firms); (ii) employing at least ten workers (small, medium, or large firms); (iii) strictly less than five years old (young firms); (iv) at least five years old (old firms). Because we do not observe foreign ownership in 2011, we cannot directly alleviate the concern that foreign firms benefited more from this policy. That said, most foreign-owned firms in the sector are likely to be large and most micro firms are likely to be domestically-owned.

Table 5 reports the results of the heterogeneity analysis. We find that the tax exemption had a positive effect on all types of firms. The point estimate for revenues is smaller for micro firms than for larger firms (though these estimates are not statistically different). The point estimates for workers and assets are very similar across size categories (though the smaller sample for larger firms leads to noisier estimates). The tax exemption has had stronger effects on younger firms than on older firms; this is likely to reflect the fact that the older firms treated by the 2013 reform may be negatively selected. Most older firms already had more than 5% exemption rate in 2011 and, hence, are not part of the analysis.

Robustness Checks. All robustness checks for our 2013 baseline results are in Tables 3 and 4, and in Appendix B.2. In our first check, we show that our baseline estimates are robust to three alternative control groups. In columns (7)-(9) of Table 3, the alternative control group 1 contains firms in ICT service sectors (as classified by the OECD).[21] In columns (1)-(3) of Table 4, the alternative control group 2 contains firms in HTKI service sectors *excluding* ICT service sectors. In columns (7)-(9) of Table 4, the alternative control group 3 only contains firms in eligible sectors. The similarity of the results across control groups suggests that the results are driven by the tax exemption and not by unrelated trends in the control groups. Note that results are not sensitive to whether we include firms in ICT service sectors, alleviating potential concerns over contamination from the treated firms to firms in related sectors.

Second, we repeat our 2013 analysis on a panel from Amadeus that starts in 2008. We first merge the raw Amadeus data with each of the four samples used in the regressions on the administrative data

[21] This control group includes the following NACE Rev 2 codes: 582 (Software publishing), 61 (Telecommunications), 62 (Computer programming, consultancy and related activities), 631 (Data processing, hosting and related activities, and web portals), and 951 (Repair of computers and communication equipment).

(i.e., the baseline sample and the three samples with alternative control groups). The merge with the administrative data allows us to classify firms in Amadeus into treatment groups. We then clean the Amadeus data using the same rules applied to the 2001 Amadeus sample.

Tables 3 and 4 present the results from the Amadeus data alongside those from the administrative data. The main advantage of the Amadeus results is that they allow us to observe firms for up to five years before the reform (relative to two years). Across alternative control groups, treated and comparison firms continue to exhibit parallel trends before 2013. In terms of the magnitude of the effects, the 95% confidence intervals overlap between specifications for the same variable, implying similar conclusions.

Estimates of the DiD coefficients are also robust to the choice of the share of exempted workers above which we consider a firm to become treated. Table B5 shows the results of estimations where treatment arises when firms jump after 2013 to at least 10, 15, 25, or 30% of workers exempted from income tax. For each of these thresholds, we report the estimates for both the yearly and pooled DiD coefficients. We present the results for both the administrative (upper panel of Table B5) and Amadeus data (lower panel of Table B5). The point estimates for the pooled DiD coefficients have overlapping confidence intervals, which cautions against over-interpreting the pattern of the point estimates. That said, the overall impression is that the higher the threshold, the lower the estimate of the effects. One plausible explanation is that the firms which experience a higher jump in the rate of workforce exemption tend to be smaller. As we have learned from the heterogeneity analysis, smaller firms experience slightly lower effects. The 20% threshold appears as a conservative choice (compared to the lower thresholds).

In a last robustness check, we revert to the definition of treatment based on a firm's sector, that we use to study the initial introduction of the policy in 2001 (see Equation (1)). This definition mitigates concerns over the potentially-endogenous firm-level jump in the share of exempted employees. Table B6 reports the results from this exercise implemented on two samples: the baseline sample from columns (1)-(3) of Table 3 and an enlarged sample. The two main takeaways established so far continue to hold, particularly for the revenues and employment variables. Namely, (i) firms in the eligible sectors continue to not be engaged in differential trends before 2013 (with respect to firms in ineligible HTKI service sectors), and (ii) after 2013, firms in eligible sectors grow relatively more. The fact that the post-2013 effects are smaller than our baseline effects in Table 3 is intuitive, given that in 2015, the average firm in eligible sectors has a share of exempted workers that is 4% higher than that in 2012 (Table 2, column (8)). This 4% jump is to be compared with the jump from under 5% pre-2013 to over 20% post-2013 that characterizes the treated firms in our baseline specification. We prefer the baseline results delivered by the specification in Equation (2) for two reasons: (i) they focus on the firms actually treated by the 2013 reform, and (ii) they control for sector-by-year shocks unrelated to this reform.

3.3 Discussion of the Firm-Level Findings

Overall, we find that both policy episodes have led to strong and lasting growth for firms in eligible IT sectors. A natural question is whether this finding is consistent with the tax break's statutory incidence on workers' personal income. We provide three arguments in favor of an affirmative answer.

First, the restrictive requirements of the tax break ensure that it only applies to particular worker-firm matches, which is likely to grant bargaining power to firms over a shared economic incidence of the tax break. Second, the tax break law stipulates that it is the firm that is responsible for preparing the necessary paperwork and applying for its workers' tax break. This requirement of an explicit "buy-in" from the firm is another reason to expect a shared economic incidence of this tax incentive. Third and last, both policy episodes have led to increases in labor productivity (measured as revenue per worker).[22] There are three plausible drivers of these improvements in labor productivity. First, this policy allows firms to pay workers higher net wages and, by doing so, to improve their motivation and efficiency. Second, the increase in employment in IT is likely to have led to sector-level economies of scale.[23] Finally, it is also likely that both policy episodes have attracted higher-ability workers to software development.

While we find evidence of a shared economic incidence of the tax break, estimating the split of the tax between firms and workers is outside the scope of this paper. As we lack worker-level data, any estimate of this split would need to be model-dependent. Moreover, the split is likely to change with time, depending on the entry of new firms into the IT sector and the long-run elasticity of the supply of programmers. Nonetheless, the firm-level effects are in line with the incentive structure of this policy.

Before comparing the magnitude of the estimates from the two firm-level exercises, let us first highlight the distinguishing features of each. When this policy was introduced in 2001, the Romanian IT sector was in its infancy. Before 2001, Romanian programmers were emigrating at high rates, lacking confidence that the Romanian IT sector was poised for growth. The introduction of the 2001 tax break shifted that perception, as it signaled that the development of the sector had become a priority for policy-makers. Whereas firms and workers who became eligible in 2001 for the tax break are still eligible, in the 2001 exercise, we only focus on the 1997 to 2005 period. This aims to isolate the effects of this tax break from those of other policies that may interact with it (such as the nationwide switch in 2005 from progressive income taxation to a flat income tax) or global shocks to the IT sector.

Moreover, the Amadeus data used in the 2001 exercise does not contain the firm-level exemption rate of workers but only the firm's sector. Hence, this exercise estimates the effects on firms whose sector becomes eligible, i.e., "intent-to-treat" effects. The typical firm in the eligible sector employed 13% more workers in 2005 relative to 2000 and relative to the typical firm in ineligible HTKI sectors. Assuming a 25% take-up among firms in the eligible sector, we convert the ITT estimates to "treatment-on-the-treated" (TOT) estimates of 52%. Given that in 2000, the median firm in the eligible sector employed three workers, a 52% increase in employment is sensible. Furthermore, the firms who took advantage of this new tax break during those initial years were likely to be positively selected from all firms in the IT sector, lending further credibility to these magnitudes.

The 2013 exercise estimates the effects on firms who actually experience a sizable decrease in labor costs. Despite this merit, this exercise also has the disadvantage of focusing on a sample of firms that in 2012 – 11 years after the initial introduction of the policy – have under 5% exempted employees. Hence, this sample of firms is likely to be negatively selected (e.g., be younger and less experienced, or smaller). This is likely to explain the smaller 10% estimate for the increase in workforce after the 2013 reform, relative to the 52% TOT estimate for 2001. Nevertheless, the purpose of the 2013 reform was precisely to improve the reach of the tax break to relevant but not yet eligible firms and workers.

[22] In 2015, the labor productivity of Romania's IT and information services (value added per hour worked, PPP) was the third-highest in the EU28, only behind that of Ireland and Luxembourg [van der Marel et al., 2020].

[23] Bartelme et al. [2018] find large sector-level economies of scale in the "Computers and Electronics" manufacturing sector.

In addition, the size of the tax incentive (as a share of the wage) is larger between 2001 and 2004 than after 2004 (hence after 2013). This is because – while programmers are fully exempt from the income tax in all years after 2001 – between 2001 and 2004 Romania had a progressive income tax (with a 40% marginal top tax rate), whereas, in 2005, it switched to a flat income tax rate of 16%. This is another plausible reason why the treatment effects are smaller in the 2013 exercise than in the 2001 exercise.

4 Sector-Level Cross-Country Analysis: 2001 Income Tax Break

4.1 Direct Effects on the Expansion of the IT Sector

This sector-level cross-country analysis of the growth of the IT sector relative to the rest of the economy complements the firm-level analysis from Section 3 in two ways. First, this analysis captures not only the intensive margin of growth (as the firm-level analysis does) but also the extensive margin (the number of firms in the IT sector). Appendix C.1 presents descriptive evidence on the extensive margin, suggesting that after 2001 the number of firms in the IT sector has increased faster in Romania than in comparable CEE countries. This increase has occurred through relatively higher entry rates and stable exit rates in Romania. Part of the entry is attributable to entirely new firms, with the remainder accounted for by incumbent firms that switched to the IT sector. While we lack the data to separately estimate the causal effects of the tax break on the extensive margin, the stronger effects that we will estimate here will reflect the fact that a sector-level analysis encompasses this margin as well.

Second, this synthetic control method (SCM) analysis allows us to alleviate concerns about potential confounding factors that may affect the IT sector globally in 2001 and that may be the true cause of the effects measured with firm-level data. One such confounding factor could be the dot-com crash of 2001. In its aftermath, U.S.-based companies may have chosen to mitigate some of the losses incurred during the crash by offshoring part of their operations in CEE countries. Global productivity or demand shocks to the IT sector could be another confounding factor for its growth in Romania. We therefore benchmark the growth of the IT sector in Romania to that in similar neighboring countries, that are likely to have been similarly affected by such confounding factors.

Empirical Strategy. We use SCM to measure the effect of the income tax break on the growth of the IT sector in Romania. SCM is a data-driven approach to small-sample studies proposed by Abadie and Gardeazabal [2003] and used to estimate treatment effects. Its intuition is that a weighted combination of countries provides a better comparison for Romania than any individual country. SCM makes explicit the relative contribution of each control country to the counterfactual, and the degree of similarity between Romania and synthetic Romania, in terms of pre-intervention outcomes and other predictors of post-intervention outcomes [Abadie et al., 2010]. The choice of weights is such that the resulting unit closely matches the treated unit over the pre-treatment period. Outcomes for the synthetic unit are then projected into the post-treatment period based on these weights. Inference is conducted using placebo tests. The same model is estimated on each untreated country, assuming that it was treated in 2001. The result is a distribution of placebo effects. If this procedure does not yield effects for untreated countries as large as the effects for Romania, then it is unlikely that the estimated effect for Romania is a result of chance.

Formally, SCM entails the following. Let J be the number of control countries ("the donor pool"), where J equals 13 in our case. Let $\boldsymbol{W} = (w_2, ..., w_{J+1})'$ be a $J \times 1$ vector of weights w_j, such that $w_j \geq 0$ and $\sum_{j=2}^{J+1} w_j = 1$. w_j is the weight of country j in synthetic Romania. SCM chooses W such that synthetic Romania most closely matches the real Romania before 2001 (i.e., in 1999 and 2000). Let $\boldsymbol{X_1}$ be a $(K \times 1)$ vector of pre-2001 values of K predictors for the relative growth of the Romanian IT sector. Similar to $\boldsymbol{X_1}$, we define $\boldsymbol{X_0}$ as the $(K \times J)$ matrix containing the values for the same predictors for the J candidate control countries. Let V be a diagonal matrix with non-negative components, whose diagonal elements represent the relative importance of these three predictor variables in the construction of synthetic Romania. The vector of weights $\boldsymbol{W}$ is chosen to minimize the objective function $(\boldsymbol{X_1} - \boldsymbol{X_0}\boldsymbol{W})'\boldsymbol{V}(\boldsymbol{X_1} - \boldsymbol{X_0}\boldsymbol{W})$, such that $w_j \geq 0$. As in Abadie and Gardeazabal [2003], Abadie et al. [2010], we allow for the choice of the weighting matrix, $\boldsymbol{V}$, to be data-driven. $\boldsymbol{V}$ allows for the pre-2001 outcomes of Romania to be closest to the outcomes of the synthetic control obtained from $\boldsymbol{W}^*(\boldsymbol{V})$.[24]

Treatment effects, α_{1t}, for Romania in post-treatment years t (2002 to 2015) are estimated as the difference between the year t outcomes for Romania and those for synthetic Romania:

$$\hat{\alpha}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}, \tag{3}$$

where Y_{1t} is the year t outcome for Romania and Y_{jt} is the same outcome in year t for control country j.

Concretely, we first contrast the within-country growth of the IT sector to the growth of all other sectors in the economy. Through SCM, we then compare this relative growth of the IT sector in Romania to the relative growth of the IT sector in synthetic Romania. An advantage of SCM is that it delivers the optimal set of weights to construct synthetic Romania and limits the researchers' degrees of freedom in the choice of the comparison group. Last, we ask whether the relative growth of the IT sector in Romania is exceptional compared to the relative growth experienced by the sector in untreated countries.

Data. The data source for the dependent variables is Eurostat, Structural Business Statistics. We require from these variables to be available in 1999 and 2000 for both Romania and all other countries in the donor pool. Also, we require these variables to appear consistently in the following years and meet minimal data quality standards. Finally, we want these variables to be relevant for this study. These conditions are met by the following three variables: number of employees, gross revenues (turnover or gross premiums written, in million euros) and production value (in million euros). Data for the predictors comes from the World Bank, World Development Indicators. We use as predictors the GDP per capita (constant LCU), the share of manufacturing value-added coming from the medium and high-tech sector, and the share of GDP coming from services.

Due to frequent missing data at the three-digit level, the IT sector is defined more broadly than the set of eligible three-digit codes, as K72 (Computer and related activities). We use as comparison sectors

[24] Throughout the paper, we implement SCM with the help of the **synth** and **synth_runner** packages in Stata [Quistorff and Galiani, 2017]. We depart from the default option of these packages by selecting the *nested* option. Hence **synth** embarks on a fully nested optimization procedure that searches among all (diagonal) positive semidefinite $\boldsymbol{V}$ matrices and sets of $\boldsymbol{W}^*$-weights for the best fitting convex combination of the control units. The fully nested optimization contains the regression based $\boldsymbol{V}$ as a starting point, but produces convex combinations that achieve even lower mean squared prediction error.

all other sectors in the economy. To obtain normalized values for the outcome variables in each year and country, the yearly absolute value of the variable in the treated sector is divided by its value in 2000, the year prior to the introduction of the policy in Romania. From these resulting yearly ratios we subtract the corresponding ratios for the comparison sectors.[25]

The donor pool of countries for the synthetic control contains Bulgaria, the Czech Republic, Estonia, Hungary, Ireland, Latvia, Lithuania, Poland, Portugal, the Slovak Republic, and Slovenia. These countries were chosen based on their geographic proximity, similarity in development, performance in the IT sector pre-2001, and data availability. See Appendix D for details on data construction.

Baseline Results. Figures 3 and 4 present the output of our SCM analysis for two dependent variables: gross revenues and employment (both normalized). The upper left panels show the evolution of these outcomes in Romania and synthetic Romania. Before 2001, the growth of the IT sector of synthetic Romania closely mimics that of Romania. From 2001 onward, both the gross revenues and employment of the Romanian IT sector experience a marked relative growth. The upper right panels show the difference between the outcomes of Romania and those of synthetic Romania. Fourteen years after, the gross revenues (employment) in the IT sector in Romania had expanded 6.52 (1.83) times more relative to the gross revenues (employment) in all other sectors, relative to the year 2000, and relative to the corresponding relative growth in synthetic Romania.

The lower left panels plot the raw paths of these normalized outcomes of the IT sector in Romania and the 11 donor countries. We notice how exceptional the growth was in Romania, compared to that in all donor countries. Last, we implement a battery of placebo tests that considers all other donor countries as potentially treated and proposes synthetic controls for each. Reassuringly, the lower right panels show that the relative growth for Romania is starker than the relative growth for all other donor countries. For at least until 2008, the actual treatment differences of gross revenues and employment growth for Romania lie outside the range of placebo differences. Formally, these results are confirmed by the almost-zero p-values until 2008 (see Table C1, Appendix C.2.1).

One might be concerned that synthetic Romania is an unreasonable proposition of SCM. Synthetic Romania is a combination of Bulgaria, the Czech Republic, and Slovakia (with weights varying with the outcome variable). Table C2 shows that synthetic Romania is reasonably similar in terms of the share of services in GDP and the share of high-tech manufacturing in total manufacturing value added. While synthetic Romania is different to Romania in terms of its GDP, our SCM analysis is relative to each country's level in 2000. Hence, this proposal of synthetic Romania seems appropriate.

Robustness Checks. All robustness checks figures and tables can be found in Appendix C.2.1. We first show that these findings are not unique to gross revenues and employment. Figure C6 (and its associated Tables C1 and C2) shows similar patterns of outstanding growth in the IT sector of Romania, this time in terms of production value. Again, we employ permutation methods to assess the statistical likelihood of our results. In the first seven years after the introduction of the policy, almost-zero p-values allow us to rule out a treatment effect of zero. Visually, the lower right panel of Figure C6 shows that the actual difference for Romania is consistently above the upper limit of placebo differences.

[25] A value of 2 in year t means that the multiplication factor of the value of the dependent variable in the treated sector in t, relative to 2000, is larger by 2 units than the counterfactual multiplication factor in comparison sectors.

One concern with deriving results from the entire 1999 to 2015 time series of Eurostat data may come from the need to rely on a crosswalk between NACE Rev 1.1 and Rev 2 sector codes. Because pre-2007 data is reported for NACE Rev 1.1 sectors and post-2007 data for NACE Rev 2 sectors, one needs a crosswalk to stitch together the time series. As the relationship between classifications is not bijective, there is no widely-used crosswalk. Appendix D provides details on own crosswalk construction.

To test whether our results are driven by the use of the full time series, we truncate the time series in 2007. We can thus study the effect of the 2001 policy using data that is consistently reported in one classification. Fortunately, we find that results for pre-2007 years are not affected by the addition of post-2007 years (see Figure C7). However, this finding does not imply that post-2007 results are not affected by the stitching of the sector-level time series. We cannot distinguish whether our weaker post-2007 results are driven by the differential effects of the financial crisis on the IT sector (compared to the rest of the economy), by a later introduction by donor countries of other policies that also favor the IT sector, or by an imprecise stitching of the sector-level time series.

A last robustness check is one in which we would contrast the growth of the IT sector to that of the same three-digit comparison sectors used in the firm-level analysis. A first constraint comes from the fact that Eurostat data is at the two-digit level. Second, the two-digit sectors 64 and 92 cannot be used, as their data is frequently missing across years and countries. Third, sector 73 is small and with noisy data, and sector 74 contains several three-digit sectors other than those we use as control to sector 722, sectors we believe are dissimilar to 722. Despite these caveats, sector 72 still exhibits a faster growth than that of comparison sectors 73 and 74. Results are available upon request.

4.2 Downstream Effects of the Expansion of the IT Sector

Given that, in 2017, the IT sector accounted for less than 2% of Romania's total employment and that only specific workers in this sector are eligible for the income tax break, one might question the wider effects of the policy. The IT sector is a sector whose inputs are broadly used by households and firms in all sectors. While the development of the IT sector most likely led to level effects as well (given its broad use), we propose a research design that allows us to credibly estimate the differential effects of the policy on sectors that relied more heavily on IT services relative to sectors that relied less.

The development of the IT sector after 2001 is likely to have boosted the development of sectors relying more heavily on IT services in two ways.[26] First, after the tax break, labor productivity (measured as revenues per worker) increased. As labor is the main input in the production of IT services, this is also likely to have improved the quality of IT services. Hence, the tax break is likely to have lowered the quality-adjusted price of IT services. Under a plausible market structure and well-behaved cost and demand conditions, this should lead to output increases in sectors purchasing more inputs from the targeted (IT) sector [Lane, 2017]. To the extent that the increase in the quality of IT inputs was not fully priced, then IT-using sectors not only experienced increases in output but also in productivity.

[26] One scenario in which this policy might have *hurt* the development of sectors relying more heavily on IT services, is one in which firms in these sectors used to produce programming services in-house. As programmers only benefit from the tax break if working for a firm in the eligible IT sector, this may have made them less likely to join ineligible sectors. This scenario is unlikely, as it was uncommon for firms in non-IT sectors to develop software in-house. Note that employees who work in maintenance are not eligible for this tax break, irrespective of their employer's sector.

Second, the IT sector has also expanded through the entry of new firms (see Section 4.1 and Appendix C.1). Figure C5 (Appendix C.1.3) suggests that part of this entry occurred through foreign direct investment in the IT sector of Romania, which has intensified since the early 2000s. This implies that the IT sector has considerably expanded the set of varieties proposed to downstream sectors. Whenever downstream sectors have a love of variety for intermediate inputs, this expansion in varieties leads to productivity gains.[27] In addition, varieties proposed by foreign-owned firms are likely to have been of higher quality than those proposed by domestic firms.

All in all, improvements in the prices, quality, and variety of IT inputs are likely to have provided a boost to downstream sectors relying more on IT. We study not only the evolution of the size of these sectors, but also their export performance. To the extent that a stronger IT sector generates productivity gains for IT-using downstream sectors, this is likely to shift trade patterns.

Empirical Strategy. To study the effect of the 2001 tax break on the expansion of downstream sectors, we employ a similar SCM to the one described in Section 4.1. The only difference between these exercises is in the definition of treated and comparison sectors. Hereafter, we define treated sectors as those downstream sectors for which the IT sector is most important as the supplier of inputs. Conversely, comparison sectors are those relying relatively less on the IT sector as an input supplier.

Data. We start from the input-output table (I-O table, henceforth) of Romania for 2000. We use the harmonized I-O table provided by the OECD, which tracks the flows of goods and services between all two-digit NACE Rev 1 sectors. Given that 2000 is the year before the unexpected introduction of the policy, inter-sector linkages are not yet affected by this policy. We then compute the share of the total input expenditures of a given sector purchased from the IT sector (NACE Rev 1 sector 72, "Computer and related activities"). Based on these shares we identify the sectors for which IT services are the most important inputs in 2000. Based on their position in this sector-level distribution of shares, we assign sectors to either a high- or low-intensity category of use of IT services. The treated high-intensity category contains sectors that are among the top 25% users of IT services. All other sectors lying in the bottom 75% constitute the control category.[28] See Table C3 in Appendix C.2.2 for details.

Similar to the SCM in Section 4.1, we rely on Eurostat data to construct the same normalized dependent variables and World Bank data for the same predictor variables. In addition, we use UN Comtrade data to study the export performance of sectors relying more on IT services, relative to those relying less. Given data availability, we use SCM to study the exports of goods alone. One notable advantage of UN Comtrade data is that it starts in 1996, offering three more years than the Eurostat data of pre-treatment years. See Appendix D for details on data construction.

Baseline Results. We first ask whether sectors with stronger upstream linkages to the IT sector experienced a more pronounced growth than sectors with weaker linkages. Figures 5 and 6 provide a visual answer for gross revenues and employment. Fourteen years after the introduction of the policy, gross revenues (employment) in sectors with high-intensity use of IT services has grown 0.75 (0.61) times

[27] See Rodríguez-Clare [1996], Goldberg et al. [2010], Carluccio and Fally [2013], Kee [2015]

[28] We also calculate the share of the total sales of the IT sector purchased by each sector. Again, we assign sectors to quarters based on these new shares. While these two classifications are conceptually different, given the I-O table of Romania for 2000, we find that there is no practical difference in the final split of sectors between the top and bottom three quartiles.

more than gross revenues (employment) in low-intensity sectors (compared to year 2000 and compared to the equivalent difference in synthetic Romania). When implementing the permutation method suggested by Abadie et al. [2010], we find that our SCM estimates lie at the upper limit of the distribution of placebo estimates. The low p-values in Table C4 (Appendix C.2.2) rule out null effects.

Next, we study the export performance of sectors relying more on IT services, compared to those relying less. Because high-intensity sectors (defined until now as those over the third quartile of the IT-usage intensity distribution) are all service sectors, we now define high-intensity sectors as those manufacturing sectors between the second and third quartile. Figure 7 depicts a striking relative growth in the export trade value of goods from high-intensity sectors in Romania (relative to those from low-intensity sectors and relative to synthetic Romania).[29] Placebo tests show that this relative growth in Romania is exceptional compared to that predicted for all other countries in the donor pool.[30] Table C6 (Appendix C.2.2) makes the same argument, formally. This evidence suggests that the development of the IT sector in Romania not only increased the output of IT-using sectors, but also improved their comparative advantage.[31]

Robustness Checks All robustness checks figures and tables can be found in Appendix C.2.2. We first show that our baseline findings on the relative growth of downstream sectors with heavier usage of IT services are not specific to gross revenues, employment, or export value. For instance, we find that the production value of IT-using sectors has also grown significantly more in Romania (see Figure C8).

Second, the SCM findings presented above are robust to different choices of the weighting matrix $\boldsymbol{V}$ and matrix of weights $\boldsymbol{W}$. Another potential concern with SCM relates to its sensitivity to the number of pre-intervention periods used in the computation of the weights. We cannot rule out this concern with the Eurostat data, as the panel only starts in 1999. However, a benefit of Comtrade data is that it allows us to observe export patterns since 1996. We run the Comtrade SCM exercise varying the number of pre-treatment years used in the estimation. Reassuringly, treatment effects remain unaltered.

Third, one might be concerned that our results are driven by specific high-intensity IT using sectors. From the beginning, in the analysis using Eurostat data, we exclude NACE Rev 1 sector 72, as this two-digit sector contains the three-digit sector 722 eligible for the income tax break. This avoids the risk of a mechanical result. In addition, as a robustness check, we also exclude NACE Rev 1 sectors 73 and 74; these sectors belong to the list of high-intensity IT users, but also contain three-digit NACE Rev 1 sectors that we use as comparison sectors in the firm-level analysis. If sectors 73 and 74 – or their subset of three-digit sectors comparable to sector 722 – were experiencing correlated shocks with those of sector 722, our findings could be affected by such shocks. Figure C9 shows that when we exclude

[29] Most of this growth is explained by the SITC Rev 1 commodity codes 54 (Medicinal and pharmaceutical products), 62 (Rubber manufactures), 73 (Road vehicles, other than motor vehicles), and 86 (Watches and clocks).

[30] Figure C12 (Appendix C.2.2) shows that after 2000 the exports of service sectors relying more on IT services also grew noticeably faster in Romania than in comparable countries. Among these sectors, those under NACE Rev 1 sector 74 (e.g., call centers, advertising, business and management consultancy, secretarial and translation activities etc.) experienced the most impressive growth. Romania's trend is compared to that of the five countries that constitute the typical synthetic Romania in all SCM exercises thus far, i.e. Bulgaria, the Czech Republic, Hungary, Lithuania, and Slovakia.

[31] The behavior of FDI flows to Romania also supports this claim. Figure C5 (Appendix C.1.3) shows that FDI in high-intensity IT using sectors grew faster than FDI in low-intensity sectors. While not the only driver behind this relative growth in FDI, Romania's IT sector is frequently mentioned among those that are most significant. For instance, this article describes how Romania's IT sector played an important role in Renault's 2007 decision to build its Technocentre in Romania.

sectors 73 and 74 the treatment effect is actually larger than the one found in the baseline Figure 6.[32]

Fourth, we also show that results have qualitatively similar patterns when we change the threshold of the grouping of sectors into the high- and low-intensity categories. Figure C10 presents results from the grouping of sectors under the median of usage of inputs from the IT sector into the low-intensity category (as opposed to under the third quartile). As expected, while the difference in the development of the high and low intensity categories becomes less stark, the general pattern is maintained.

Fifth, one might also worry that relying on Romania's I-O table from 2000 to construct the high- and low-intensity treatment categories is a concern in itself. As an alternative, we use the classification of sectors proposed by van Ark et al. [2003]. Sectors are assigned one of the following six categories based on U.S. measures of pre-2000 ICT (information and communication technology) intensity from Stiroh [2002]: ICT-producing manufacturing, ICT-producing services, ICT-using manufacturing, ICT-using services, non-ICT manufacturing, and non-ICT services. We exclude sectors in ICT-producing manufacturing and ICT-producing services, as they might be directly affected by the tax break. We group sectors in ICT-using manufacturing and ICT-using services into the high-intensity category, and sectors in non-ICT manufacturing and non-ICT services into the low-intensity category. While the patterns obtained with this grouping are noisier than those obtained with our preferred grouping, we still find a stronger relative growth in Romania in ICT-using sectors compared to non-ICT using, and compared to synthetic Romania. We assess our initial grouping to be superior, as it is more narrowly defined around the treated sector (NACE Rev 1 sector 722) than the one proposed by Stiroh [2002].[33]

Finally, we check whether our results are robust to the exclusion of the second half of the Eurostat time series. As explained in Section 4.1, the lengthening of sector-level time series to include 2007 to 2015 relies on an inherently-imprecise crosswalk between the NACE Rev 2 and NACE Rev 1 classifications. Figure C11 shows that results for the years under the NACE Rev 1 classification (1999 to 2006) are identical to those for the same years obtained using the full time series (1999 to 2015). This concern does not affect the Comtrade SCM results, as the Comtrade data is available in a unique classification.

In addition to these robustness checks, the timing of the relative growth of IT-using sectors speaks against concerns of reverse causality, i.e., it is the development of downstream sectors using IT intensively that actually boosted the development of the IT sector.[34]

4.3 Discussion of the Sector-Level Cross-Country Findings

In Section 4.1, we show that, since 2001, the IT sector in Romania has grown significantly faster compared to the rest of the sectors in Romania and compared to the same relative growth in similar countries. This finding gives us confidence that the effects we measure are plausibly caused by the tax break to programmers introduced in Romania, and not by other global supply- or demand-side shocks to the IT sector. This (relative) growth of the IT sector has occurred both on the intensive (through the growth of incumbent firms) and extensive margin (through the entry of new firms in the sector).

[32] The main SCM exercise using Comtrade data excludes, by construction, all service sectors, hence sectors 72, 73, and 74.

[33] ICT contains several other (significantly larger) sectors than 722, unrelated to the policy we study.

[34] This does not exclude the possibility of a feedback loop between the development of IT-using sectors and the IT sector itself. In a 2016 *Reuters* article, Florin Talpeş (a pioneer in Romania's IT sector) advised new entrants in the IT sector to focus on developing technology for the now-mature automotive sector (e.g., driver-less technology or car connectivity).

In Section 4.2, we find that since 2001, IT-using sectors in Romania have grown significantly faster compared to non-IT using sectors in Romania and the same relative growth in similar countries. This pattern provides support to the conjecture that improvements in the quality-adjusted prices and variety of IT services benefit more sectors relying more heavily on these services.

Several pieces of evidence lend credibility to the magnitude of our SCM estimates. First, while the estimated magnitude of the relative growth rate of the IT sector in Romania might seem impressive, this magnitude is lower than the actual growth of the IT sector in Romania. Our SCM estimate for 2015 for the number of workers is 1.83, i.e., the number of workers in IT grew 1.83 times faster between 2000 and 2015 compared to the number of workers in the rest of the Romanian economy and compared to the same relative growth in synthetic Romania. In the raw data, the employment in IT in Romania grew six times, from 13,691 workers in 2000 to 81,780 workers in 2015. Similarly, gross revenues grew 14-fold over the same period. Hence, our estimates attenuate the actual growth of the IT sector in Romania after 2001, as they control for broader trends in the rest of the Romanian economy and similar economies. Second, consistent with the firm-level results, we find stronger increases in revenues in the IT sector than in employment. Third, the relative growth of downstream sectors with stronger links to the IT sector is not as large as the relative growth of the IT sector itself. As one would expect, the sector directly receiving the tax incentive grows faster than downstream sectors benefiting from the incentive indirectly.[35]

It is important to emphasize that there are features of either the available data or of the empirical strategies that do not recommend a direct comparison of the sector-level and the firm-level evidence. First, due to data constraints, the sector-level and firm-level growth rates in IT are measured relative to different comparison groups.[36] The choice of the reference group can affect the exact magnitude of the estimates. Second, the sector-level evidence also allows for growth through firm entry, whereas the firm-level evidence characterizes incumbent firms alone.

Finally, the time frame of analysis is also important for the magnitudes and their interpretation. We conduct each firm-level exercise in a relatively narrow time window around the two policy of interest: 1997 to 2005 for the 2001 introduction of the income tax break, and 2011 to 2015 for the 2013 amendment to the tax break. To improve identification, the firm-level analysis, therefore, uses short-term variation in the tax conditions of firms in the IT sector in Romania. In the long term, there are general equilibrium effects (such as those on the supply of programmers) or unrelated shocks (such as the global financial crisis) that would have hindered the interpretation of long-term firm-level growth estimates.

By virtue of their long time frame (1999 to 2015), the sector-level cross-country findings are likely to capture not only the direct effects of the initial introduction of the policy in 2001, but also those of other developments in Romania and abroad that differentially affect the IT sector of Romania since 2001. For instance, to the extent that in the early 2000s the development of the IT sector in Central and Eastern Europe was at the cusp of multiple potential equilibria, this policy is likely to have acted as a timely signal to both local and foreign firms. In a world with first-mover advantage and path dependence, this policy is likely to have tilted the balance towards an equilibrium favorable to the IT sector of Romania. In

[35] Figure C5 (Appendix C.1.3) confirms the same intuition: the IT sector itself is the one becoming more attractive for FDI, followed by downstream sectors using IT services intensively.

[36] The sector-level cross-country data is at the two-digit level, with frequent missing values in the two-digits containing the three-digit codes used for comparison in the firm-level analysis. Hence, our baseline sector-level estimates are with respect to the rest of the economy, whereas the firm-level estimates are with respect to firms in certain three-digit sector codes.

this light, it is likely that a sizable part of the subsequent growth of the IT sector is due to a snowballing effect of the signal of the policy, as opposed to the size of the actual incentive it provides.

Relatedly, the magnitude of the effects on IT-using sectors captures more than the indirect incentive granted by this tax break. It also reflects the idea that in an economy with coordination failures – due to economies of scale and imperfect tradability of services associated to skill-intensive manufacturing – government policy can move the economy towards the "high-wage, high-tech equilibrium" [Rodrik, 1996]. Moreover, it suggests the possibility of strong complementarities between this industrial policy and the FDI attraction and trade opportunities that followed Romania's joining of the EU in 2007 [as in Topalova and Khandelwal, 2011]. While Romania's comparison countries also joined the EU in 2004 or 2007, only Romania saw such a distinctive growth in the exports of IT-intensive downstream sectors.

To conclude, while it is outside the scope of this paper to disentangle the direct effects of this income tax break and those circumstances that may have amplified or dampened its effects in the aggregate and in the long-run, it is reassuring that the sector-level cross-country evidence and the firm-level evidence paint an overall consistent picture by which the income tax break has been effective in its objective to boost the development of the IT sector in Romania.

5 Back-of-the-Envelope Cost Estimates of the Policy

One key policy concern is the cost of this policy to the government, given that the IT workers who benefit from this income tax exemption are among the highest-paid workers in Romania. Another concern might be that our findings of large and persistent firm- and sector-level growth after the introduction and extension of the income tax break may seem disproportionate compared to the tax break itself. To address these concerns, we estimate the cost of this policy to the government or put differently, the implicit incentive to the IT sector. Under the assumption that the administrative costs of this policy are negligible (both to firms and the government), we define this cost as the tax revenues foregone due to the exemption of the flat personal income tax of 16% owed on the taxable income of all exempted workers. We estimate this cost for 2015, the last year of our firm-level analysis, assuming that the tax exemption was canceled unexpectedly at the beginning of 2015.

We assume that exempted employees would retain full-time employment in the same or related sectors. Then, in our first scenario we assume that in the short term, gross wages would remain unchanged, as employment contracts cannot be easily changed in the short-term. Therefore, our first estimate uses the actual average wages of the exempted employees. In the absence of the policy, the currently exempted employees may also move to other ineligible HTKI service sectors or other ineligible ICT service sectors. Our second and third estimates are based on the average wages in those two types of service sectors. Table 6 shows our estimates for these three scenarios.

Overall, the estimates show that the tax receipts foregone in 2015 due to the policy vary between 56 and 92 million euros. To put these numbers into perspective, they represent between 2.7 and 4.5% of the value added of the IT sector, and between 4.5 and 7.4% of its gross wage bill. Note that we compute these shares out of the value added or wage bill of the entire sector, including firms with little to no exemption of their employees. For firms with a large share of employees exempt from the income tax, the share between the "forgiven" tax bill and their value added or wage bill is significantly larger. We conclude that the income tax exemption is not only a signal of commitment of the Romanian government to the development of the IT sector but is also a sizable incentive. This incentive is also likely to act as an indirect incentive to other sectors, particularly those sectors that rely heavily on IT services.

6 Conclusion

This paper examines the effects of a unique industrial policy introduced by Romania in 2001: a personal income tax break for workers with specific IT-relevant bachelor's degrees and who work directly on software development for a firm with an eligible IT sector code. In 2013, the tax break law was amended to allow for a significantly larger list of eligible sector codes for firms and eligible bachelor's degrees for workers. We use both policy episodes to bring plausibly-causal estimates of their firm and sector-level effects. Across empirical strategies and measures of firm size, we find that the policy and its reform led to strong and lasting growth for IT firms. This is in line with a shared economic incidence of the tax incentive between firms and workers. Sector-level cross-country evidence makes the additional point that the growth after 2001 of the IT sector in Romania is unlikely to be driven by confounding factors, as similar neighboring countries fail to show comparable growth. Moreover, we find evidence of wider benefits from this policy for the Romanian economy. Namely, we find that sectors relying relatively more on IT inputs grew faster than those relying less.

Our results suggest that this policy has been effective in its goal to support the development of the IT sector and to reallocate resources (such as high-skill labor) towards this sector. This is an important achievement, because many industrial policies only allow specific groups to extract rents, without actually affecting resource allocations. Moreover, this shift was made towards a high-skill/high-wage sector, a policy priority in both developed and developing countries.

Establishing whether this policy was also efficient is beyond the scope of this project, but a fruitful area for future research. A number of special features of this policy motivate such additional research. First, this policy was not designed to rescue a "sunset sector" (as has been the case in several East Asian industrial policies). On the contrary, it targeted a sector under-developed in Romania at the time, but generally seen as crucial for growth. Second, the effects on IT-using downstream sectors suggest that this policy may have also mitigated (inter-sector) coordination failures coming from scale economies and imperfectly-tradable services (here, IT services) useful for skill-intensive downstream sectors [as in Rodrik, 2004]. Last, this policy involved reductions in labor taxes, as opposed to the vastly more common reductions in corporate taxes or state aid packages. With worker-level data, one could study the distributional implications of this policy design relative to the common designs.

References

Alberto Abadie and Javier Gardeazabal. The Economic Costs of Conflict: A Case Study of the Basque Country. American Economic Review, 93(1):113–132, March 2003.

Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. Journal of the American Statistical Association, 105(490):493–505, 2010.

Philippe Aghion, Jing Cai, Mathias Dewatripont, Luosha Du, Ann Harrison, and Patrick Legros. Industrial Policy and Competition. American Economic Journal: Macroeconomics, 7(4):1–32, October 2015.

Phillipe Aghion, Julian Boulanger, and Elie Cohen. Rethinking Industrial Policy. Bruegel Policy Brief, 4, 2011.

Adam Banai, Peter Lang, Gabor Nagy, and Martin Stancsics. Waste of Money or Growth Opportunity: The Causal Effect of EU Subsidies on Hungarian SMEs. Economic Systems, 44, 2020.

Dominick Bartelme, Arnaud Costinot, Dave Donaldson, and Andres Rodriguez-Clare. External Economies of Scale and Industrial Policy: A View from Trade. Working Paper, 2018.

Paul Beaudry, Mark Doms, and Ethan Lewis. Should the Personal Computer Be Considered a Technological Revolution? Evidence from U.S. Metropolitan Areas. Journal of Political Economy, 118(5):988–1036, 2010.

Bruce A Blonigen. Industrial Policy and Downstream Export Performance. Economic Journal, 126(595):1635–1659, 2016.

Nicholas Bloom, Raffaella Sadun, and John Van Reenen. Americans Do IT Better: US Multinationals and the Productivity Miracle. American Economic Review, 102(1):167–201, 2012.

Philipp Boeing. The allocation and effectiveness of chinas r&d subsidies – evidence from listed firms. Research Policy, 45(9):1774–1789, 2016.

Jing Cai and Ann Harrison. Industrial Policy in China: Some Unintended Consequences? Industrial and Labor Relations Review, 2019.

Juan Carluccio and Thibault Fally. Foreign Entry and Spillovers with Technological Incompatibilities in the Supply Chain. Journal of International Economics, 90(1):123–135, 2013.

Zhao Chen, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Xu. Notching R&D Investment with Corporate Income Tax Cuts in China. Working Paper, November 2018.

Reda Cherif and Fuad Hasanov. The Return of the Policy That Shall Not Be Named: Principles of Industrial Policy. Technical report, International Monetary Fund, 2019.

Chiara Criscuolo, Ralf Martin, Henry Overman, and John Van Reenen. Some Causal Effects of an Industrial Policy. American Economic Review, 2018.

Luosha Du, Ann Harrison, and Gary Jefferson. FDI Spillovers and Industrial Policy: The Role of Tariffs and Tax Holidays. World Development, 64:366–383, 2014.

Eurofound. Employment Effects of Reduced Non-Wage Labour Costs. Luxembourg, 2017.

European Commission. A Study on R&D Tax Incentives. Taxation Papers, Working Paper 52, 2014.

European Commission. Industrial Policy Strategy: Investing in a Smart, Innovative and Sustainable Industry. European Commission Policy Report, 2017.

Chris Forman, Avi Goldfarb, and Shane Greenstein. The Internet and Local Wages: A Puzzle. American Economic Review, 102(1):556–75, 2012.

Rikard Forslid and Karen Helene Midelfart. Internationalisation, Industrial Policy and Clusters. Journal of International Economics, 66(1):197–213, 2005.

Rigmor Kringelholt Fowkes, João Sousa, and Neil Duncan. Evaluation of Research and Development Tax Credit. HM Revenue and Customs HMRC Working Paper, 17, 2015.

Sourafel Girma, Holger Görg, Eric Strobl, and Frank Walsh. Creating Jobs through Public Subsidies: An Empirical Analysis. Labour Economics, 15:1179–1199, 2008.

Pinelopi Koujianou Goldberg, Amit Kumar Khandelwal, Nina Pavcnik, and Petia Topalova. Imported Intermediate Inputs and Domestic Product Growth: Evidence from India. Quarterly Journal of Economics, 125(4):1727–1767, 2010.

Austan Goolsbee. Does Government R&D Policy Mainly Benefit Scientists and Engineers? American Economic Review, 88:298–302, 1998.

Holger Görg, Michael Henry, and Eric Strobl. Grant Support and Exporting Activity. Review of Economics and Statistics, 90(1):168–174, 2008.

Bruce Greenwald and Joseph E Stiglitz. Helping Infant Economies Grow: Foundations of Trade Policies for Developing Countries. American Economic Review, 96(2):141–146, 2006.

Irem Guceri. Will the Real R&D Employees Please Stand Up? Effects of Tax breaks on Firm-Level Outcomes. International Tax and Public Finance, 25:1–63, 2018.

Ann Harrison and Andrés Rodríguez-Clare. Trade, Foreign Investment, and Industrial Policy. Handbook of Development Economics, 5, 2010.

Albert Hirschman. The Strategy of Economic Development. Yale University Press, 1958.

Kenan Huremović and Fernando Vega-Redondo. Production Networks. Working Paper, 2016.

Oleg Itskhoki and Benjamin Moll. Optimal Development Policies with Financial Frictions. Econometrica, 2018.

Dale W Jorgenson, Mun S Ho, and Kevin J Stiroh. A Retrospective Look at the US Productivity Growth resurgence. Journal of Economic perspectives, 22(1):3–24, 2008.

Omar Joya and Eric Rougier. Do (All) Sectoral Shocks Lead to Aggregate Volatility? Empirics from a Production Network Perspective. European Economic Review, 113:77–107, 2019.

Réka Juhász. Temporary Protection and Technology Adoption: Evidence from the Napoleonic Blockade. American Economic Review, 2018.

Ulrich Kaiser and Johan M. Kuhn. Worker-Level and Firm-Level Effects of a Wage Subsidy Program for Highly Educated Labor: Evidence from Denmark. Research Policy, 45:1939–1943, 2016.

Şebnem Kalemli-Özcan, Bent Sørensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcan Yeşiltaş. How to Construct Nationally Representative Firm Level Data from the ORBIS Global Database. Note, 2015.

Aki Kangasharju. Do Wage Subsidies Increase Employment in Subsidized Firms? Economica, 74:51–67, 2007.

Hiau Looi Kee. Local Intermediate Inputs and the Shared Supplier Spillovers of Foreign Direct Investment. Journal of Development Economics, 112:56 – 71, 2015. ISSN 0304-3878.

Richard Kneller and Jonathan Timmis. ICT and Exporting: The Effects of Broadband on the Extensive Margin of Business Service Exports. Review of International Economics, 24(4):757–796, 2016.

Nathaniel Lane. Manufacturing Revolutions: Industrial Policy and Networks in South Korea. Working Paper, 2017.

Nathaniel Lane. The New Empirics of Industrial Policy. Review Article Requested by Dani Rodrik for a Special Issue on Industrial Policy. Journal of Industry, Competition and Trade, 20(3), 2020.

Ernest Liu. Industrial Policies in Production Networks. Quarterly Journal of Economics, 134(4), 2019.

Boris Lokshin and Pierre Mohnen. Do R&D Tax Incentives Lead to Higher Wages for R&D Workers? Evidence from The Netherlands. Research Policy, 42(3):823 – 830, 2013.

Stefano Lombardi, Oskar Nordström Skans, and Johan Vikström. Targeted Wage Subsidies and Firm Performance. Labour Economics, 53:33–45, 2018.

Jonas Månsson and A.M.M. Shahiduzzaman Quoreshi. Evaluating Regional Cuts in the Payroll Tax from a Firm Perspective. Annals of Regional Science, 54:323–347, 2015.

Howard Pack and Larry E Westphal. Industrial Strategy and Technological Change: Theory Versus Reality. Journal of development economics, 22(1):87–128, 1986.

Madalina Pruna and John Soleanicov. Rethinking Romanian Industrial Policy. Technical report, MPA/ID Thesis, John F. Kennedy School of Government, Harvard University, 2012.

Brian Quistorff and Sebastian Galiani. The synth_runner Package: Utilities to Automate Synthetic Control Estimation Using **synth**, 2017. URL https://github.com/bquistorff/synth_runner. Version 1.6.0.

Andrés Rodríguez-Clare. Multinationals, Linkages, and Economic Development. American Economic Review, 86(4):852–873, 1996.

Dani Rodrik. Coordination Failures and Government policy: A Model with Applications to East Asia and Eastern Europe. Journal of International Economics, 40(1):1 – 22, 1996.

Dani Rodrik. Industrial Policy for the Twenty-first Century. Technical report, UNIDO Working Paper, 2004.

Dani Rodrik. Normalizing Industrial Policy. Commission on Growth and Development Working Paper No. 3, 2008.

Emmanuel Saez, Benjamin Schoefer, and David Seim. Payroll Taxes, Firm Behavior, and Rent Sharing: Evidence from a Young Workers' Tax Cut in Sweden. American Economic Review, 109(5):1717–1763, 2019.

Claudia Steinwender. Real Effects of Information Frictions: When the States and the Kingdom Became United. American Economic Review, 108(3):657–96, 2018.

Kevin J. Stiroh. Information Technology and the U.S. Productivity Revival: What Do the Industry Data Say? American Economic Review, 92(5):1559–1576, 2002.

Patricia Succar. The Need for Industrial Policy in LDC's-A Re-statement of the Infant Industry Argument. International Economic Review, pages 521–534, 1987.

Chad Syverson. What Determines Productivity? Journal of Economic Literature, 49(2):326–365, 2011.

Petia Topalova and Amit Khandelwal. Trade Liberalization and Firm Productivity: The Case of India. Review of Economics and Statistics, 93(3):995–1009, 2011.

Bart van Ark, Robert Inklaar, and Robert H. McGuckin. ICT and Productivity in Europe and the United States: Where Do the Differences Come From? Working Paper, 2003.

Bart Van Ark, Mary O'Mahony, and Marcel P. Timmer. The Productivity Gap between Europe and the United States: Trends and Causes. Journal of Economic Perspectives, 22(1):25–44, 2008.

Erik van der Marel, Fredrik Erixon, Oscar Guinea, and Philipp Lamprecht. Are Services Sick? How Going Digital Can Cure Services Performance. Global Economic Dynamics, Bertelsmann Stiftung Report, 2020.

Figures

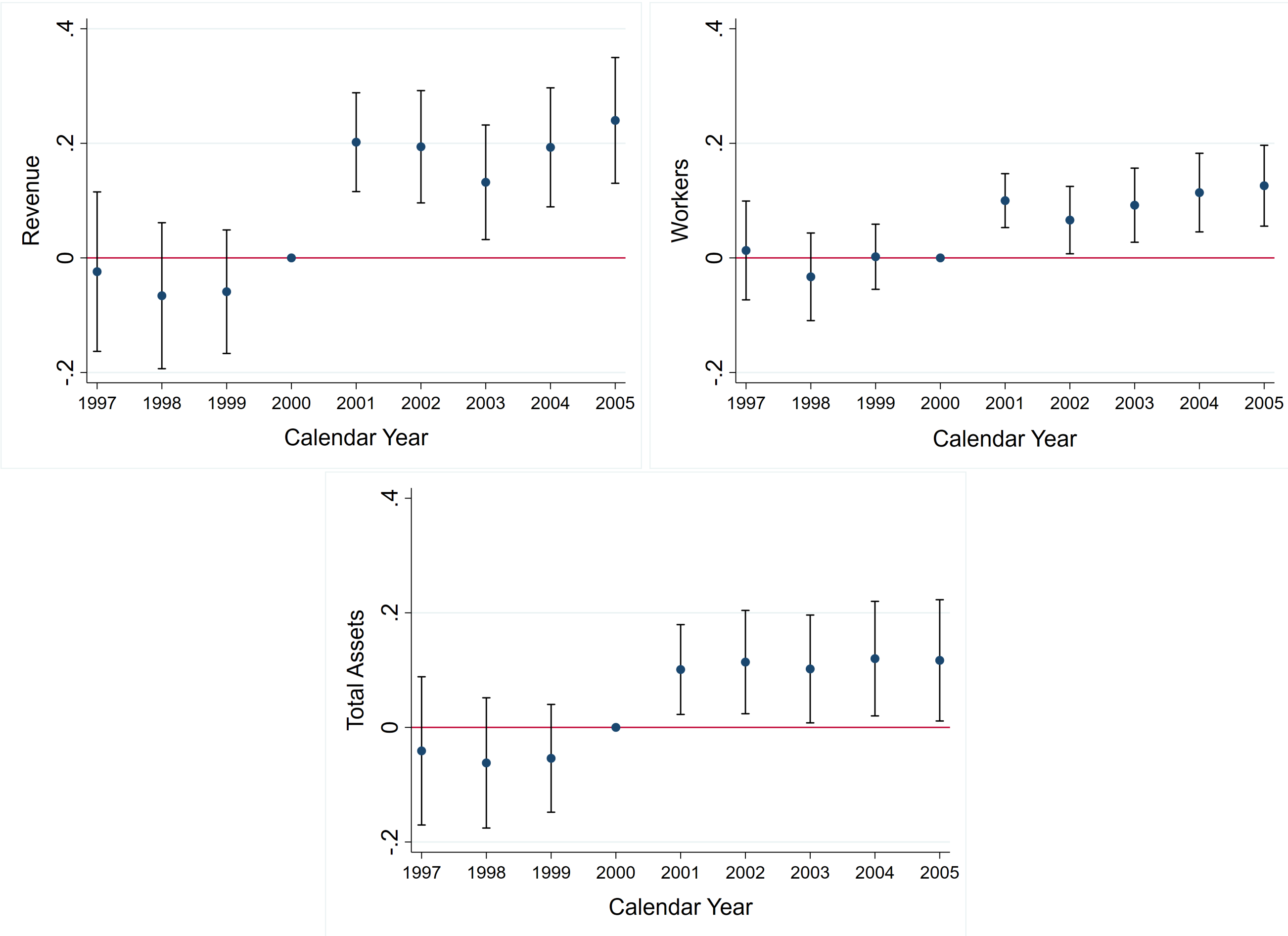

Figure 1: DiD Estimates of the Effects of the 2001 Income Tax Break

Notes: In this exercise, we study the firm-level impact of the introduction of the 2001 income tax break to workers in IT. Figure 1 plots the baseline estimates of the yearly DiD coefficients from Equation (1), $\beta_{DiD,t}$, together with their 95% confidence intervals. We consider three firm-level outcome variables: log operating revenues, log number of workers, and log total assets. The coefficients for the year 2000, the year prior to the introduction of the tax break, are normalized to zero. Treated firms are those in the NACE Rev 1 sector 722 (Software consultancy and supply). Firms in the baseline control group (used in these figures) are in high-tech knowledge-intensive service sectors (as classified by the Eurostat). All regressions in this table use Amadeus data. The regression model includes firm and calendar year fixed effects. See columns (1)-(3) from Table 1 for more details.

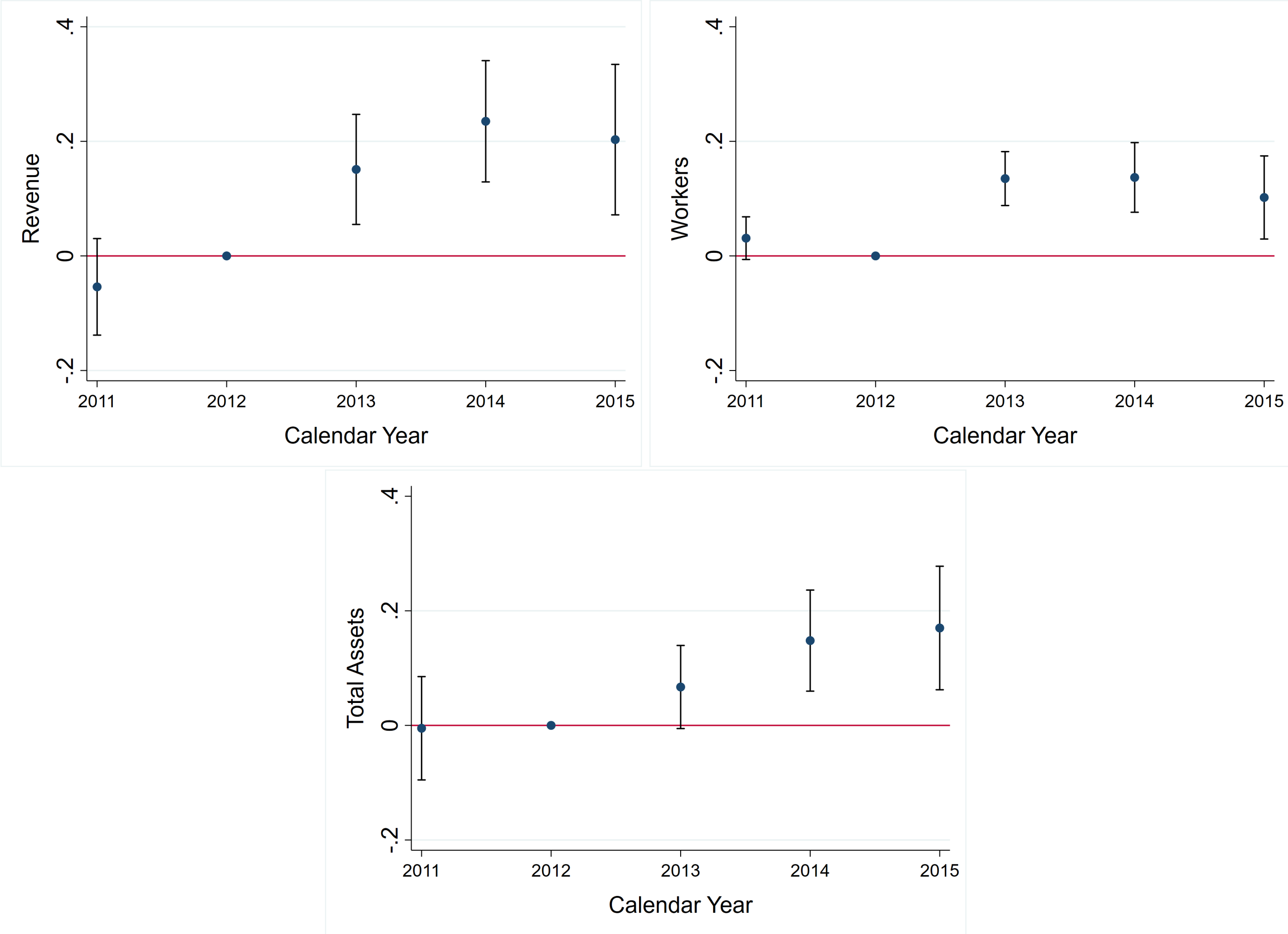

Figure 2: DiD Estimates of the Effects of the 2013 Reform

Notes: In this exercise, we study the firm-level impact of the introduction of the 2013 reform to the 2001 tax break. Figure 2 plots the baseline estimates of the yearly DiD coefficients from Equation (2), $\beta_{DiD,t}$, together with their 95% confidence intervals. We consider three firm-level outcome variables: log revenues, log number of workers, and log total assets. The coefficients for the year 2012, the year prior to the 2013 reform, are normalized to zero. Treated firms are those whose share of income tax exempted workers jumps from under 5% to over 20% after 2013. Firms in the baseline control group (used in these figures) are in high-tech knowledge-intensive service sectors (as classified by the Eurostat) and have an under 5% share of income tax exempted workers throughout the entire 2011-2015 period. This exercise builds on administrative tax data collected by the Ministry of Finance. The regression model includes firm and sector-by-year fixed effects. See columns (1)-(3) from Table 3 for more details.

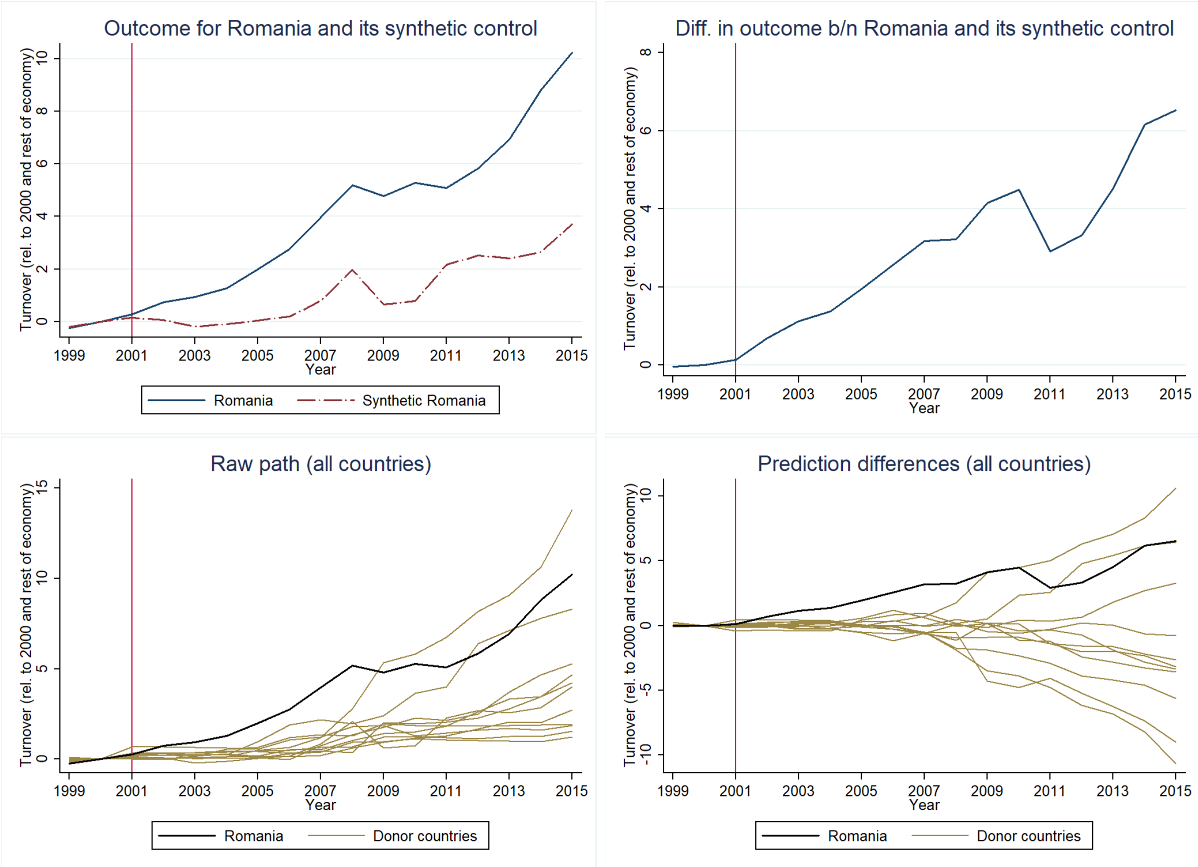

Figure 3: IT Sector Vs. Rest of the Economy. SCM with Outcome Variable: Gross Revenues ("Turnover or Gross Premiums Written") - Million Euros (Normalized)

Notes: In this exercise we use the synthetic control method introduced in Section 4.1 to study the sector-level direct effects of the introduction of the 2001 law granting an income tax break to workers in IT. All figures have as dependent variable the country-level (normalized) gross revenues ("Turnover or gross premiums written - million euros"). The yearly absolute value of the dependent variable in the treated sector is divided by its value in 2000, the year prior to the introduction of the income tax break in Romania. From these resulting yearly ratios we subtract the corresponding ratios for the comparison sectors. The treated sector is K72 ("Computer and related activities," including "Software consultancy and supply " and "Publishing of software"). We use as comparison sectors all other sectors in the economy (all except K72). The data source for the dependent variable is Eurostat, Structural Business Statistics, Annual detailed enterprise statistics on services (NACE Rev 1.1). Data for the predictors comes from the World Bank, World Development Indicators. We use as predictors the "GDP per capita (constant LCU)," "Medium and high-tech sector (% manufacturing value added)" and "Services, etc., value added (% of GDP)." All figures are an output of the **synth_runner** package for Stata [Quistorff and Galiani, 2017], with the *nested* option specified.

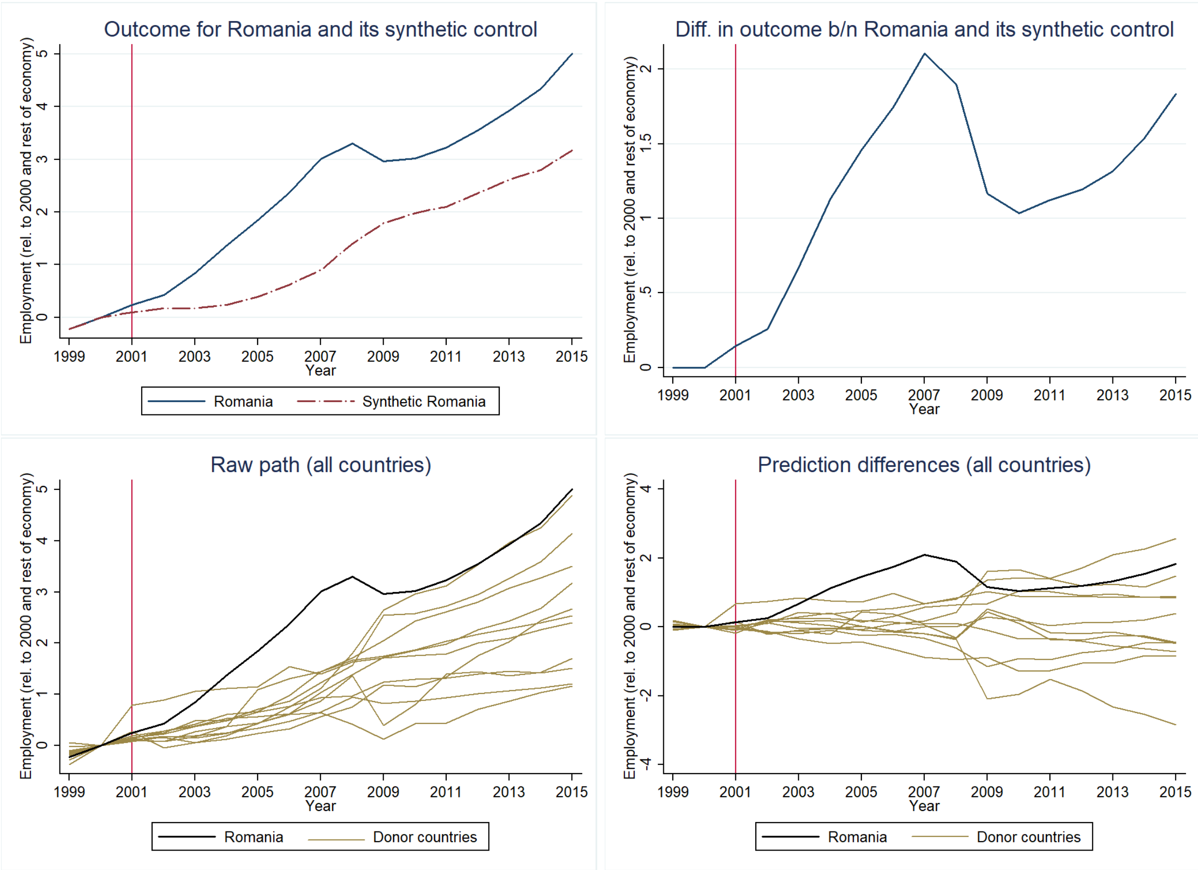

Figure 4: IT Sector Vs. Rest of the Economy. SCM with Outcome Variable: "Employees - Number" (Normalized)

Notes: In this exercise we use the synthetic control method introduced in Section 4.1 to study the sector-level direct effects of the introduction of the 2001 law granting an income tax break to workers in IT. All figures have as dependent variable the country-level (normalized) "Employees - number" The yearly absolute value of the dependent variable in the treated sector is divided by its value in 2000, the year prior to the introduction of the income tax break in Romania. From these resulting yearly ratios we subtract the corresponding ratios for the comparison sectors. The treated sector is K72 ("Computer and related activities," including "Software consultancy and supply " and "Publishing of software"). We use as comparison sectors all other sectors in the economy (all except K72). The data source for the dependent variable is Eurostat, Structural Business Statistics, Annual detailed enterprise statistics on services (NACE Rev 1.1). Data for the predictors comes from the World Bank, World Development Indicators. We use as predictors the "GDP per capita (constant LCU)," "Medium and high-tech sector (% manufacturing value added)" and "Services, etc., value added (% of GDP)." All figures are an output of the **synth_runner** package for Stata [Quistorff and Galiani, 2017], with the *nested* option specified.

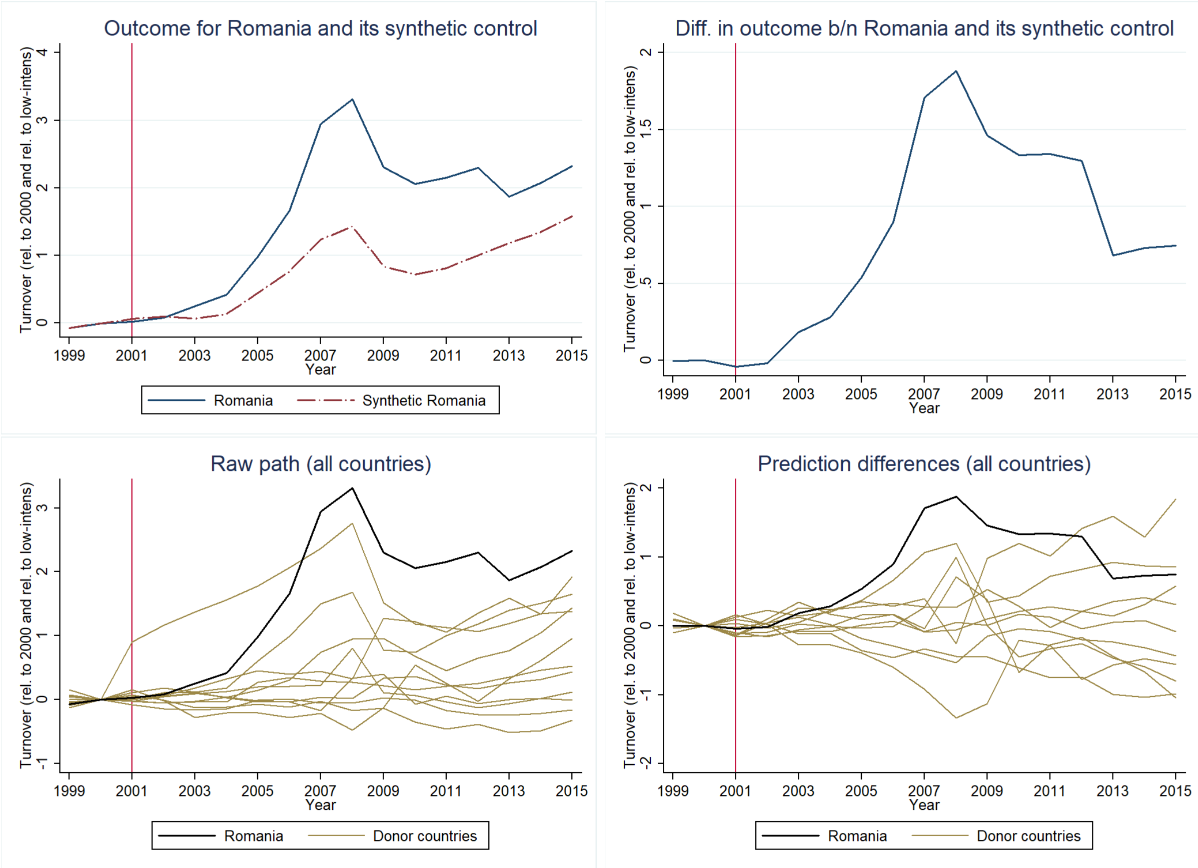

Figure 5: IT-Using Sectors Vs. Non-IT Using Sectors. SCM with Outcome Variable: Gross Revenues ("Turnover or Gross Premiums Written") - Million Euros (Normalized)

Notes: In this exercise we use the synthetic control method introduced in Section 4.2 to study the sector-level downstream effects of the introduction of the 2001 law granting an income tax break to workers in IT. All figures have as dependent variable the country-level (normalized) gross revenues ("Turnover or gross premiums written - million euros"). The yearly absolute value of the dependent variable in the treated sector is divided by its value in 2000, the year prior to the introduction of the income tax break in Romania. From these resulting yearly ratios we subtract the corresponding ratios for the comparison sectors. The treated sectors are those that use K72 ("Computer and related activities," including "Software consultancy and supply" and "Publishing of software") services at high-intensity. We exclude K72 itself from this category. Sectors that have a low-intensity of use of K72 services serve as comparison sectors. The data source for the dependent variable is Eurostat, Structural Business Statistics, Annual detailed enterprise statistics on services (NACE Rev 1.1). Data for the predictors comes from the World Bank, World Development Indicators. We use as predictors the "GDP per capita (constant LCU)," "Medium and high-tech sector (% manufacturing value added)" and "Services, etc., value added (% of GDP)." All figures are an output of the **synth_runner** package for Stata [Quistorff and Galiani, 2017], with the *nested* option specified.

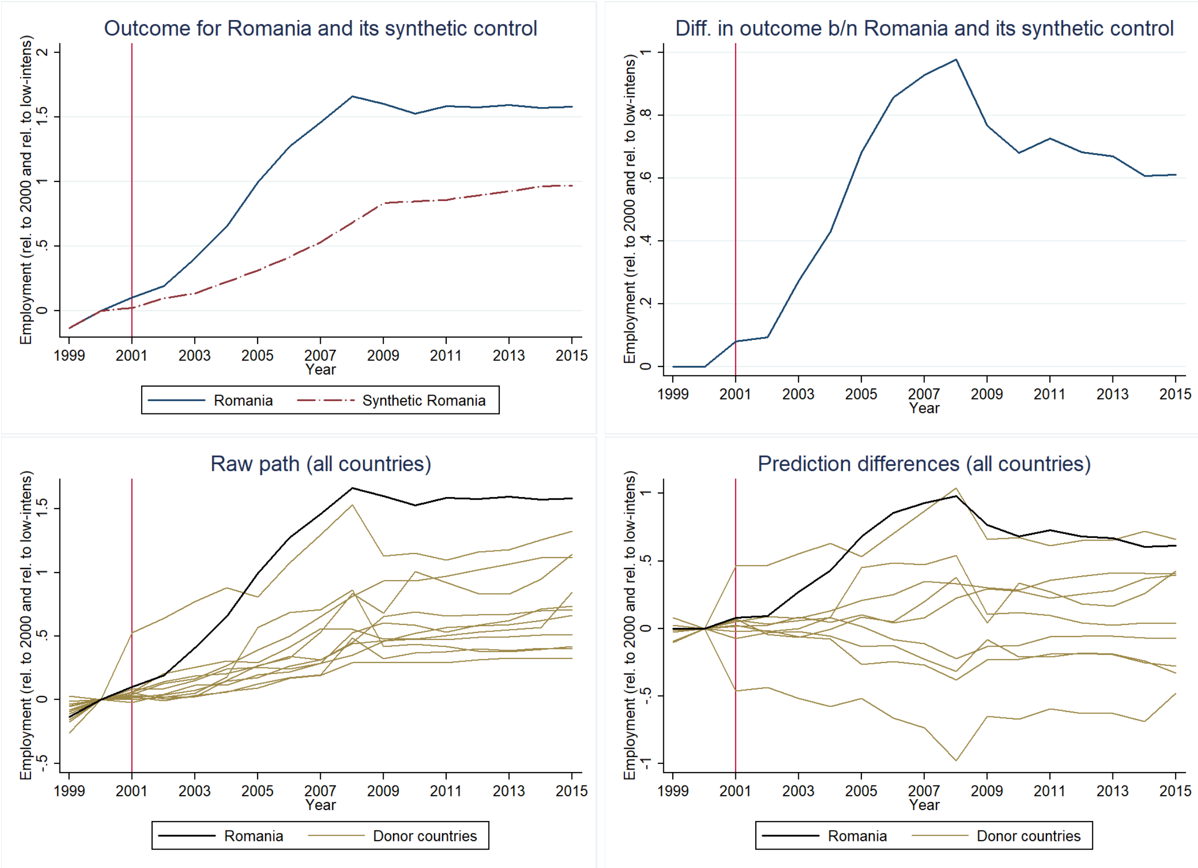

Figure 6: IT-Using Sectors Vs. Non-IT Using Sectors. SCM with Outcome Variable: "Employees - Number" (Normalized)

Notes: In this exercise we use the synthetic control method introduced in Section 4.2 to study the sector-level downstream effects of the introduction of the 2001 law granting an income tax break to workers in IT. All figures have as dependent variable the country-level (normalized) "Employees - number." The yearly absolute value of the dependent variable in the treated sector is divided by its value in 2000, the year prior to the introduction of the income tax break in Romania. From these resulting yearly ratios we subtract the corresponding ratios for the comparison sectors. The treated sectors are those that use K72 ("Computer and related activities," including "Software consultancy and supply" and "Publishing of software") services at high-intensity. We exclude K72 itself from this category. Sectors that have a low-intensity of use of K72 services serve as comparison sectors. The data source for the dependent variable is Eurostat, Structural Business Statistics, Annual detailed enterprise statistics on services (NACE Rev 1.1). Data for the predictors comes from the World Bank, World Development Indicators. We use as predictors the "GDP per capita (constant LCU)," "Medium and high-tech sector (% manufacturing value added)" and "Services, etc., value added (% of GDP)." All figures are an output of the **synth_runner** package for Stata [Quistorff and Galiani, 2017], with the *nested* option specified.

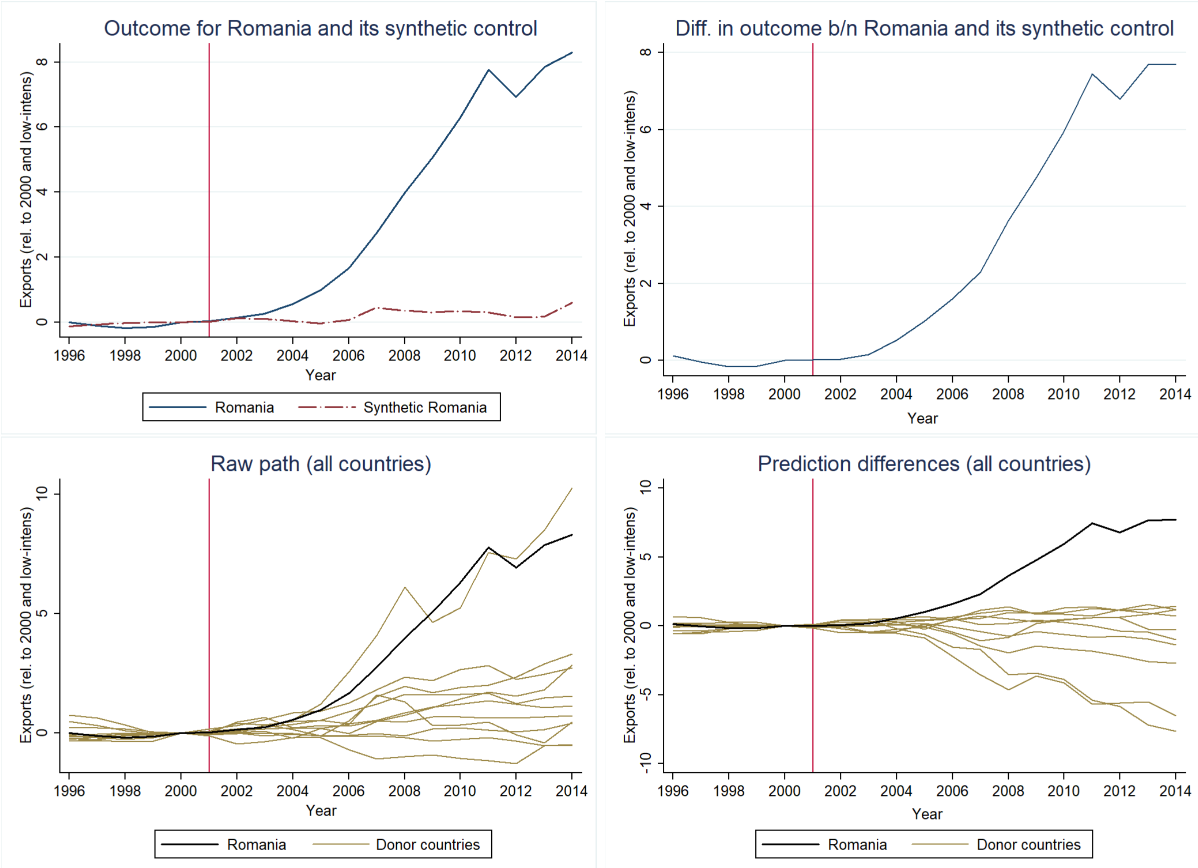

Figure 7: IT-Using Sectors Vs. Non-IT Using Sectors. SCM with Outcome Variable: "Goods Export Value" (Normalized)

Notes: In this exercise we use the synthetic control method introduced in Section 4.2 to study the sector-level downstream effects of the introduction of the 2001 law granting an income tax break to workers in IT. All figures have as dependent variable the country-level (normalized) "(Goods Export) Trade Value (US$)." The yearly absolute value of the dependent variable in the treated sector is divided by its value in 2000, the year prior to the introduction of the income tax break in Romania. From these resulting yearly ratios we subtract the corresponding ratios for the comparison sectors. The treated sectors are those that use K72 ("Computer and related activities," including "Software consultancy and supply" and "Publishing of software") services at high-intensity. K72 itself is excluded from this category. Sectors that have a low-intensity of use of K72 services serve as comparison sectors. The data source for the dependent variable is UN Comtrade, Goods Exports. Data for the predictors comes from the World Bank, World Development Indicators. We use as predictors the "GDP per capita (constant LCU)," "Medium and high-tech sector (% manufacturing value added)" and "Services, etc., value added (% of GDP)." All figures are an output of the **synth_runner** package for Stata [Quistorff and Galiani, 2017], with the *nested* option specified.

Tables

Table 1: Difference-in-Difference Around 2001 Income Tax Break: Main Results for the Baseline and Alternative Control Groups

	Baseline sample			Alternative control group 1			Alternative control group 2		
Dep. var.	Revenues (1)	Workers (2)	Assets (3)	Revenues (4)	Workers (5)	Assets (6)	Revenues (7)	Workers (8)	Assets (9)
Yearly effects									
$\beta_{DiD,1997}$	-0.024	0.013	-0.041	-0.026	-0.005	-0.005	-0.043	0.014	0.005
	(0.071)	(0.044)	(0.066)	(0.101)	(0.063)	(0.091)	(0.083)	(0.050)	(0.077)
$\beta_{DiD,1998}$	-0.066	-0.033	-0.062	-0.124	-0.064	-0.038	-0.075	-0.023	-0.026
	(0.065)	(0.039)	(0.058)	(0.089)	(0.053)	(0.081)	(0.074)	(0.044)	(0.068)
$\beta_{DiD,1999}$	-0.059	0.002	-0.054	-0.026	0.010	0.053	-0.066	0.014	-0.029
	(0.055)	(0.029)	(0.048)	(0.074)	(0.040)	(0.066)	(0.062)	(0.033)	(0.057)
$\beta_{DiD,2001}$	0.202***	0.100***	0.101**	0.198***	0.109***	0.114**	0.250***	0.122***	0.166***
	(0.044)	(0.024)	(0.040)	(0.059)	(0.032)	(0.053)	(0.051)	(0.028)	(0.048)
$\beta_{DiD,2002}$	0.194***	0.066**	0.114**	0.218***	0.088**	0.197***	0.252***	0.098***	0.183***
	(0.050)	(0.030)	(0.046)	(0.066)	(0.039)	(0.061)	(0.058)	(0.035)	(0.055)
$\beta_{DiD,2003}$	0.132***	0.092***	0.102**	0.208***	0.099**	0.230***	0.143**	0.144***	0.142**
	(0.051)	(0.033)	(0.048)	(0.069)	(0.042)	(0.064)	(0.059)	(0.038)	(0.057)
$\beta_{DiD,2004}$	0.193***	0.114***	0.120**	0.298***	0.136***	0.266***	0.191***	0.163***	0.147**
	(0.053)	(0.035)	(0.051)	(0.074)	(0.046)	(0.069)	(0.062)	(0.040)	(0.060)
$\beta_{DiD,2005}$	0.240***	0.126***	0.117**	0.301***	0.140***	0.259***	0.229***	0.183***	0.135**
	(0.056)	(0.036)	(0.054)	(0.077)	(0.048)	(0.072)	(0.066)	(0.043)	(0.063)
Adjusted R^2	0.819	0.822	0.833	0.832	0.849	0.849	0.813	0.836	0.830
# Observations	25,518	25,518	25,518	12,203	12,203	12,203	15,931	15,931	15,931
# Firms	4,336	4,336	4,336	2,120	2,120	2,120	2,696	2,696	2,696

Notes: In this exercise we study the firm-level impact of the introduction of the 2001 law granting an income tax break to workers in IT. All regressions in this table use Amadeus data. In this table we report the point estimates of the difference-in-difference coefficients of interest from the Equation (1) (upper panel) and from the pooled version of Equation (1) (lower panel), i.e., $\beta_{DiD,t}$ and β_{DiD} respectively. The outcome variables used are log operating revenues, log number of workers, and log total assets. The year prior to the introduction of the law (2000) is the reference year. Treated firms are those in the NACE Rev 1 sector 722 (Software consultancy and supply). Firms in the baseline control group – columns (1)-(3) – are in high-tech knowledge-intensive service sectors (as classified by the Eurostat). Firms in the alternative control group 1 – columns (4)-(6) – are either part of untreated IT service sectors or R&D sectors. Firms in the alternative control group 2 – columns (7)-(9) – are part of high-tech knowledge-intensive service sectors (as classified by the Eurostat) *excluding* ICT service sectors. All specifications include firm and calendar year fixed effects. The number of observations is the same for the yearly and pooled versions of the same regression. Robust standard errors, clustered at the firm-level, in parenthesis. ***,**,* denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 2: Difference-in-Differences Design Around 2013 Reform: "First Stage" Effects on the Share of Income Tax Exempted Employees

Dep. var.	1 if more than 20% exempted employees				Share of exempted employees			
Sample	Baseline (1)	Alternative 1 (2)	Alternative 2 (3)	Eligible (4)	Baseline (5)	Alternative 1 (6)	Alternative 2 (7)	Eligible (8)
Elig. sect. × d_{2011}	-0.002 (0.006)	-0.002 (0.006)	-0.002 (0.006)		-0.003 (0.003)	-0.002 (0.003)	-0.003 (0.003)	
Elig. sect. × d_{2013}	0.023*** (0.005)	0.024*** (0.005)	0.023*** (0.005)		0.012*** (0.003)	0.012*** (0.003)	0.012*** (0.003)	
Elig. sect. × d_{2014}	0.050*** (0.006)	0.050*** (0.006)	0.049*** (0.006)		0.028*** (0.004)	0.028*** (0.004)	0.027*** (0.004)	
Elig. sect. × d_{2015}	0.071*** (0.007)	0.072*** (0.007)	0.071*** (0.007)		0.038*** (0.005)	0.039*** (0.005)	0.038*** (0.005)	
d_{2011}				-0.002 (0.006)				-0.002 (0.003)
d_{2013}				0.024*** (0.005)				0.012*** (0.003)
d_{2014}				0.049*** (0.006)				0.028*** (0.004)
d_{2015}				0.071*** (0.007)				0.038*** (0.005)
Adjusted R^2	0.174	0.166	0.102	0.004	0.152	0.145	0.087	0.003
# Observations	39,052	37,349	28,269	21,601	39,052	37,349	28,269	21,601
# Firms	8,990	8,551	6,547	4,986	8,990	8,551	6,547	4,986

Notes: In this exercise, we study the "first stage" effects of the 2013 income tax reform on the firm-level share of workforce exemption from the income tax. All regressions in this table use administrative data. Columns (1)-(4) use as the dependent variable a dummy variable that takes value 1 for firm i in year t if firm i has more than 20% of its workers exempted from the income tax in year t. Columns (5)-(8) use as the dependent variable the share of workers of firm i who are exempted from the income tax in year t. The year prior to the amendment of 2013 (2012) is the reference year. Firms in the baseline control group – columns (1) and (5) – are in high-tech knowledge-intensive service sectors (as classified by the Eurostat). Firms in the alternative control group 1 – columns (2) and (6) – are in ICT service sectors (as classified by the OECD). Firms in the alternative control group 2 – columns (3) and (7) – are part of high-tech knowledge-intensive service sectors (as classified by the Eurostat) *excluding* ineligible ICT service sectors (as classified by the OECD). Firms in alternative control 3 are only part of eligible sectors. The samples used in this table are different from the main samples in Table 3 and Table 4 in that the former keep *all* firms in HTKI (high-tech knowledge-intensive) service sectors, in ICT service sectors, in HTKI service sectors *excluding* the ICT service sectors, and in eligible sectors, whereas the latter bring additional restrictions on the firm-level share of workforce exemption before and after 2013. In addition to interaction terms between a dummy taking value 1 if the sector of firm i is eligible for the income tax break of its eligible workers (Eligible sector) and a year dummy (d_t), columns (1), (2), (3), (5), (6) and (7) include firm fixed effects and year fixed effects (d_t, whose estimates are omitted from the table). In addition to the year fixed effects (d_t) whose estimates are reported in the table, columns (4) and (8) also include firm fixed effects. ***,**,* denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3: Difference-in-Differences Design Around 2013 Reform: Main Results for the Baseline and Alternative Control Groups (1/2)

	Administrative, baseline sample			Amadeus, baseline sample			Administrative, alternative control 1			Amadeus, alternative control 1		
	Revenues	Workers	Assets	Revenues	Workers	Assets	Revenues	Workers	Assets	Revenues	Workers	Assets
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\beta_{DiD,2008}$				-0.092	-0.059	0.026				-0.098	-0.061	0.023
				(0.074)	(0.046)	(0.079)				(0.074)	(0.046)	(0.080)
$\beta_{DiD,2009}$				-0.003	-0.038	-0.000				-0.006	-0.039	-0.004
				(0.063)	(0.039)	(0.065)				(0.063)	(0.039)	(0.065)
$\beta_{DiD,2010}$				-0.028	0.003	0.035				-0.029	0.004	0.033
				(0.060)	(0.034)	(0.055)				(0.060)	(0.034)	(0.055)
$\beta_{DiD,2011}$	-0.054	0.031	-0.005	-0.057	0.009	0.012	-0.065	0.031	-0.008	-0.070	0.009	0.008
	(0.043)	(0.019)	(0.046)	(0.051)	(0.026)	(0.047)	(0.043)	(0.019)	(0.046)	(0.052)	(0.026)	(0.047)
$\beta_{DiD,2013}$	0.151***	0.135***	0.067*	0.147***	0.110***	0.098***	0.147***	0.136***	0.064*	0.143***	0.111***	0.094***
	(0.049)	(0.024)	(0.037)	(0.038)	(0.023)	(0.036)	(0.049)	(0.024)	(0.037)	(0.039)	(0.023)	(0.036)
$\beta_{DiD,2014}$	0.235***	0.137***	0.148***	0.215***	0.110***	0.143***	0.232***	0.138***	0.143***	0.212***	0.111***	0.137***
	(0.054)	(0.031)	(0.045)	(0.050)	(0.030)	(0.045)	(0.054)	(0.031)	(0.045)	(0.050)	(0.030)	(0.045)
$\beta_{DiD,2015}$	0.203***	0.102***	0.170***	0.172***	0.080**	0.190***	0.203***	0.105***	0.168***	0.173***	0.082**	0.188***
	(0.067)	(0.037)	(0.055)	(0.064)	(0.038)	(0.058)	(0.067)	(0.037)	(0.055)	(0.064)	(0.038)	(0.058)
Adjusted R^2	0.848	0.910	0.905	0.792	0.855	0.849	0.851	0.899	0.896	0.798	0.847	0.843
# Observations	27,491	27,491	27,491	32,715	32,715	32,715	25,916	25,916	25,916	31,651	31,651	31,651
# Firms	6,146	6,146	6,146	4,837	4,837	4,837	5,769	5,769	5,769	4,676	4,676	4,676

Notes: In this exercise we study the firm-level impact of the 2013 expansion to the income tax break law of 2001. We present both the baseline results on the administrative in columns (1)-(3) and their robustness to alternative data sources and control groups. The outcome variables used are log revenues, log number of workers, and log total assets. The difference between columns (1)-(3) and columns (4)-(6) (between columns (7)-(9) and (10)-(12)) stems from the source of the data over which we estimate the model in Equation (2). Columns (1)-(3) and (7)-(9) use the administrative data described in Section 3.2. Columns (4)-(6) and (10)-(12) use data from Amadeus, a Bureau van Dijk product. Firms in the baseline control group – columns (1)-(6) – are part of high-tech knowledge-intensive service sectors (as classified by the Eurostat). Firms in the alternative control group 1 – columns (7)-(12) – are in ICT service sectors (as classified by the OECD). Robust standard errors, clustered at the firm-level, in parenthesis. All specifications include firm fixed effects. ***,**,* denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4: Difference-in-Differences Design Around 2013 Reform: Robustness of the Results to Alternative Control Groups (2/2)

	Administrative, alternative control 2			Amadeus, alternative control 2			Administrative, alternative control 3			Amadeus, alternative control 3		
	Revenues	Workers	Assets	Revenues	Workers	Assets	Revenues	Workers	Assets	Revenues	Workers	Assets
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\beta_{DiD,2008}$				-0.084	-0.045	0.028				-0.087	-0.046	0.025
				(0.077)	(0.046)	(0.083)				(0.077)	(0.046)	(0.083)
$\beta_{DiD,2009}$				-0.004	-0.030	-0.009				-0.006	-0.031	-0.013
				(0.065)	(0.039)	(0.068)				(0.066)	(0.039)	(0.068)
$\beta_{DiD,2010}$				-0.034	0.011	0.008				-0.032	0.013	0.007
				(0.062)	(0.034)	(0.056)				(0.062)	(0.034)	(0.056)
$\beta_{DiD,2011}$	-0.051	0.025	-0.019	-0.074	0.003	-0.009	-0.055	0.026	-0.023	-0.079	0.003	-0.014
	(0.044)	(0.020)	(0.048)	(0.053)	(0.026)	(0.049)	(0.044)	(0.020)	(0.048)	(0.053)	(0.026)	(0.049)
$\beta_{DiD,2013}$	0.173***	0.137***	0.069*	0.151***	0.115***	0.097***	0.170***	0.139***	0.065*	0.149***	0.118***	0.093**
	(0.047)	(0.024)	(0.037)	(0.039)	(0.024)	(0.037)	(0.048)	(0.024)	(0.037)	(0.040)	(0.024)	(0.037)
$\beta_{DiD,2014}$	0.242***	0.134***	0.147***	0.213***	0.114***	0.141***	0.239***	0.136***	0.143***	0.212***	0.116***	0.136***
	(0.056)	(0.031)	(0.046)	(0.051)	(0.030)	(0.046)	(0.056)	(0.032)	(0.046)	(0.052)	(0.031)	(0.046)
$\beta_{DiD,2015}$	0.203***	0.104***	0.166***	0.170***	0.090**	0.192***	0.204***	0.106***	0.164***	0.171***	0.092**	0.189***
	(0.068)	(0.037)	(0.056)	(0.066)	(0.039)	(0.059)	(0.068)	(0.037)	(0.056)	(0.066)	(0.039)	(0.059)
Adjusted R^2	0.835	0.919	0.904	0.787	0.868	0.847	0.822	0.902	0.875	0.779	0.859	0.817
# Observations	17,238	17,238	17,238	20,550	20,550	20,550	10,934	10,934	10,934	13,227	13,227	13,227
# Firms	3,882	3,882	3,882	3,050	3,050	3,050	2,473	2,473	2,473	1,994	1,994	1,994

Notes: In this exercise we check the robustness of the baseline results from Table 3 on the firm-level impact of the 2013 amendment to the income tax break law of 2001. The outcome variables used are log revenues, log number of workers, and log total assets. The difference between columns (1)-(3) and columns (4)-(6) (between columns (7)-(9) and (10)-(12)) stems from the source of the data over which we estimate the model in Equation (2). Columns (1)-(3) and (7)-(9) use the administrative data described in Section 3.2. Columns (4)-(6) and (10)-(12) use data from Amadeus, a Bureau van Dijk product. Firms in the alternative control group 2 – columns (1)-(6) – are part of high-tech knowledge-intensive service sectors (as classified by the Eurostat) *excluding* ineligible ICT service sectors. Firms in the alternative control group 3 – columns (7)-(12) – belong only to eligible sectors. Thus the comparison is between firms under the 5% threshold of exempted employees with firms that jump over the 20% threshold after 2013 (all in eligible sectors). Robust standard errors, clustered at the firm-level, in parenthesis. All specifications include firm fixed effects. ***,**,* denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5: Difference-in-Differences Design Around 2013 Reform: Heterogeneity of the Baseline Results

	Revenues (1)	Workers (2)	Assets (3)	Revenues (4)	Workers (5)	Assets (6)
Panel A: Size	*Micro: < 10 workers*			*Small, medium, large: ≥ 10 workers*		
β_{DiD}	0.213***	0.113***	0.128***	0.251***	0.110	0.129
	(0.055)	(0.028)	(0.046)	(0.085)	(0.082)	(0.082)
Adjusted R^2	0.753	0.772	0.850	0.893	0.878	0.949
# Observations	23,903	23,903	23,903	3,588	3,588	3,588
# Firms	5,384	5,384	5,384	762	762	762
Panel B: Age	*Young: < 5 years old*			*Old: ≥ 5 years old*		
β_{DiD}	0.271***	0.149***	0.163**	0.141***	0.072**	0.065
	(0.088)	(0.045)	(0.068)	(0.053)	(0.032)	(0.050)
Adjusted R^2	0.758	0.836	0.828	0.890	0.934	0.937
# Observations	10,244	10,244	10,244	17,247	17,247	17,247
# Firms	2,358	2,358	2,358	3,788	3,788	3,788

Notes: In this exercise we explore the heterogeneity of the baseline effects of the 2013 expansion to the income tax break law of 2001. The outcome variables used are log revenues, log number of workers, and log total assets. For brevity, we implement the pooled version of the DiD Equation (2) on the baseline sample of 6,146 firms (see Table 3, columns (1)-(3)). Firms in the baseline control group are in high-tech knowledge-intensive service sectors (as classified by the Eurostat). The sample is split in two parts based on the number of workers or age of the firm (both in 2011). The DiD regressions are run separately on each part of the baseline sample. All specifications include firm and sector-by-year fixed effects. Robust standard errors, clustered at the firm-level, in parenthesis. ***,**,* denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 6: Back-of-the-Envelope Calculation of Foregone Tax Revenues in 2015

Scenario	(1)	(2)	(3)
Equivalent Subsidy (Mil. Euro)	92.63	58.46	56.10
% Value Added	4.45%	2.81%	2.69%
% Revenues	2.23%	1.41%	1.35%
% Production	2.34%	1.47%	1.41%
% Wage Bill	7.40%	4.67%	4.48%

Notes: This table reports back-of-the-envelope values of the equivalent subsidy extended to the IT sector (i.e., the foregone tax revenues from the income tax break to exempted workers). In all three scenarios the workers exempted in 2015 are assumed to maintain full employment were the exemption to be removed at the beginning of that year. Scenario (1) is one in which the wages of exempted employees would be kept exactly the same. Scenario (2) is one in which exempted employees would be paid the average wage in ineligible high-tech knowledge-intensive sectors. Scenario (3) is one in which exempted employees would be paid the average wage in ineligible ICT service sectors. Wages come from the administrative data used in the baseline analysis, whereas the value-added, revenues, production, and wage bill of eligible sectors come from the Structural Business Statistics of Eurostat.

Appendix A   Context on Policies Relevant to the IT Sector in Romania

Appendix A.1   Details on Romania's Income Tax Break for Workers in IT

Table A1: 2001 Income Tax Break and Its Subsequent Amendments: Eligibility Criteria for the Income Tax Break (1/2)

Order Number	4079/ 268/ 1480/ 2001	661/ 444/ 2196/ 2001	250/ 189/ 748/ 2004	539/ 225/ 1479/ 2013	217/ 4172/ 1348/ 835/ 2015	872/ 5932/ 2284/ 2903/ 2016
Effective date	08/05/2001	11/22/2001	06/29/2004	09/10/2013	07/30/2015	01/09/2017
NACE Code	7220 (Rev 1)	7220 (Rev 1)	**7221, 7222 (Rev 1.1)**	**5821, 5829, 6201, 6202, 6209 (Rev 2)**	5821, 5829, 6201, 6202, 6209 (Rev 2)	5821, 5829, 6201, 6202, 6209 (Rev 2)
Occupations	Analyst,	Analyst,	Analyst,	Analyst,	Analyst,	Analyst,
	Programmer,	Programmer,	Programmer,	Programmer,	Programmer,	Programmer,
	Computer Systems Designer, Engineer or Programmer of IT Systems, Database Manager,	Computer Systems Designer, Engineer or Programmer of IT Systems, Database Manager,	Computer Systems Designer, Engineer or Programmer of IT Systems, Database Manager,	Computer Systems Designer, Engineer or Programmer of IT Systems, Database Manager,	Computer Systems Designer, Engineer or Programmer of IT Systems, Database Manager,	Computer Systems Designer, Engineer or Programmer of IT Systems, Database Manager,
	Software Engineer and Manager of IT Projects	Software Engineer and Manager of IT Projects	Software Engineer and Manager of IT Projects	Software Engineer and Manager of IT Projects	Software Engineer and Manager of IT Projects	Software Engineer and Manager of IT Projects
					Computer System Programmer	**Computer System Programmer**
Unit within the Firm	Unit Specialized in IT	Unit Specialized in IT	Unit Specialized in IT	Unit Specialized in IT	Unit Specialized in IT	Unit Specialized in IT

Extra Eligibility Criteria for the Income Tax Break, in Addition to Those from Table A1 (2/2)

Order Number: 4079/ 268/ 1480/ 2001

Eligible Major during Higher Education for Exempted Worker: Automation, Computers, Computer Science, Cybernetics, Mathematics, Electronics.

Minimum Annual Revenue from Software Development: Annual income of at least 10,000 U.S. dollars per employee benefiting from the income tax break.

Balance for Software Development Income: A balance is required, in which the income from software development needs to be explicitly reported.

Order Number: 661/444/2196/2001

Eligible Major during Higher Education for Exempted Worker: Automation and Industrial Computer Science; Computers, Electrical Engineering and Computers; Electronics; Applied Electronics, Electronics and Telecommunications, Communications; Mathematics, Mathematical Computer Science; Computer Science, Computer Science and Economics, Applied Computer Science, Cybernetics and Computer Science and Economics, Cybernetics and Economic Prediction, Accounting and Computer Science and Management.

Minimum Annual Revenue from Software Development: Annual income, in preceding year, of at least 10,000 U.S. dollars per employee benefiting in a given year from the income tax break.

Balance for Software Development Income: A balance is required, in which the income from software development needs to be explicitly reported.

Order Number: 250/189/748/2004

Eligible Major during Higher Education for Exempted Worker: Automation and Industrial Computer Science; Computers, Electrical Engineering and Computers; Electronics; Applied Electronics, Electronics and Telecommunications, Communications; Mathematics, Mathematical Computer Science; Computer Science, Computer Science and Economics, Applied Computer Science, Cybernetics and Computer Science and Economics, Cybernetics and Economic Prediction, Accounting and Computer Science and Management.

Minimum Annual Revenue from Software Development: Annual income, in preceding year, of at least 10,000 U.S. dollars per employee benefiting in a given year from the income tax break.

Balance for Software Development Income: A balance is required, in which the income from software development needs to be explicitly reported.

Order Number: 539/225/1479/2013

Eligible Major during Higher Education for Exempted Worker: Automation and Industrial Computer Science; Computers, Electrical Engineering and Computers; Electronics; Applied Electronics, Electronics and Telecommunications, Communications; Mathematics, Mathematical Computer Science; Computer Science, Computer Science and Economics, Applied Computer Science, Cybernetics and Computer Science and Economics, Cybernetics and Economic Prediction, Accounting and Computer Science and Management. **Newly eligible majors:** Industrial Computer Science, Applied Computer Science in Electrical Engineering, Applied Computer Science in Material Engineering, Mathematics and Applied Computer Science in Engineering; Cybernetics and Economics; Physics and Computer Sci-

ence; Chemistry and Computer Science; Automation and Applied Computer Science, Equipment for Modeling, Simulation and Computerized Warfare, Engineering of Multimedia Systems; Technologies and Telecommunication Systems, Remote Controls and Electronics in Transportation; Transmissions and Military Electronic Equipment.
Minimum Annual Revenue from Software Development: Annual income, in preceding year, of at least 10,000 U.S. dollars per employee benefiting in a given year from the income tax break.
Balance for Software Development Income: A balance is required, in which the income from software development needs to be explicitly reported.

Order Number: 217/4172/1348/835/2015 and 872/5932/2284/2903/2016
Eligible Major during Higher Education for Exempted Worker: A diploma issued after a form of higher education, irrespective of major.
Minimum Annual Revenue from Software Development: Annual income, in preceding year, of at least 10,000 U.S. dollars per employee benefiting in a given year from the income tax break.
Balance for Software Development Income: A balance is required, in which the income from software development needs to be explicitly reported. **New:** New firms or firms undergoing a restructuring during that fiscal year are exempted from this requirement.

Appendix A.2 Other Policies Relevant for the IT sector

Appendix A.2.1 State Aid Program

Between 2011 and 2016, Romania introduced several state aid programs supporting job creation.[37] The most relevant for the sectors studied was the program created by Government Decision 797/2012. It supported large investments in new technologies with an IT component and job creation of at least of 200 new jobs. While firms in most manufacturing, energy and service sectors were eligible, mainly firms in high-tech knowledge-intensive (HTKI) sectors benefited from it. We drop from our sample of analysis firms that have benefited from this State Aid. The program created by Government decision 332/2014 (meant to support large investments, job creation, and regional development) also benefited several firms in HTKI sectors. Firms in HTKI sectors were also eligible for programs supporting SMEs and start-ups, such as “Start-up Nation,” but these programs were smaller and less likely to affect major investments or job creation (e.g., “Start-up Nation” had an upper limit of approximately 44,000 euros).

Appendix A.2.2 Other Tax Exemptions

Programmers are not the only category of workers exempted from the personal income tax in Romania. Two other categories of workers exempted from the income tax could be employed by IT firms, without being programmers: workers with serious disabilities[38] and since 2016 workers in research and development (defined broadly, with no requirement to work in software development).[39] For companies with at least 50 employees, it was compulsory that at least 4% of their workers have disabilities. When

[37] Government decisions 797/2012, 322/2014 and 807/2014.
[38] See Law 448/2006 for details.
[39] See Order 4947/899/2018/1840/906/2016 published in September 2016 for details.

firms could not comply, they had to pay a given amount to support the inclusion of people with disabilities. Despite these other exemptions, the vast majority of exempted employees in eligible sectors were exempted due to the tax break for workers in IT. In October 2017,[40] workers benefiting from the law under study in this article represented 96% of the exempted employees in eligible sectors.

Table A2: Correspondence Table between NACE sector Codes Rev 2 and Rev 1.1

NACE Rev 2	Description Rev 2	NACE Rev 1.1	Description Rev 1.1	Comments
5821	Publishing of computer games	7221	Publishing of software	Publishing of computer games
5821	Publishing of computer games	724	Database activities	On-line computer games publishing
5829	Other software publishing	7221	Publishing of software	All software publishing, except computer games publishing
5829	Other software publishing	724	Database activities	All on-line software publishing, except computer games on-line publishing
6201	Computer programming activities	7221	Publishing of software	Software programming
6201	Computer programming activities	7222	Other software consultancy and supply	Includes: Analysis, design and programming of systems ready to use: - development, production, supply and documentation of made-to-order software based on orders from specific users - writing of programs following directives of the user - web page design
6201	Computer programming activities	724	Database activities	Designing of structure and content of database
6202	Computer consultancy activities	721	Hardware consultancy	All
6202	Computer consultancy activities	7222	Other software consultancy and supply	Analysis, design and programming of systems ready to use: - analysis of the user's needs and problems, consultancy on the best solution
6209	Other information technology and computer service activities	3002	Manufacture of computers and other information processing equipment	Installation of personal computers and peripheral equipment
6209	Other information technology and computer service activities	7222	Other software consultancy and supply	Software installation services
6209	Other information technology and computer service activities	726	Other computer related activities	NACE 1.1 class 72.60 was an ”empty class”.

Notes: Source Eurostat.

[40] In 2017 we can observe the reason of the income tax exemption, which we cannot observe in previous years.

Appendix B Additional Firm-Level Evidence

Appendix B.1 Descriptive Statistics with Firm-Level Data

Appendix B.1.1 Amadeus Data: 1997 - 2005

Table B1: Descriptive Statistics for Firms in Treated and Comparison Sectors in Year 2000

Variable	# Firms	Mean	Median	SD
Eligible sector				
Revenues	825	128.73	27	321.65
Workers	825	7.28	3	12.10
Assets	825	67.64	12	191.91
Baseline comparison sectors				
Revenues	2,117	133.61	27	369.44
Workers	2,117	7.07	3	12.89
Assets	2,117	64.10	10	188.12

Notes: Table B1 reports descriptive statistics for the baseline sample used to study the firm-level impact of the introduction in 2001 of the income tax break for workers in IT (see results in Table 1). The data source is Amadeus. These statistics pertain to the year 2000, the year prior to the introduction of the policy of interest. The upper panel reports summary statistics for the firms in the eligible sector (722, NACE Rev 1) and for the three outcome variables of interest (operating revenues, number of workers, and total assets). The lower panel reports summary statistics for the firms that belong to the baseline set of comparison sectors (i.e., high-tech knowledge-intensive service sectors, as classified by Eurostat). The unit of measure for the operating revenues and assets variables is thousands of euros.

Appendix B.1.2 Administrative Data: 2011 - 2015

Table B2: Descriptive Statistics of Firms in Year 2011, by Firm Category

Sector eligible Exempted after 2013	Not eligible Not exempted		Eligible Not exempted		Eligible Exempted	
	Mean	SD	Mean	SD	Mean	SD
Revenues	1,694.33	27,974.28	193.85	1,083.31	336.28	2,171.49
Workers	22.01	218.63	3.84	13.73	7.22	25.36
Assets	2,312.87	34,902.15	231.64	2,019.81	210.65	1,621.17
Micro	0,83		0,94		0,89	
Small	0,13		0,05		0,08	
Medium	0,04		0,01		0,02	
Large	0,01		0,00		0,00	
Young	0,34		0,35		0,41	

Notes: This table reports descriptive statistics for 2011 (two years before the expansion of the income tax break law), for the three types of firms that we include in our baseline sample, and for the three outcome variables of interest (revenues, number of workers, and total assets). The first sample includes firms in sectors comparable to the IT sector. These comparable sectors were, however, not targeted by the income tax break law. The second and third samples contain firms in the eligible sectors. The second sample contains firms with less than 5% of employees exempted from the income tax in each year of our 2011 to 2015 sample. The third sample contains firms that had less than 5% of exempted employees in 2011 and 2012, but jumped to over 20% of exempted employees after 2013. The unit of measure for the revenues and assets variables is thousands of euros.

Table B3: Percentage of Firms with Exempted Employees in Different Samples

Percentage of firms with at least 1 exempted employee, raw dataset

Year	HTKI	ICT	Eligible
2011	18%	19%	35%
2012	18%	19%	34%
2013	20%	21%	37%
2014	22%	23%	40%
2015	23%	24%	42%

Percentage of firms with at least 1 exempted employee, clean dataset used as main sample

Year	HTKI	ICT	Eligible
2011	20%	21%	36%
2012	19%	20%	35%
2013	21%	22%	39%
2014	23%	24%	42%
2015	25%	26%	45%

Percentage of firms with 20% or more exempted employees, clean dataset used as main sample

Year	HTKI	ICT	Eligible
2011	16%	16%	28%
2012	16%	17%	29%
2013	17%	18%	31%
2014	19%	19%	34%
2015	20%	21%	36%

Notes: Table B3 reports the share of firms in a given sample that have at least 1 exempted employee or that have more than 20% of their employees exempted from the income tax for each year between 2011 and 2015. By construction, all firms in ineligible sectors have less than 5% of their employees exempted from the income tax (employees who are likely to have a disability, another criterion on which this tax break can be granted). The sample of firms in "HTKI" services (HTKI stands for "high-tech knowledge-intensive") from the lower two panels ("cleaned dataset used as main sample") is the baseline sample whose results are reported in Table 3, columns (1)-(3). The sample of firms in "ICT" services sectors from the lower two panels ("cleaned dataset used as main sample") is the robustness check sample whose results are reported under the "alternative control group 1" header in Table 3, columns (4)-(6). The sample of firms in "eligible" sectors from the two lower panels ("cleaned dataset used as main sample") contains only firms in sectors whose IT workers are eligible to the income tax break and is the robustness check sample whose results are reported under the "alternative control group 3" header in Table 4 (Appendix B.2).

Table B4: Predictors of Firm-level Exemption Performance

	1 if more than 20% exempted workers			Share of exempted workers		
	OLS	Probit	Logit	OLS	Probit	Logit
	(1)	(2)	(3)	(4)	(5)	(6)
Foreign-owned	0.105***	0.101***	0.100***	0.068***	0.063***	0.062***
	(0.022)	(0.022)	(0.021)	(0.015)	(0.014)	(0.014)
Small	0.085***	0.083***	0.083***	0.030***	0.029***	0.030***
	(0.017)	(0.017)	(0.017)	(0.011)	(0.010)	(0.010)
Medium	0.218***	0.215***	0.213***	0.098***	0.090***	0.089***
	(0.034)	(0.035)	(0.035)	(0.021)	(0.020)	(0.020)
Large	0.242***	0.239***	0.238***	0.129***	0.119***	0.117***
	(0.075)	(0.079)	(0.079)	(0.044)	(0.042)	(0.040)
Young	-0.004	-0.003	-0.003	0.017**	0.018**	0.018**
	(0.011)	(0.011)	(0.011)	(0.009)	(0.009)	(0.009)
County % exempt.	0.573***	0.569***	0.573***	0.406***	0.403***	0.406***
	(0.062)	(0.061)	(0.061)	(0.044)	(0.044)	(0.043)
d_{2011}	-0.011*	-0.011*	-0.012*	-0.013***	-0.014***	-0.014***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
d_{2013}	0.020***	0.021***	0.021***	0.012***	0.013***	0.013***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
d_{2014}	0.030***	0.032***	0.032***	0.017***	0.018***	0.018***
	(0.007)	(0.007)	(0.007)	(0.005)	(0.005)	(0.005)
d_{2015}	0.047***	0.048***	0.048***	0.023***	0.024***	0.024***
	(0.008)	(0.008)	(0.008)	(0.005)	(0.006)	(0.006)
Adjusted/Pseudo R^2	0.042	0.033	0.033	0.030	0.019	0.019
# Observations	21,601	21,601	21,601	21,601	21,601	21,601
# Firms	4,986	4,986	4,986	4,986	4,986	4,986

Notes: Table B4 explores the predictors of the firm-level share of income tax exempted workers. We only focus on the sample of firms in eligible sectors, as only in those sectors workers can benefit from an income tax break based on their IT-related occupation. Columns (1) and (4) report the results of OLS models, columns (2) and (5) of probit (fractional response for the share of exempted workers) models, and columns (3) and (6) of logit (fractional response for the share of exempted workers) models. Columns (1)-(3) use as the dependent variable a dummy variable that takes value 1 for firm i in year t if firm i has more than 20% of its workers exempted from the income tax in year t. Columns (4)-(6) use as the dependent variable the share of workers of firm i who are exempted from the income tax in year t. The sample used in this table is different from the "alternative control 3" sample in Table 4 in that the former keeps *all* firms in eligible sectors, whereas the latter brings additional restrictions on the firm-level share of workforce exemption before and after 2013. The dependent variables and sample used in this table are the same as those used in Table 2, which explores the first stage effects of the 2013 reform on the firm-level share of workforce exemption from the income tax. The explanatory variables of size and age are contemporaneous. Data on foreign ownership in 2016 comes from Amadeus. We assume that if a firm is foreign-owned in 2016 it has been foreign-owned throughout 2011 to 2015. The reference category contains firms that are domestically-owned, have a micro size, and are older than five years. The reference year in 2012.

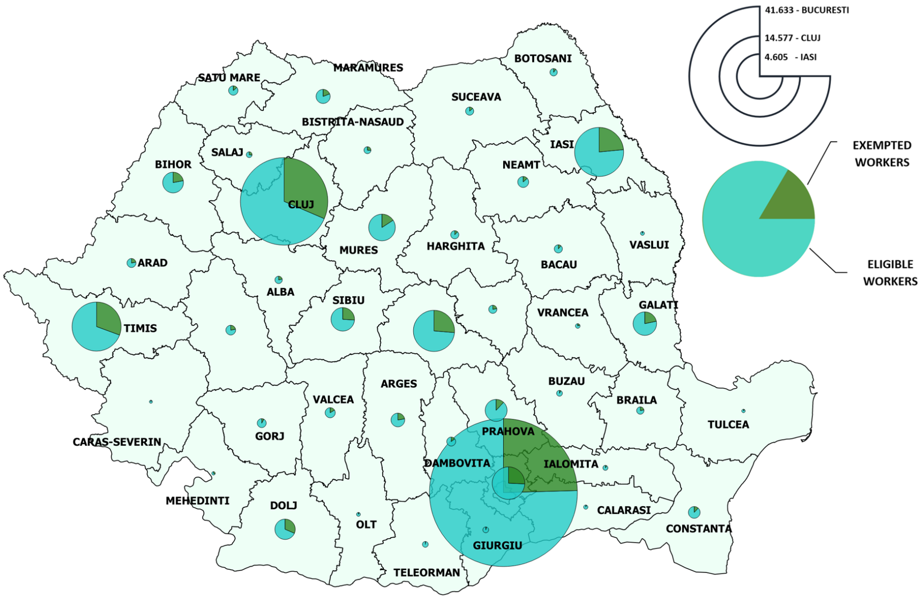

Figure B1: County-Level Number of Employees in the Eligible IT Sector in 2015 and the Share of these Employees Actually Exempted from the Income Tax

Notes: The legend reports the total number of employees in the eligible IT sectors in 2015 in the three counties with the largest number of such employees, i.e., București, Cluj, and Iași. This map uses the sample from the 2011 to 2015 administrative dataset kept for the econometric analysis, as described in Section 3.2.

Figure B1 shows the spatial distribution in 2015 of the number of workers in the sectors eligible for the income tax exemption and the share of these workers who are actually benefiting from this exemption. The unit of analysis is a county (*județ*). This map highlights the striking spatial concentration of the IT sector in Romania's leading counties: București (the capital), Cluj, Iași, and Timiș.

This map also draws attention to heterogeneity in the share of employees in the IT sector who actually benefit from the tax exemption. This share reaches a peak in Timiș (around 46%) and a low in Teleorman (under 5%). This heterogeneity is likely to reflect differences across workers and firms in their ability to meet the requirements of the tax break (firms in less developed counties are more likely to hire workers with a profile that does not meet the educational criteria, such firms are also less likely to earn more than 10,000 U.S. dollars per eligible employee etc.). Discrepancies are also likely to reflect differential abilities across firms to fill in the necessary paperwork for a worker to be granted the tax break. From Table B4 we learn that IT firms located in counties with a higher share of exemption of the total workforce in IT are more likely to have a higher share of workforce exemption themselves. This finding points to potentially localized knowledge spillovers on the necessary procedures to benefit from the tax break and the importance of a strong local pool of skilled programmers.

We also compute the county-level number of workers in the eligible sectors per 100,000 inhabitants. While Bucureşti remains the leader in the country with 2,257 workers in the eligible sectors per 100,000 inhabitants, Cluj continues to stand out in a comparable position with 2,078 such workers per 100,000 inhabitants. The following three performers are Timiş, Braşov, and Iaşi, each having between 580 to 660 workers in IT per 100,000 inhabitants. The lowest-ranked 30 counties are home to less than 100 workers in the eligible sectors per 100,000 inhabitants.

Appendix B.2 Additional Robustness Checks on the Firm-Level Effects of the 2013 Reform to the Income Tax Break

Table B5: Difference-in-Differences Design Around 2013 Reform: Robustness of the Baseline Results to the Threshold Choice

Administrative data	Threshold 10%			Threshold 15%			Threshold 20% (Baseline)			Threshold 25%			Threshold 30%		
	Revenues	Workers	Assets	Revenues	Workers	Assets	Revenues	Workers	Assets	Revenues	Workers	Assets	Revenues	Workers	Assets
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
$\beta_{DiD,2011}$	-0.069*	0.020	-0.048	-0.054	0.031	-0.022	-0.054	0.031	-0.005	-0.043	0.038*	0.003	-0.007	0.048**	0.018
	(0.038)	(0.018)	(0.043)	(0.041)	(0.019)	(0.046)	(0.043)	(0.019)	(0.046)	(0.044)	(0.020)	(0.046)	(0.041)	(0.021)	(0.046)
$\beta_{DiD,2013}$	0.176***	0.147***	0.088***	0.168***	0.147***	0.080**	0.151***	0.135***	0.067*	0.120**	0.128***	0.060	0.103**	0.118***	0.042
	(0.047)	(0.022)	(0.033)	(0.049)	(0.023)	(0.035)	(0.049)	(0.024)	(0.037)	(0.050)	(0.024)	(0.038)	(0.052)	(0.025)	(0.039)
$\beta_{DiD,2014}$	0.273***	0.169***	0.167***	0.255***	0.160***	0.155***	0.235***	0.137***	0.148***	0.198***	0.118***	0.149***	0.187***	0.100***	0.135***
	(0.052)	(0.029)	(0.042)	(0.054)	(0.031)	(0.043)	(0.054)	(0.031)	(0.045)	(0.056)	(0.031)	(0.047)	(0.058)	(0.032)	(0.050)
$\beta_{DiD,2015}$	0.237***	0.139***	0.180***	0.222***	0.122***	0.164***	0.203***	0.102***	0.170***	0.153**	0.075**	0.163***	0.150**	0.065*	0.172***
	(0.062)	(0.035)	(0.051)	(0.065)	(0.036)	(0.053)	(0.067)	(0.037)	(0.055)	(0.068)	(0.037)	(0.057)	(0.070)	(0.039)	(0.059)
β_{DiD}	0.259***	0.143***	0.165***	0.239***	0.130***	0.142***	0.220***	0.112***	0.130***	0.177***	0.091***	0.122***	0.151***	0.075***	0.107**
	(0.048)	(0.026)	(0.039)	(0.050)	(0.027)	(0.041)	(0.050)	(0.027)	(0.042)	(0.050)	(0.027)	(0.043)	(0.051)	(0.028)	(0.044)
# Observations	27,950	27,950	27,950	27,666	27,666	27,666	27,491	27,491	27,491	27,346	27,346	27,346	27,144	27,144	27,144

Amadeus data	Threshold 10%			Threshold 15%			Threshold 20% (Baseline)			Threshold 25%			Threshold 30%		
$\beta_{DiD,2008}$	-0.125*	-0.073*	-0.013	-0.112	-0.064	-0.001	-0.092	-0.059	0.026	-0.051	-0.026	0.063	-0.019	-0.013	0.057
	(0.065)	(0.041)	(0.070)	(0.070)	(0.044)	(0.075)	(0.074)	(0.046)	(0.079)	(0.074)	(0.046)	(0.081)	(0.078)	(0.049)	(0.086)
$\beta_{DiD,2009}$	-0.034	-0.043	-0.030	-0.016	-0.046	-0.021	-0.003	-0.038	-0.000	0.050	-0.015	0.034	0.083	-0.014	0.035
	(0.056)	(0.036)	(0.058)	(0.060)	(0.038)	(0.062)	(0.063)	(0.039)	(0.065)	(0.063)	(0.040)	(0.066)	(0.066)	(0.042)	(0.069)
$\beta_{DiD,2010}$	-0.066	-0.015	0.002	-0.020	0.007	0.029	-0.028	0.003	0.035	-0.002	0.009	0.034	0.014	0.005	0.034
	(0.055)	(0.031)	(0.050)	(0.056)	(0.032)	(0.052)	(0.060)	(0.034)	(0.055)	(0.061)	(0.035)	(0.057)	(0.063)	(0.036)	(0.059)
$\beta_{DiD,2011}$	-0.062	-0.001	-0.008	-0.042	0.012	0.015	-0.057	0.009	0.012	-0.014	0.026	0.027	0.006	0.037	0.034
	(0.046)	(0.023)	(0.042)	(0.049)	(0.025)	(0.045)	(0.051)	(0.026)	(0.047)	(0.050)	(0.027)	(0.047)	(0.050)	(0.028)	(0.047)
$\beta_{DiD,2013}$	0.155***	0.105***	0.108***	0.147***	0.108***	0.098***	0.147***	0.110***	0.098***	0.147***	0.106***	0.098**	0.132***	0.102***	0.080**
	(0.036)	(0.021)	(0.033)	(0.038)	(0.022)	(0.035)	(0.038)	(0.023)	(0.036)	(0.040)	(0.024)	(0.038)	(0.042)	(0.025)	(0.039)
$\beta_{DiD,2014}$	0.236***	0.126***	0.164***	0.221***	0.115***	0.143***	0.215***	0.110***	0.143***	0.210***	0.100***	0.148***	0.194***	0.087***	0.133***
	(0.046)	(0.028)	(0.042)	(0.049)	(0.029)	(0.043)	(0.050)	(0.030)	(0.045)	(0.053)	(0.031)	(0.047)	(0.055)	(0.033)	(0.050)
$\beta_{DiD,2015}$	0.203***	0.108***	0.201***	0.178***	0.084**	0.174***	0.172***	0.080**	0.190***	0.158**	0.059	0.188***	0.155**	0.058	0.199***
	(0.058)	(0.035)	(0.053)	(0.062)	(0.037)	(0.055)	(0.064)	(0.038)	(0.058)	(0.067)	(0.040)	(0.060)	(0.070)	(0.042)	(0.061)
β_{DiD}	0.243***	0.134***	0.159***	0.210***	0.116***	0.128***	0.206***	0.113***	0.124***	0.171***	0.088***	0.110**	0.142***	0.079***	0.101**
	(0.043)	(0.026)	(0.043)	(0.044)	(0.027)	(0.044)	(0.046)	(0.028)	(0.047)	(0.047)	(0.028)	(0.048)	(0.048)	(0.029)	(0.049)
# Observations	33,301	33,301	33,301	32,948	32,948	32,948	32,715	32,715	32,715	32,504	32,504	32,504	32,257	32,257	32,257

Notes: In this exercise, we check the robustness of the baseline results from Table 3 on the firm-level impact of the 2013 amendment to the income tax break law of 2001. The samples in this table differ in two ways. The upper panel uses the administrative data, while the lower panel uses the Amadeus data. Columns differ in the threshold above which we consider a firm to be treated by the 2013 amendment, i.e., instead of allowing $Exempted_{isk}$ to become 1 only when the firm passes above the 20% threshold of exempted employees, we change this threshold to 10, 15, 25 and 30%. In each sample firms that have above 5% exempted employees post-2013 but under the given threshold are excluded. Robust standard errors, clustered at the firm-level, in parenthesis. All specifications include firm and sector-by-year fixed effects.***,**,* denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Table B6: Robustness of the Baseline Results to the Definition of Treatment. Here Treatment Based on Firm Sector

	Baseline 2013 Spec.			Baseline 2001 Spec.			Enlarged 2001 Spec.		
	Revenues	Workers	Assets	Revenues	Workers	Assets	Revenues	Workers	Assets
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\beta_{DiD,2011}$	-0.054	0.031	-0.005	-0.006	0.000	-0.013	-0.019	-0.006	-0.025
	(0.043)	(0.019)	(0.046)	(0.021)	(0.009)	(0.021)	(0.020)	(0.009)	(0.021)
$\beta_{DiD,2013}$	0.151***	0.135***	0.067*	0.039	0.040***	-0.022	0.059**	0.054***	-0.010
	(0.049)	(0.024)	(0.037)	(0.025)	(0.011)	(0.018)	(0.024)	(0.011)	(0.018)
$\beta_{DiD,2014}$	0.235***	0.137***	0.148***	0.096***	0.081***	0.007	0.125***	0.104***	0.025
	(0.054)	(0.031)	(0.045)	(0.031)	(0.015)	(0.024)	(0.030)	(0.015)	(0.024)
$\beta_{DiD,2015}$	0.203***	0.102***	0.170***	0.074**	0.096***	0.006	0.102***	0.123***	0.021
	(0.067)	(0.037)	(0.055)	(0.035)	(0.017)	(0.029)	(0.034)	(0.017)	(0.028)
Adjusted R^2	0.848	0.910	0.905	0.847	0.910	0.904	0.848	0.908	0.904
# Observations	27,491	27,491	27,491	27,491	27,491	27,491	28,856	28,856	28,856
# Firms	6,146	6,146	6,146	6,146	6,146	6,146	6,437	6,437	6,437

Notes: In this exercise we study the effects of the 2013 expansion to the income tax break law of 2001 using the same DiD design used to study the initial introduction of the law in 2001. Specifically, treatment is assigned to firms based on whether their sector was eligible for the income tax break law or not, interacted with a dummy activated when $t \geq 2013$ (as in Equation (1), for year 2001), as opposed to their workforce exemption status. Columns (4)-(6) use the same baseline sample as in columns (1)-(3) from Table 3. In columns (7)-(9), to ensure comparability with the baseline sample, we only keep firms with under 5% exemption rate before 2013 – irrespective of their post-2013 exemption rate. For reference, in columns (1)-(3), we reproduce columns (1)-(3) of Table 3, estimated using the specification in Equation (2). In this table, we use the baseline control group, which includes firms in high-tech knowledge-intensive service sectors (as classified by the Eurostat). Also, in this table, we focus on the administrative data. The outcome variables used are log revenues, log number of workers, and log total assets. Robust standard errors, clustered at the firm-level, in parenthesis. All specifications in columns (4)-(9) include firm, year, and sector fixed effects. Note that sector-by-year fixed effects cannot be included in those columns, as in Equation (2). ***,**,* denotes statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix C Additional Sector-Level Evidence

Appendix C.1 Descriptive Statistics with Sector-Level Data

Appendix C.1.1 Sector Code Switches of Existing Firms to the IT Sector

One channel through which the income tax break law could have boosted the growth of the Romanian IT sector is by changing incentives of firms (workers) to operate in the eligible sector. Using a sector-level transition matrix from the National Bank of Romania, we study transitions into the eligible sector (NACE Rev 1 code 722) before and after the introduction of the 2001 income tax break. While this evidence is descriptive, it highlights abnormal trends in transitions after 2001.

Figure C1 plots the absolute number of firms alive in year $(t-1)$ and with an economic activity different from 722, which switched in year t to the NACE Code 722. Figure C2 plots the year t share of firms in NACE Code 722 alive in year $(t-1)$ coming from firms with a different economic activity in $(t-1)$ that have switched their year t economic activity to 722. We notice that while switching one's sector towards 722 was a trend occurring both before and long after the passing of the 2001 law, immediately after 2001, switching to NACE Rev 1 sector 722 became a visibly more prevalent practice. From 2004 onward, the momentum of this switching practice is lost, which suggests that most justifiable switches occurred immediately after the passage of the law.

One caveat to this exercise is that we cannot pin down the exact reasons behind these switches in firms' main sector code. Hence, this type of growth of the sector may be to some extent artificial, if firms switching their main sectoral activity were already conducting most of their activity in the eligible sector, but had a misassigned main sector. As the law made the income tax break eligibility dependent on a firm's main sector code, the law may have incentivized corrections in firms' sector codes.

Most switches in firms' main sector code occurred from sector codes complementary to the eligible sector (e.g., 721 - "Consulting in the field of computing equipment / hardware" and 726 - "Other activities related to computer science"). We conjecture that these switching firms were multi-activity firms that decided to focus on software development once the income tax break for programmers became available. One might be concerned that these sector code switches were meant to deceit tax authorities. We do not believe that this concern is justified because the sector of the firm was only one of several strict requirements for a worker to qualify for the exemption. In particular, the firm had to provide thorough evidence that revenues of at least 10,000 U.S. dollars per exempted employee came from software development and that employees involved in this software development had eligible bachelor degrees.

For these reasons, the evidence in Figures C1 and C2 suggests that switches were likely to be motivated by an actual switch of focus of the firm on software development. Even if part of these switches may not have resulted in a stronger focus on software development (had some of these firms had a previously misassigned main sector), it is plausible to expect that the income tax break law has strengthened incentives to focus on software development. Moreover, the stark jump in switches around 2001 is a convincing evidence of the salience of the law.

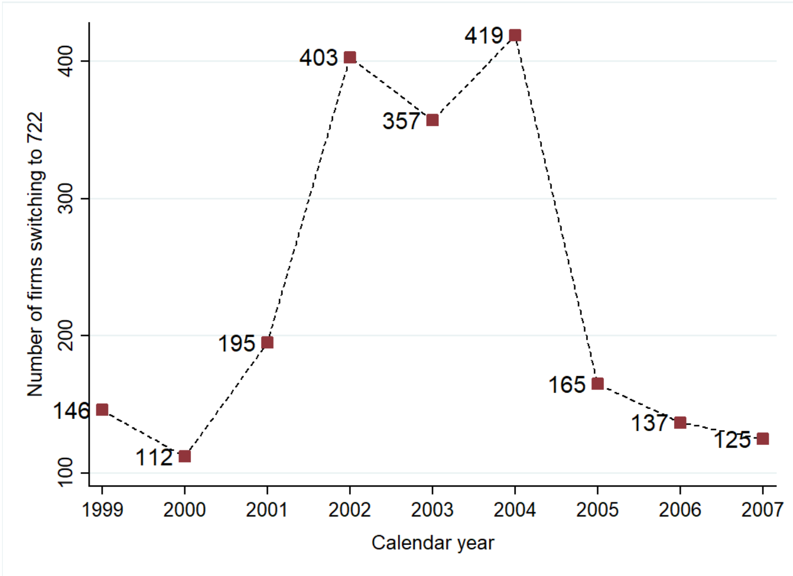

Figure C1: Romania: Number of firms alive in year $(t-1)$ and switching their economic activity in t towards the beneficiary sector, 722

Notes: Data source: National Bank of Romania. This graph plots the number of firms alive in year $(t-1)$ (with an economic activity different from 722) that switched in year t to the NACE Code 722 (7221 and 7222, Software consultancy and supply). We notice two peaks in this practice in 2002 and 2004, the years after the first and second laws of income tax exemption for workers in sector 722 were passed in Romania.

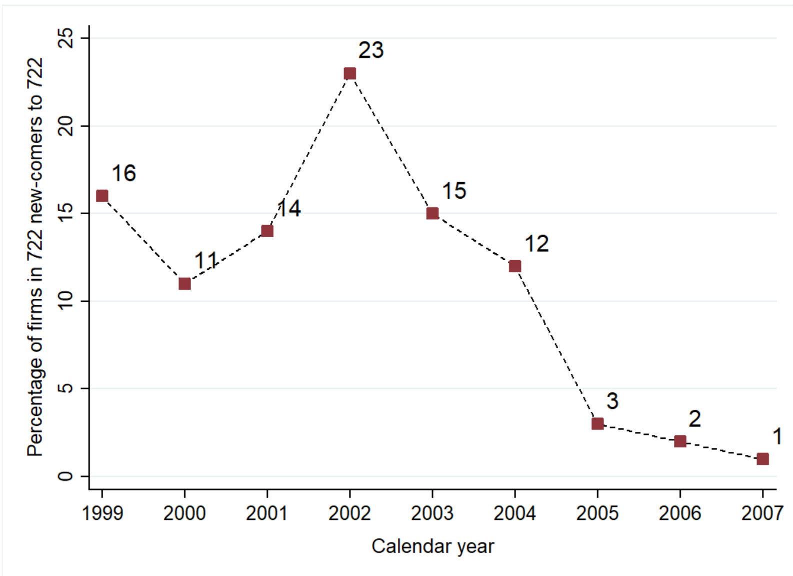

Figure C2: Romania: Among firms in NACE code 722 in year t and alive in year $(t-1)$ (with any economic activity), the percentage of new-comers in t to 722

Notes: Data source: National Bank of Romania. This graph plots the year t share of firms in NACE Code 722 (7221 and 7222, Software consultancy and supply) alive in year $(t-1)$ coming from firms with a different economic activity in $(t-1)$ that have switched their year t economic activity to 722. We notice a clear peak in this practice in 2002, the year after the first law of income tax exemption for workers in sector 722 was passed in Romania.

Appendix C.1.2 Firm Entries and Exits in the IT Sector

The income tax break may have also affected the growth of the software sector by increasing the birth rate of new firms in the sector and/or decreasing firm death rates. In this section, we provide descriptive evidence on the evolution of firm birth and death rates in the sector (722, in NACE Rev 1.1 classification) in the years before and after the introduction of the policy. The birth/entry (death/exit) rate at time t is defined as the number of firm births/entries (deaths/exits) relative to the population of active firms at the beginning of time t. Our analysis is based on sector-level data from the Eurostat Business Demography database. Figures C3 and C4 show the evolution of birth and death rates in the software sector between 2000 and 2007 for Romania and four other comparable Central and Eastern European (CEE) countries. These other countries were chosen based on data availability for the entire 2000 to 2007 period and on their similarity to Romania (i.e., they are transition economies, at similar stages of development, and with similar economic structures).

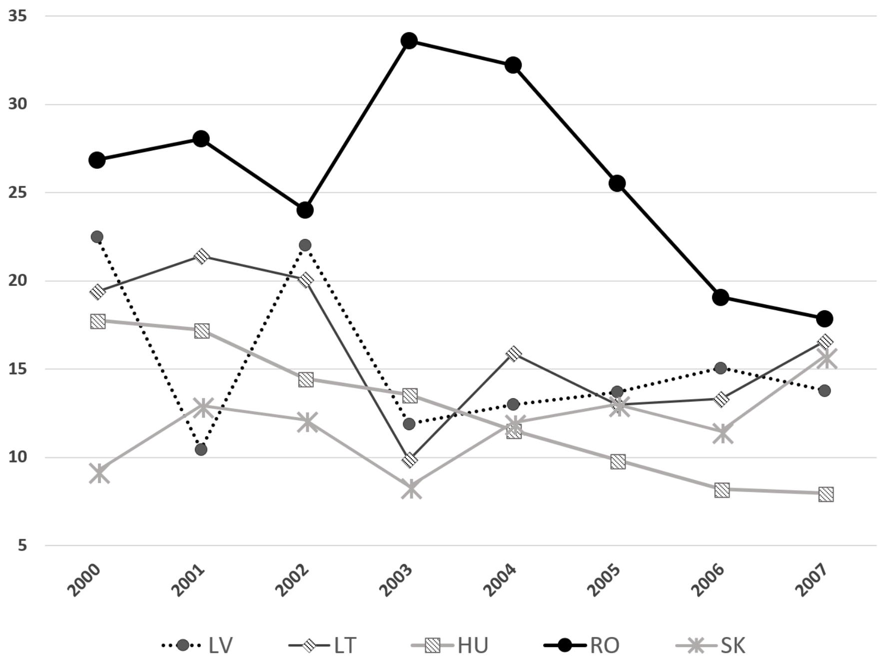

Figure C3: Entry Rates of Firms in the "Software Consultancy and Supply" sector in Romania and Comparable Countries from Central and Eastern Europe

Notes: Data source: Sector-level data from the Eurostat Business Demography database. The entry rate in year t in a given country is defined as the number of firm entries (births) relative to the population of active firms in that country at the beginning of t.

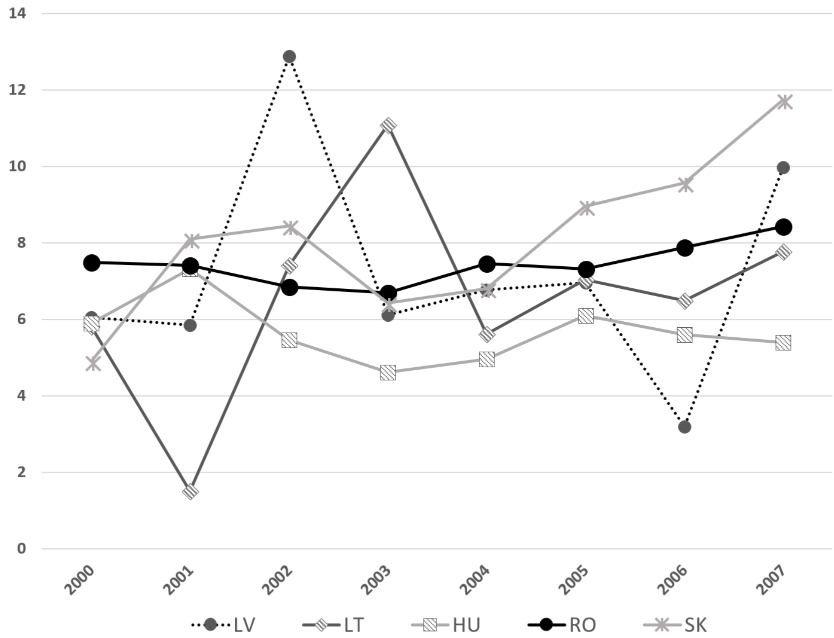

Figure C4: Death Rates of Firms in the "Software Consultancy and Supply" Sector in Romania and Comparable Countries from Central and Eastern Europe

Notes: Data source: Sector-level data from the Eurostat Business Demography database. The exit rate in year t in a given country is defined as the number of firm exits (deaths) relative to the population of active firms in that country at the beginning of t.

Figure C3 shows that the birth rate of firms in the IT sector in Romania experienced a notable increase in 2003 and 2004 relative to its pre-tax break levels.[41] No similar increase can be observed for any of the other four CEE countries. There is a two-year lag in the peak of the firm birth rate in Romania, which is consistent with a need for entrepreneurs to be reassured that the tax break was not to be reverted in the short-term. Figure C4 shows the evolution of firm death rates in the IT sector in Romania and in comparable CEE countries. While the death rate in Romania does not seem to be affected by the introduction of the tax break, throughout the period of analysis Romanian firm death rates remain less volatile than those in comparable countries. The combined trends in firm birth and death rates lead to an increasing trend in the stock of firms in the IT sector in Romania.

Appendix C.1.3 Firm Entry through Foreign Direct Investment in IT

In terms of foreign ownership in the IT sector of Romania, the early 2000s were the turning point. Between 2000 and 2004, the first multinational firms decided to offshore part of their operations to Romania [Pruna and Soleanicov, 2012]. Moreover, during the same period, there were also smaller foreign companies that acquired Romanian firms, e.g., Adobe Systems Inc. who acquired InterAKT [Pruna and Soleanicov, 2012]. Ever since, Romania has continued attracting steady FDI inflows into the IT sector, to the point that in 2017 the biggest players in the IT sector of Romania were multinationals (Oracle, IBM, Ericsson, and Endava). Bitdefender is the only Romanian firm with comparable operating

[41] As shown in Appendix C.1.3, from 2003 to 2016, the stock of FDI in the IT sector of Romania has been multiplied by twenty. A sizable share of the firm entry into the IT sector is likely to be due to foreign firm entry.

revenue from its operations in Romania.[42] In 2017 foreign-owned companies produced 73% of the gross revenues and employed 59% of the total number of workers in the Software and IT Services of Romania.

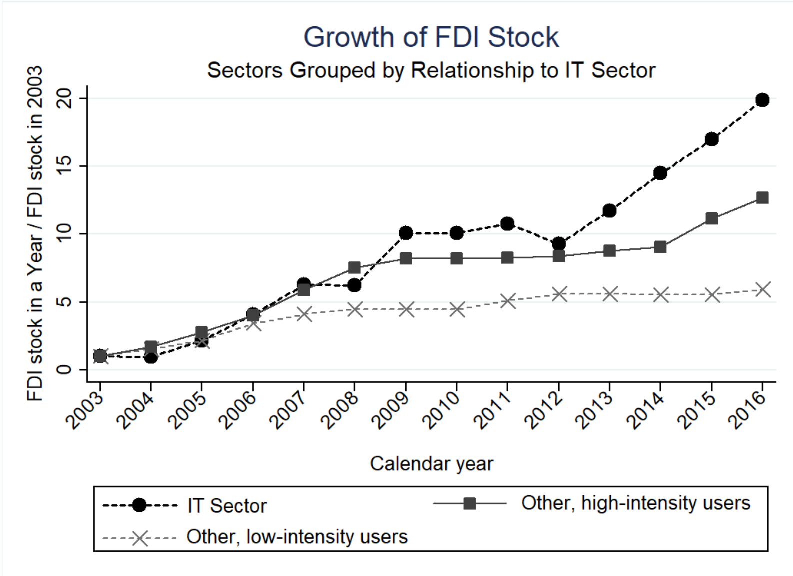

Figure C5: Growth of the Stock of FDI in Romania, by Relationship to the IT Sector

Notes: Data sources: National Bank of Romania (for historical sector-level FDI stocks) and OECD (for the Romanian input-output table for the year 2000, at the two-digit NACE Rev 1 level). This figure plots the growth of the FDI stock in Romania between 2003 and 2016 (defined as the yearly value of the FDI stock divided by the relevant FDI stock value in 2003) for three categories of sectors. The first category is the sector 72 ("Computer and related activities") itself. The second category called "Other, high-intensity users" contains all two-digit NACE Rev 1 sectors that are in the top quarter of the distribution of sector-level shares of input purchases from sector 72 in a sector's total input expenditure (except sector 72). These shares are computed based on Romania's input-output table for the year 2000. The last category of sectors called "Other, low-intensity users" contains all other sectors for which IT inputs are less important. The following are the NACE Rev 1 codes and names of the sectors in the "Other, high-intensity users" category: 50-52 (Wholesale and retail trade; repairs), 60-63 (Transport and storage), 70 (Real estate activities), 71 (Renting of machinery and equipment), 73-74 (R&D and other business activities), 75 (Public administration and defence; compulsory social security), 80 (Education), and 85 (Health and social work).

An ideal dataset to study the causal effects of the income tax break for workers in IT on Foreign Direct Investment (FDI, henceforth) into the IT sector of Romania would have to start before 2001, include not only Romania but also comparable countries, and be disaggregated at the sector-level. Unfortunately, we have not found such a dataset. For Romania, the longest sector-level FDI time series (from the National Bank of Romania) starts in 2003. With this data, we study the growth of the FDI stock for three categories of sectors: the IT sector itself, the group of sectors that are "high-intensity users" of IT services, and the "low-intensity users".[43] Figure C5 plots the growth of the FDI stock in Romania between 2003 and 2016 (defined as the yearly value of the FDI stock divided by the relevant FDI stock value in 2003) for these three categories of sectors. The figure shows that sectors with the highest share of IT services in their input expenditure are those experiencing the highest growth of their FDI stock since 2003. The IT sector itself is experiencing an even starker growth in its FDI stock. The importance of foreign revenues has also been growing steadily over the years, to the point that in 2017 total foreign revenues were three times larger than domestic revenues.[44]

[42] See December 2017 article from the *Romanian Journal*.
[43] See Section 4.2 for details on how we assign sectors into either the high- or low-usage groups.
[44] Source of last two sets of statistics: the Association of Employers in the Software and Services Industry or ANIS.

Appendix C.2 Additional Robustness Checks on the Sector-Level Cross-Country Effects of the 2001 Tax Break

Appendix C.2.1 IT Sector Vs. Rest of the Economy

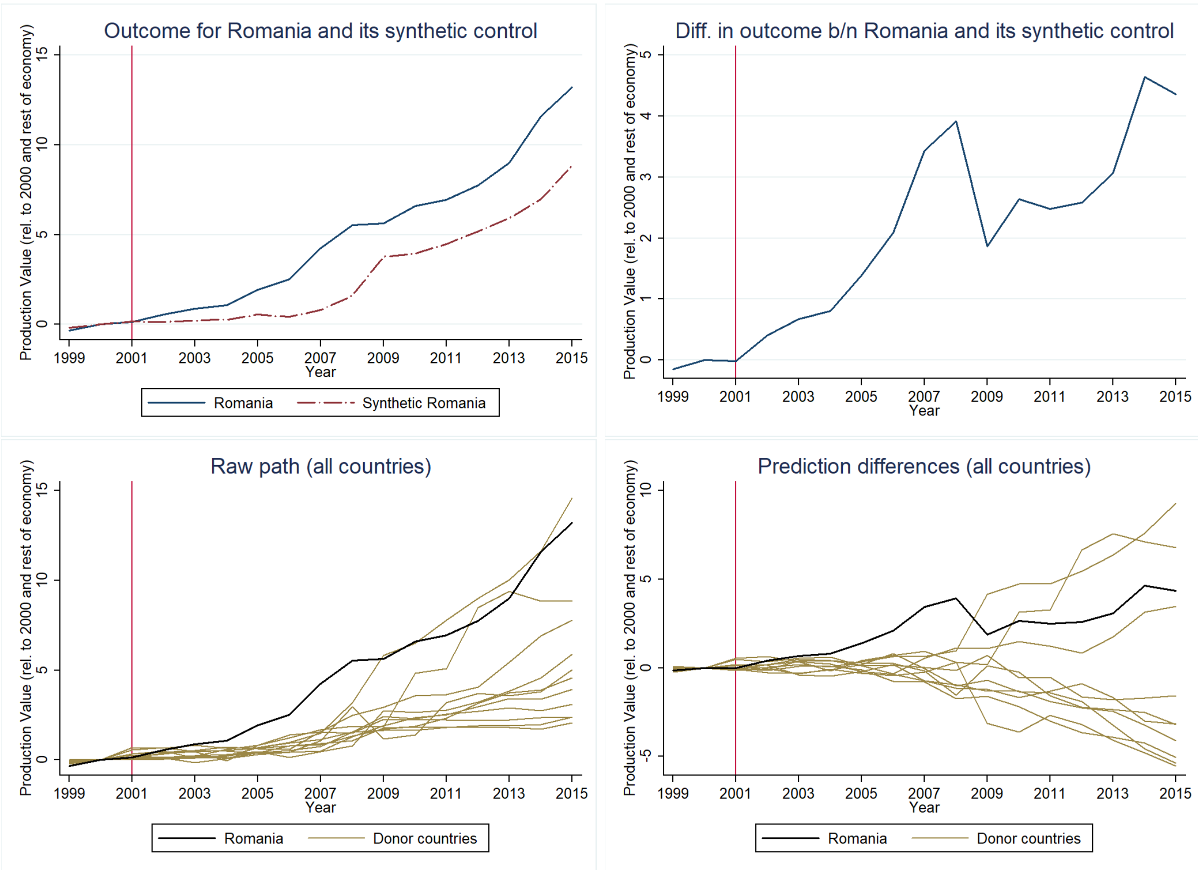

Figure C6: IT Sector Vs. Rest of Economy. SCM with Outcome Variable: "Production Value" - Million Euros (Normalized)

Notes: In this exercise we use the synthetic control method introduced in Section 4.1 to study the sector-level direct effects of the introduction of the 2001 law granting an income tax break to workers in IT. All figures have as dependent variable the country-level (normalized) " Production value - million euros" The yearly absolute value of the dependent variable in the treated sector is divided by its value in 2000, the year prior to the introduction of the income tax break in Romania. From these resulting yearly ratios we subtract the corresponding ratios for the comparison sectors. The treated sector is K72 ("Computer and related activities," including "Software consultancy and supply " and "Publishing of software"). We use as comparison sectors all other sectors in the economy (all except K72). The data source for the dependent variable is Eurostat, Structural Business Statistics, Annual detailed enterprise statistics on services (NACE Rev 1.1). Data for the predictors comes from the World Bank, World Development Indicators. We use as predictors the "GDP per capita (constant LCU)," "Medium and high-tech sector (% manufacturing value added)" and "Services, etc., value added (% of GDP)." All figures are an output of the **synth_runner** package for Stata [Quistorff and Galiani, 2017], with the *nested* option specified.

Table C1: Post-treatment Results: Effects, p-values, Standardized p-values. Outcomes: Gross Revenues, Employment, and Production Value

	Gross Revenues			**Employment**			**Production Value**		
Post-treatment year	$\hat{\alpha}_{1t}$	p-values	Standardized p-values	$\hat{\alpha}_{1t}$	p-values	Standardized p-values	$\hat{\alpha}_{1t}$	p-values	Standardized p-values
2001	0.13	0.27	0.55	0.14	0.27	0.00	-0.01	1.00	1.00
2002	0.69	0.00	0.27	0.26	0.09	0.00	0.41	0.10	0.50
2003	1.13	0.00	0.27	0.67	0.09	0.00	0.67	0.00	0.80
2004	1.37	0.00	0.27	1.13	0.00	0.00	0.80	0.00	0.30
2005	1.95	0.00	0.36	1.46	0.00	0.00	1.39	0.00	0.40
2006	2.57	0.00	0.36	1.75	0.00	0.00	2.10	0.00	0.50
2007	3.17	0.00	0.36	2.10	0.00	0.00	3.43	0.00	0.30
2008	3.21	0.00	0.27	1.90	0.00	0.00	3.92	0.00	0.40
2009	4.14	0.09	0.27	1.17	0.27	0.00	1.87	0.20	0.70
2010	4.49	0.09	0.27	1.03	0.45	0.00	2.64	0.30	0.70
2011	2.92	0.36	0.27	1.12	0.36	0.00	2.48	0.40	0.70
2012	3.32	0.45	0.36	1.19	0.18	0.00	2.59	0.40	0.90
2013	4.52	0.36	0.36	1.32	0.18	0.00	3.06	0.50	0.90
2014	6.15	0.36	0.36	1.54	0.18	0.00	4.64	0.30	0.70
2015	6.52	0.27	0.36	1.83	0.18	0.00	4.36	0.50	0.80

Notes: In this exercise we use the synthetic control method (SCM) introduced in Section 4.1 to study the sector-level direct effects of the introduction of the 2001 law granting an income tax break to workers in IT. This table presents the SCM estimates of the treatment effects ($\hat{\alpha}_{1t}$) of the 2001 policy on three outcome variables (relative gross revenues, relative employment, and the relative production value) in Romania (where 1=Romania). This table corresponds to the treatment effects plotted in Figures 3, 4, and C6. Treatment effects, $\hat{\alpha}_{1t}$, for Romania in post-treatment years t (2001 to 2015) are estimated as the difference between the year t outcomes for Romania and those for synthetic Romania (for details see Equation (3) in Section 4.1). The p-values are constructed using permutation tests. To calculate a p-value for each post-treatment period, the **synth_runner** package for Stata [Quistorff and Galiani, 2017] finds the proportion of placebo effects that are as large as the main effect. The standardized p-values are the studentized p-values. See Table C2 for the combination of countries and weights that comprise the synthetic control.

Table C2: Predictor balance, averaged between 1999 and 2000. Synthetic Romania = 0.104 Czech Republic and 0.896 Slovakia (Gross Revenues); = 0.534 Bulgaria and 0.466 Czech Republic (Employment); = 0.339 Bulgaria and 0.661 Lithuania (Production Value)

	Romania	Synthetic Romania		
		Gross Revenues	Employment	Prod. Value
GDP per capita (constant LCU, $)	9,663	9,796	15,972	5,341
Services (% of GDP)	53.23	60.32	60.07	62.59
High-tech manufacturing (% mfg. value added)	23.85	35.69	34.05	23.85

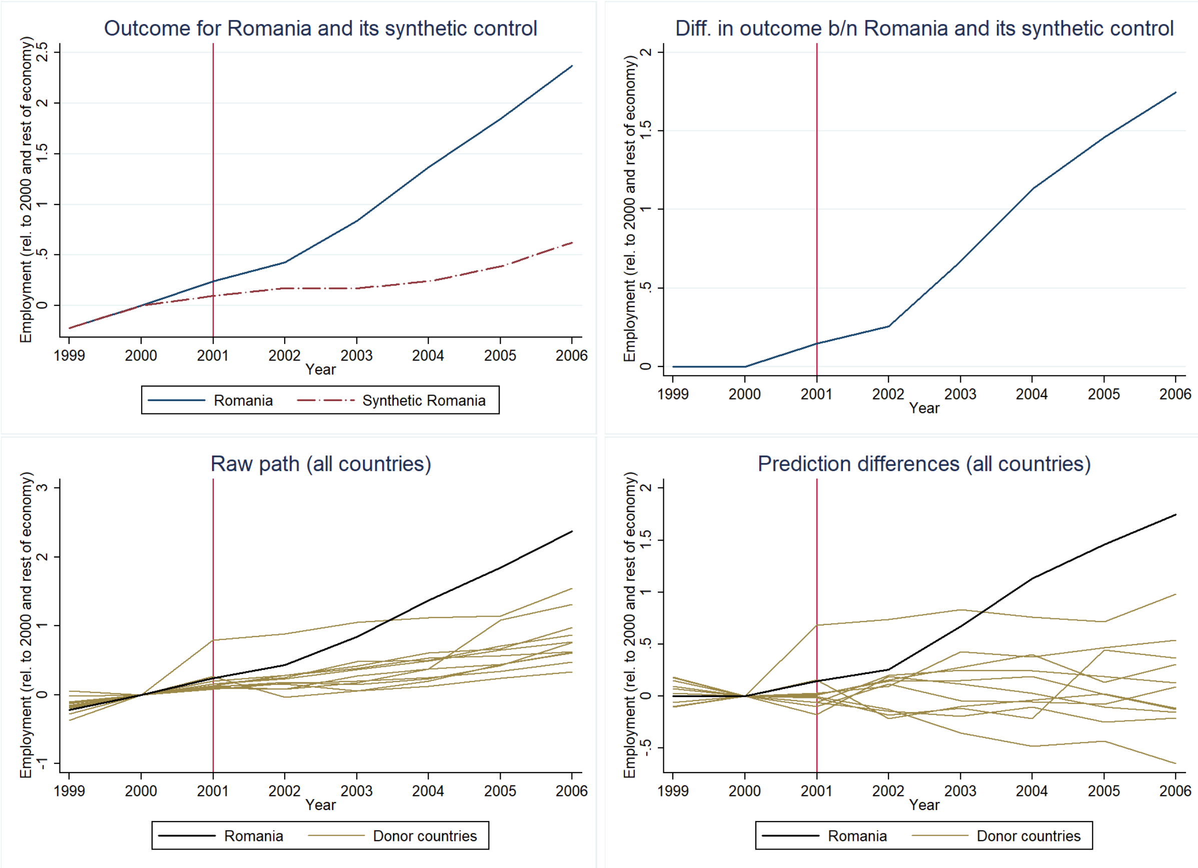

Figure C7: IT Sector Vs. Rest of Economy. SCM with Outcome Variable: "Employees - Number" (Normalized). Robustness to the Correspondence between NACE Sector Classifications

Notes: In this exercise we use the synthetic control method introduced in Section 4.1 to study the sector-level direct effects of the introduction of the 2001 law granting an income tax break to workers in IT. This exercise differs from the one presented in Figure 4 in the period of analysis. Instead of considering the entire 1999 to 2015 period, we cut the analysis in 2006 (the last year where the data is available in the NACE Rev 1 classification). This tests the robustness of our results to the correspondence we develop between the NACE Rev 1 and NACE Rev 2 classifications. All figures have as dependent variable the country-level (normalized) "Employees - number." The yearly absolute value of the dependent variable in the treated sector is divided by its value in 2000, the year prior to the introduction of the income tax break in Romania. From these resulting yearly ratios we subtract the corresponding ratios for the comparison sectors. The treated sector is K72 ("Computer and related activities," including "Software consultancy and supply " and "Publishing of software"). We use as comparison sectors all other sectors in the economy (all except K72). The data source for the dependent variable is Eurostat, Structural Business Statistics, Annual detailed enterprise statistics on services (NACE Rev 1.1). Data for the predictors comes from the World Bank, World Development Indicators. We use as predictors the "GDP per capita (constant LCU)," "Medium and high-tech sector (% manufacturing value added)" and "Services, etc., value added (% of GDP)." All figures are an output of the **synth_runner** package for Stata [Quistorff and Galiani, 2017], with the *nested* option specified.

Appendix C.2.2 IT-Using Sectors Vs. Non-IT Using Sectors

Table C3: sectors with Consistently-Available Data in Eurostat, by Quarter of IT-use Intensity

Under Quartile 1	**Between Quartiles 1-2**
Wood and products of wood and cork (20)	Manufacture of textiles and textile products (17)
Basic metals (27)	Manufacture of wearing apparel; dressing, dyeing of fur (18)
Fabricated metal products (28)	Manufacture of pulp, paper and paper products (21)
Electricity, gas and water supply (40-41)	Publishing, printing and reproduction of recorded media (22)
	Manufacture of machinery and equipment n.e.c. (29)
	Manufacture of furniture; manufacturing n.e.c. (36)
	Recycling (37)
	Construction (45)
Between Quartiles 2-3	**Over Quartile 3**
Manufacture of chemicals and chemical products (24)	Wholesale and retail trade (50-52)
Manufacture of rubber and plastic products (25)	Transport and storage (60-63)
Manufacture of other non-metallic mineral products (26)	Real estate activities (70)
Manufacture of office machinery and computers (30)	Renting of machinery and equipment (71)
Manufacture of electrical machinery and apparatus (31)	Computer and related activities (72)
Manufacture of radio, television and communication (32)	R&D and other business activities (73-74)
Manufacture of medical, precision, optical instruments (33)	
Manufacture of motor vehicles, trailers, semi-trailers (34)	
Manufacture of other transport equipment (35)	
Hotels and restaurants (55)	

Notes: This table groups sectors with consistently-available data in Eurostat (across variables, years, and countries - see Appendix D for details) based on their dependence on the IT sector for inputs. To establish this dependence, we start from the input-output table (I-O table, henceforth) of Romania for the year 2000. We use the harmonized I-O table provided by the OECD, that tracks the flows of goods and services between all two-digit NACE Rev 1 sectors in Romania. We then compute the share of the total input expenditures of a given sector purchased from the IT sector (NACE Rev 1 sector 72, "Computer and related activities"). It is based on this distribution of shares that we compute the quartiles mentioned in this table. In the main specification of the inter-sector SCM analysis we exclude sector 72 altogether from the analysis, as this is the treated sector itself. We also assign sectors over quartile three into the high-intensity category and all other sectors into the low-intensity. Note that the sectors actually used in the analysis and presented in this table are a subset of all sectors in the economy, as not all sectors had consistently-available data. Nevertheless, the grouping of these sectors by quarters was based on the full I-O matrix, which includes all sectors in the economy.

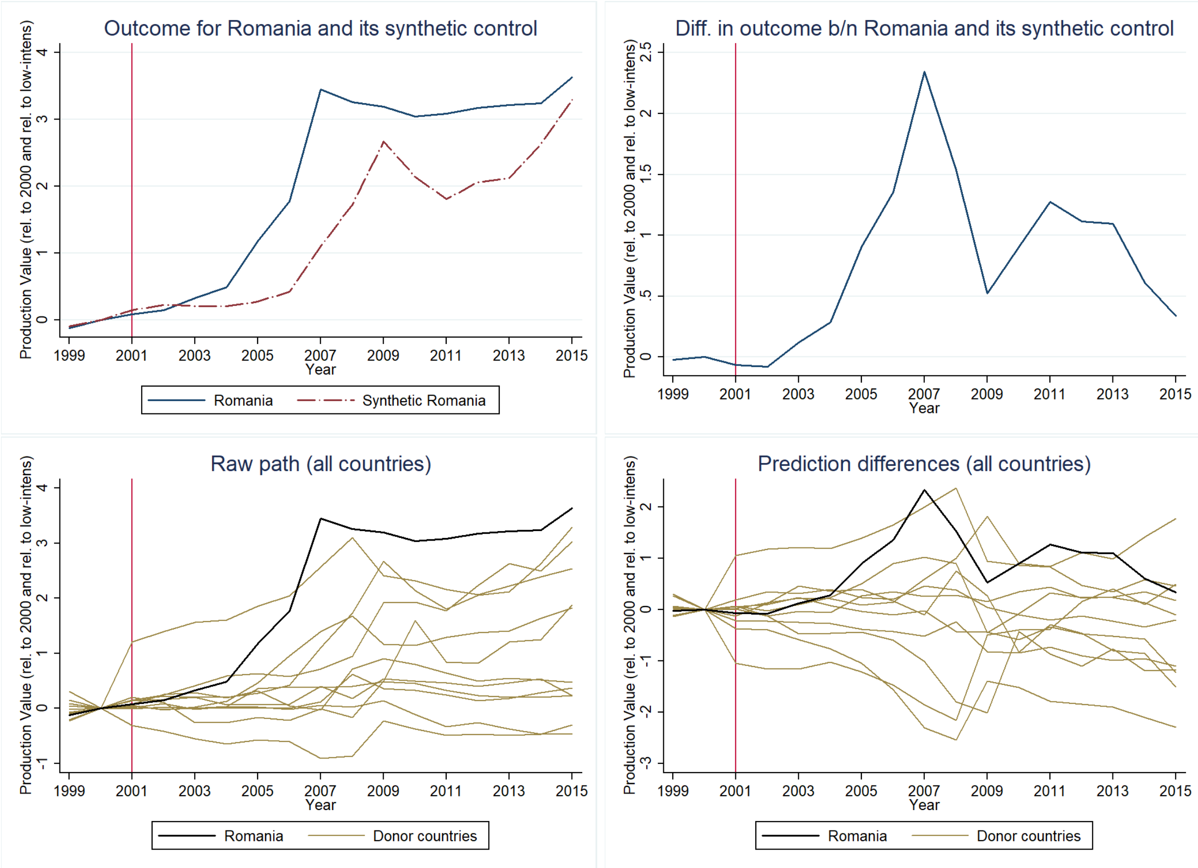

Figure C8: IT-Using Sectors Vs. Non-IT Using Sectors. SCM with Outcome Variable: "Production Value" - Million Euros (Normalized)

Notes: In this exercise we use the synthetic control method introduced in Section 4.2 to study the sector-level downstream effects of the introduction of the 2001 law granting an income tax break to workers in IT. All figures have as dependent variable the country-level (normalized) " Production value - million euros." The yearly absolute value of the dependent variable in the treated sector is divided by its value in 2000, the year prior to the introduction of the income tax break in Romania. From these resulting yearly ratios we subtract the corresponding ratios for the comparison sectors. The treated sectors are those that use K72 ("Computer and related activities," including "Software consultancy and supply" and "Publishing of software") services at high-intensity. We exclude K72 itself from this category. Sectors that have a low-intensity of use of K72 services serve as comparison sectors. The data source for the dependent variable is Eurostat, Structural Business Statistics, Annual detailed enterprise statistics on services (NACE Rev 1.1). Data for the predictors comes from the World Bank, World Development Indicators. We use as predictors the "GDP per capita (constant LCU)," "Medium and high-tech sector (% manufacturing value added)" and "Services, etc., value added (% of GDP)." All figures are an output of the **synth_runner** package for Stata [Quistorff and Galiani, 2017], with the *nested* option specified.

Table C4: Post-treatment Results: Effects, p-values, Standardized p-values. Outcomes: Gross Revenues, Employment, and Production Value

	Gross Revenues			**Employment**			**Production Value**		
Post-treatment year	$\hat{\alpha}_{1t}$	p-values	Standardized p-values	$\hat{\alpha}_{1t}$	p-values	Standardized p-values	$\hat{\alpha}_{1t}$	p-values	Standardized p-values
2001	-0.04	0.80	0.50	0.08	0.22	0.22	-0.07	0.64	0.64
2002	-0.02	1.00	0.50	0.09	0.22	0.11	-0.08	0.91	0.64
2003	0.18	0.30	0.50	0.27	0.22	0.11	0.12	0.82	0.55
2004	0.28	0.00	0.40	0.43	0.22	0.11	0.29	0.55	0.45
2005	0.54	0.00	0.40	0.68	0.00	0.11	0.90	0.27	0.27
2006	0.90	0.00	0.40	0.86	0.00	0.11	1.36	0.27	0.18
2007	1.71	0.00	0.40	0.93	0.00	0.11	2.34	0.00	0.18
2008	1.88	0.00	0.40	0.98	0.11	0.11	1.53	0.36	0.36
2009	1.46	0.00	0.40	0.77	0.00	0.11	0.52	0.45	0.27
2010	1.33	0.00	0.40	0.68	0.00	0.11	0.90	0.09	0.27
2011	1.34	0.00	0.40	0.73	0.00	0.11	1.27	0.09	0.27
2012	1.30	0.10	0.40	0.68	0.00	0.11	1.11	0.18	0.27
2013	0.68	0.30	0.40	0.67	0.00	0.11	1.10	0.09	0.27
2014	0.73	0.30	0.50	0.61	0.22	0.11	0.61	0.45	0.36
2015	0.75	0.50	0.50	0.61	0.11	0.11	0.34	0.73	0.55

Notes: In this exercise we use the synthetic control method (SCM) introduced in Section 4.2 to study the sector-level downstream effects of the introduction of the 2001 law granting an income tax break to workers in IT. This table presents the SCM estimates of the treatment effects ($\hat{\alpha}_{1t}$) of the 2001 policy on three outcome variables (relative gross revenues, relative employment, and the relative production value) in Romania (where 1=Romania). This table corresponds to the treatment effects plotted in Figures 5, 6, and C8. Treatment effects, $\hat{\alpha}_{1t}$, for Romania in post-treatment years t (2001 to 2015) are estimated as the difference between the year t outcomes for Romania and those for synthetic Romania (for details see Equation (3) in Section 4.1). The p-values are constructed using permutation tests. To calculate a p-value for each post-treatment period, the **synth_runner** package for Stata [Quistorff and Galiani, 2017] finds the proportion of placebo effects that are as large as the main effect. The standardized p-values are the studentized p-values. See Table C5 for the combination of countries and weights that comprise the synthetic control.

Table C5: Predictor balance, averaged between 1999 and 2000. Synthetic Romania = 0.341 Bulgaria and 0.659 Lithuania (Gross Revenues); = 0.363 Bulgaria + 0.002 Hungary + 0.625 Lithuania (Employment) ; = Bulgaria (Production Value)

	Romania	Synthetic Romania		
		Gross Revenues	Employment	Prod. Value
GDP per capita (constant LCU, $)	9,663	5,342	9,616	5,771
Services (% of GDP)	53.23	62.59	62.53	60.85
High-tech manufacturing (% mfg. value added)	23.85	23.87	24.15	30.53

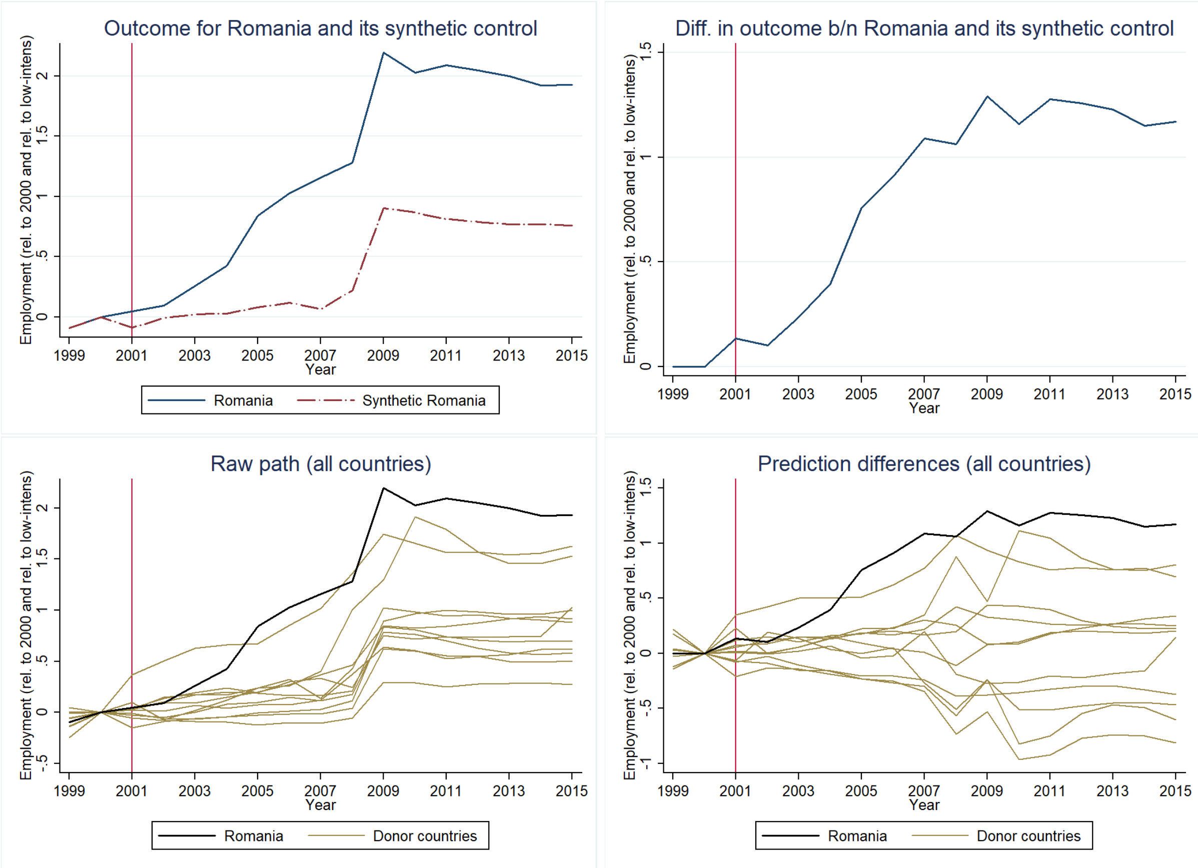

Figure C9: IT-Using Sectors Vs. Non-IT Using Sectors. SCM with Outcome Variable: "Employees - Number" (Normalized). Robustness to Exclusion of Sectors Comparable to IT from High-Intensity Sectors Category

Notes: In this exercise we use the synthetic control method introduced in Section 4.2 to study the sector-level downstream effects of the introduction of the 2001 law granting an income tax break to workers in IT. This exercise differs from the one presented in Figure 6 in its exclusion of sectors K73 and K74 from the analysis. Both sectors are otherwise part of the list of sectors that rely heavily on IT services (top 25% users of IT services). We used these sectors as comparison sectors in the firm-level analysis. All figures have as dependent variable the country-level (normalized) "Employees - number." The yearly absolute value of the dependent variable in the treated sector is divided by its value in 2000, the year prior to the introduction of the income tax break in Romania. From these resulting yearly ratios we subtract the corresponding ratios for the comparison sectors. The treated sectors are those that use K72 ("Computer and related activities," including "Software consultancy and supply" and "Publishing of software") services at high-intensity. We exclude K72 itself from this category. Sectors that have a low-intensity of use of K72 services serve as comparison sectors. The data source for the dependent variable is Eurostat, Structural Business Statistics, Annual detailed enterprise statistics on services (NACE Rev 1.1). Data for the predictors comes from the World Bank, World Development Indicators. We use as predictors the "GDP per capita (constant LCU)," "Medium and high-tech sector (% manufacturing value added)" and "Services, etc., value added (% of GDP)." All figures are an output of the **synth_runner** package for Stata [Quistorff and Galiani, 2017], with the *nested* option specified.

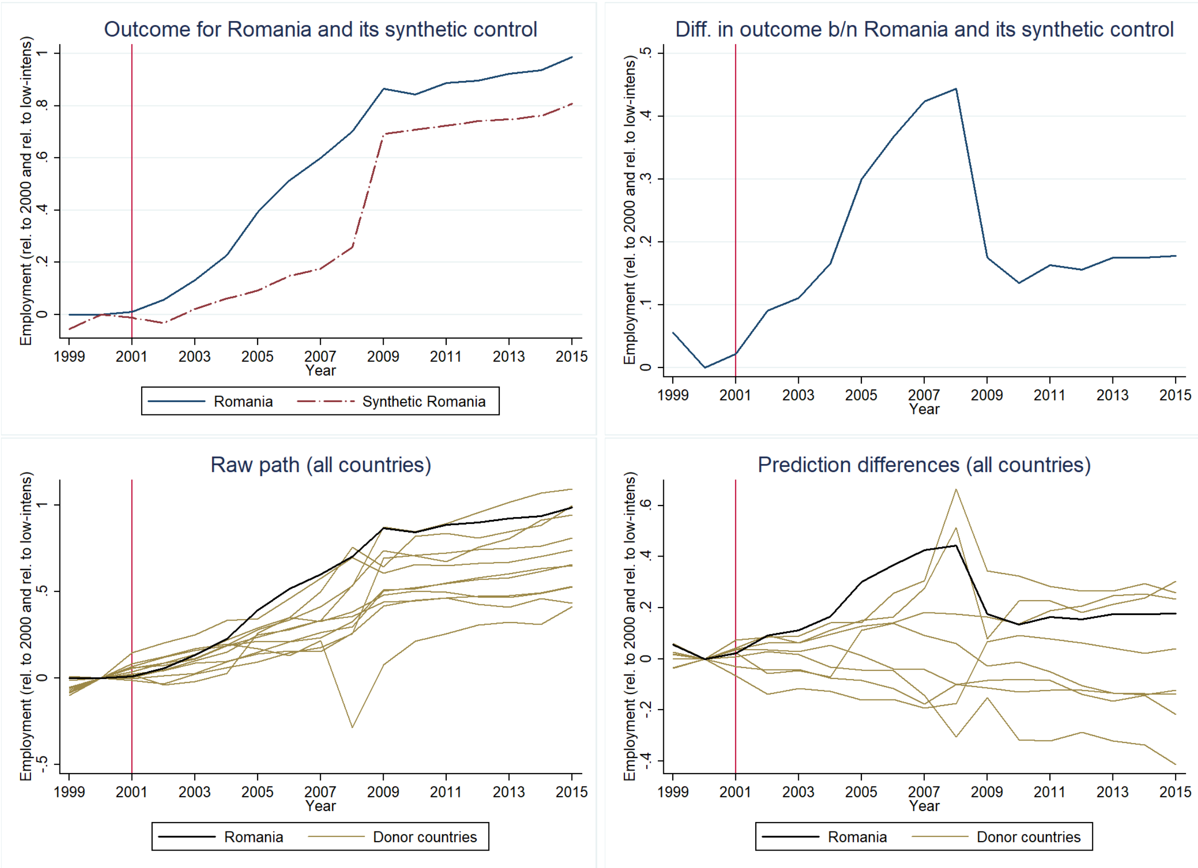

Figure C10: IT-Using Sectors Vs. Non-IT Using Sectors. SCM with Outcome Variable: "Employees - Number" (Normalized). Robustness to Grouping of Quarters into High- and Low-Intensity

Notes: In this exercise we use the synthetic control method introduced in Section 4.2 to study the sector-level downstream effects of the introduction of the 2001 law granting an income tax break to workers in IT. This exercise differs from the one presented in Figure 6 in the grouping of quarters into high- and low-intensity IT users. The main results compare the top 25% sectors in terms of IT services use (high-intensity category) with the bottom 75% (low-intensity category). In this exercise we compare the top 50% (high-intensity category) to the bottom 50% (low-intensity category). All figures have as dependent variable the country-level (normalized) "Employees - number." The yearly absolute value of the dependent variable in the treated sector is divided by its value in 2000, the year prior to the introduction of the income tax break in Romania. From these resulting yearly ratios we subtract the corresponding ratios for the comparison sectors. The treated sectors are those that use K72 ("Computer and related activities," including "Software consultancy and supply" and "Publishing of software") services at high-intensity. We exclude K72 itself from this category. Sectors that have a low-intensity of use of K72 services serve as comparison sectors. The data source for the dependent variable is Eurostat, Structural Business Statistics, Annual detailed enterprise statistics on services (NACE Rev 1.1). Data for the predictors comes from the World Bank, World Development Indicators. We use as predictors the "GDP per capita (constant LCU)," "Medium and high-tech sector (% manufacturing value added)" and "Services, etc., value added (% of GDP)." All figures are an output of the **synth_runner** package for Stata [Quistorff and Galiani, 2017], with the *nested* option specified.

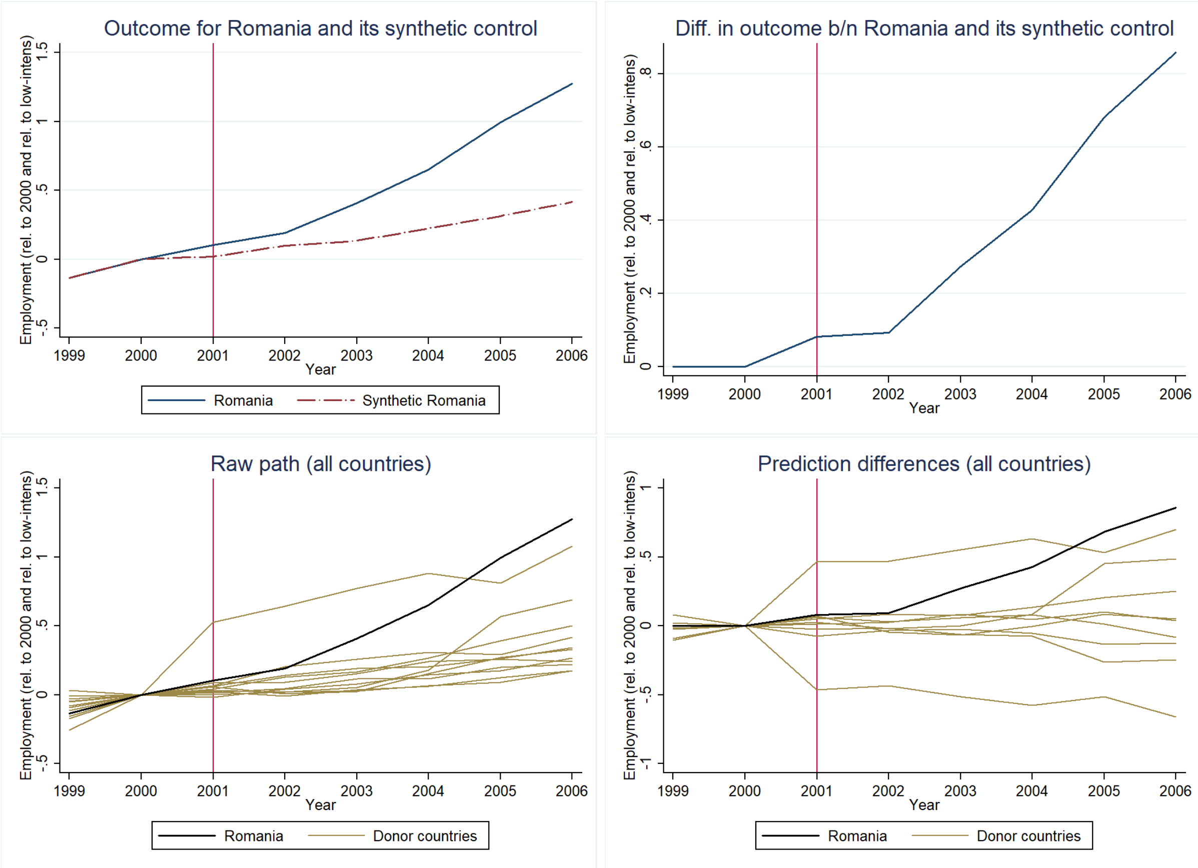

Figure C11: IT-Using Sectors Vs. Non-IT Using Sectors. SCM with Outcome Variable: "Employees - Number" (Normalized). Robustness to the Correspondence between NACE Sector Classifications

Notes: In this exercise we use the synthetic control method introduced in Section 4.2 to study the sector-level downstream effects of the introduction of the 2001 law granting an income tax break to workers in IT. This exercise differs from the one presented in Figure 6 in the period of analysis. Instead of considering the entire 1999 to 2015 period, we cut the analysis in 2006 (the last year where the data is available in the NACE Rev 1 classification). This tests the robustness of our results to the correspondence we develop between the NACE Rev 1 and NACE Rev 2 classifications. All figures have as dependent variable the country-level (normalized) "Employees - number." The yearly absolute value of the dependent variable in the treated sector is divided by its value in 2000, the year prior to the introduction of the income tax break in Romania. From these resulting yearly ratios we subtract the corresponding ratios for the comparison sectors. The treated sectors are those that use K72 ("Computer and related activities," including "Software consultancy and supply" and "Publishing of software") services at high-intensity. We exclude K72 itself from this category. Sectors that have a low-intensity of use of K72 services serve as comparison sectors. The data source for the dependent variable is Eurostat, Structural Business Statistics, Annual detailed enterprise statistics on services (NACE Rev 1.1). Data for the predictors comes from the World Bank, World Development Indicators. We use as predictors the "GDP per capita (constant LCU)," "Medium and high-tech sector (% manufacturing value added)" and "Services, etc., value added (% of GDP)." All figures are an output of the **synth_runner** package for Stata [Quistorff and Galiani, 2017], with the *nested* option specified.

Table C6: Post-treatment Results: Effects, p-values, Standardized p-values. Outcome: (Goods Export) Trade Values

Post-treatment year	$\hat{\alpha}_{1t}$	p-values	Standardized p-values
2001	0.02	0.90	0.90
2002	0.03	0.80	0.70
2003	0.16	0.80	0.40
2004	0.53	0.10	0.10
2005	1.03	0.00	0.20
2006	1.60	0.10	0.10
2007	2.30	0.10	0.10
2008	3.64	0.10	0.10
2009	4.76	0.00	0.10
2010	5.96	0.00	0.10
2011	7.46	0.00	0.10
2012	6.79	0.00	0.10
2013	7.69	0.00	0.10
2014	7.70	0.00	0.10

Notes: In this exercise we use the synthetic control method (SCM) introduced in Section 4.2 to study the sector-level downstream effects of the introduction of the 2001 law granting an income tax break to workers in IT. This table presents the SCM estimates of the treatment effects ($\hat{\alpha}_{1t}$) of the 2001 policy on the relative (goods export) trade values in Romania (where 1=Romania). This table corresponds to the treatment effects plotted in Figure 7. Treatment effects, $\hat{\alpha}_{1t}$, for Romania in post-treatment years t (2001 to 2015) are estimated as the difference between the year t outcomes for Romania and those for synthetic Romania (for details see Equation (3) in Section 4.1). The p-values are constructed using permutation tests. To calculate a p-value for each post-treatment period, the **synth_runner** package for Stata [Quistorff and Galiani, 2017] finds the proportion of placebo effects that are as large as the main effect. The standardized p-values are the studentized p-values. See Table C7 for the combination of countries and weights that comprise the synthetic control.

Table C7: Predictor balance, averaged between 1996 and 2000. Synthetic Romania = 0.706 Czech Republic, 0.240 Lithuania, and 0.053 Slovakia

	Romania	
	Actual	Synthetic
GDP per capita (constant LCU, $)	9,771	20,689
Services (% of GDP)	47.89	58.41
High-tech manufacturing (% mfg. value added)	24.33	32.60

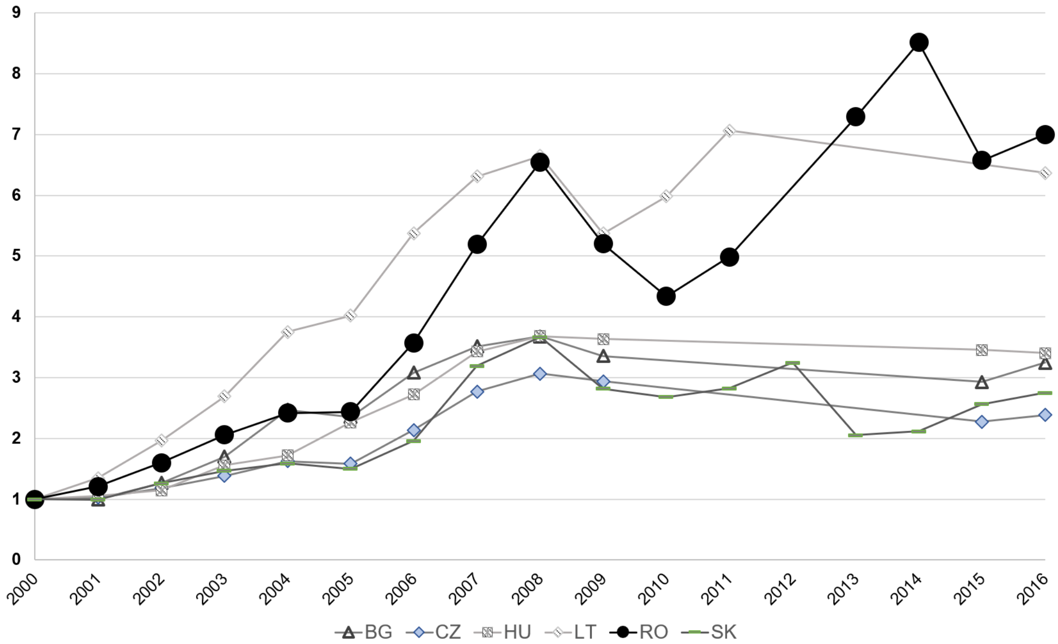

Figure C12: Trends in Export Value of IT-using Service Sectors

Notes: This figure plots the evolution of export trade values for service sectors relying more on IT services. These sectors are NACE Rev 1 sectors 51 (Wholesale trade and commission trade), 60 (Land transport), 62 (Air transport), 63 (Cargo handling and storage), 74 (Other business activities), and 75 (Public administration and defence). In Romania, most of the impressive growth experienced between 2001 and 2016 is explained by sectors under NACE Rev 1 sector 74 (e.g., call centers, advertising, business and management consultancy, secretarial and translation activities etc.). The figure includes Romania and five other countries that appear in SCM exercises as part of synthetic Romania, i.e. Bulgaria, the Czech Republic, Hungary, Lithuania, and Slovakia. This evolution is with respect to the export value in year 2000. This figure uses UN Comtrade data, EBOPS 2002. A similar graph can be made comparing the growth of IT-using service sectors to the growth of less IT-using service sectors. Romania continues to stand out in a relative comparison as well.

Appendix D Data Construction for the Synthetic Control

Appendix D.1 Eurostat Data

Downloading the Data. The first step in this process was to download data from Eurostat, using the Structural Business Statistics: Annual Detailed Enterprise Statistics tables. The data were available to download in batches, after having selected the sectors, countries, years, and variables of interest. These selections are detailed below:

Sectors: Data were available for download at the one-, two-, three-, and four-digit sector level. For the main analysis, we selected and downloaded data at the two-digit sector level due to high levels of missing data at the three- and four-digit level. According to the Rev 1.1 sector classification, there were 59 possible two-digit sectors for which we might like to have data. Our process enabled us to find corresponding data as early as 1999 for 48 of these sectors. Missing sectors include A01 (“Agriculture, hunting and related service activities”), A02 (“Forestry, logging and related service activities”), B05 (“Fishing, fish farming and related service activities”), L75 (“Public administration and defence; compulsory social security”), M80 (“Education”), N85 (“Health and social work”), O90 (“Sewage and refuse disposal, sanitation and similar activities”), O91 (“Activities of membership organizations n.e.c.”), O92 (“Recreational, cultural and sporting activities”), O93 (“Other service activities”), and P95 (“Activities of households as employers of domestic staff"). For each two-digit sector, data were downloaded in two pieces: firstly as per the Rev 1.1 sector classification and secondly as per the Rev 2 sector classification. Since the transition from Rev 1.1 to Rev 2 in the year 2007 was not one-to-one, fractions of two-digit sectors were carefully linked across the two classifications according to the degree of mutually overlapping three-digit sectors between any pair of two-digit sectors. More detail on this crosswalk follows in the "Creation of Crosswalk between Rev 1.1 and Rev 2 sector Codes" section.

Countries: Data were downloaded for the following fourteen countries: Bulgaria, Cyprus, the Czech Republic, Estonia, Hungary, Ireland, Latvia, Lithuania, Malta, Poland, Portugal, Romania, the Slovak Republic, and Slovenia. Data were widely available for most countries, with frequent missing values for Cyprus and Malta across the sectors, variables, and years of interest.

Variables: Data were downloaded for the following five variables: Gross revenues (Turnover or gross premiums written) - million euros (V12110), Production value - million euros (V12120), Employees - number (V16130), Gross investment in tangible goods - million euros (V15110), and Enterprises - number (V11110).

Years: For each sector, country, and year, relevant data were downloaded for as many years as possible, ranging from 1995 to 2016. There was low data availability prior to 1999 and after 2015 across the set of countries, sectors, and variables of interest.

After having downloaded all available data for five variables from 1995 to 2016 across 14 countries and 48 sectors, there was a further selection of years, countries, and sectors to include in the final analysis based on having a sufficiently high level of non-missing data to complete the analysis. These selections are detailed in the "Selection of Years, Countries, and Sectors for Final Analysis" section.

Creation of Crosswalk between Rev 1.1 and Rev 2 Sector Codes. The SCM analysis compares pre-trends from before and after 2001, between Romania and synthetic Romania, for sectors with high- versus low-intensity of IT use. Since the Rev 1.1 sector classification was in effect for the treatment period of interest, our analysis seeks to classify sectors as either high- or low-intensity of IT use according to their Rev 1.1 sector classification. For this reason, the data downloaded as per Rev 2 sector classifications (for years 2007 onwards) needed to be stitched together with the data downloaded as per Rev 1.1 sector classifications (for years prior to 2007) according to a correspondence or crosswalk between the two sector classifications.

An extensive correspondence detailing equivalence between two-, three-, or four-digit sectors (as applicable and as dictated by the transition) between Rev 1.1 and Rev 2 was obtained from Eurostat and used for constructing a crosswalk that could link sectors across sector classifications at the two-digit level. The revision between Rev 1.1 and Rev 2 was in most cases not a one-to-one transition. For six of the 59 Rev 1.1 sectors, one two-digit Rev 2 sector corresponded to a single two-digit Rev 1.1 sector. Six of the 59 two-digit Rev 1.1 sectors were either split into two or more corresponding two-digit Rev 2 sectors or were part of a merge into one corresponding two-digit Rev 2 sector. The remaining 47 of 59 sectors were combinations of splits and merges, that is, some portion of the two-digit sectors as per Rev 1.1 were split into two or more sectors as per Rev 2, where they were merged with other portions of two-digit Rev 1.1 sectors that were similarly split among two or more Rev 2 sectors.

Two primary methods were used to create a crosswalk to link some fraction of each two-digit Rev 2 sector to a corresponding two-digit Rev 1.1 sector. In the first method ("gross revenues and count"), the average gross revenues across all 14 countries of interest was used as a proxy for the size of each Rev 1.1 sector at the three-digit level. Data for this portion of the task was similarly downloaded from Eurostat, using the Structural Business Statistics: Annual Detailed Enterprise Statistics tables. This information on sector size was used to calculate what portion of a two-digit Rev 2 sector stemmed from distinct three-digit Rev 1.1 sectors, and in turn, to which two-digit Rev 1.1 sectors these three-digit sectors corresponded. This involved calculating the average gross revenues in 2007 across all 14 countries of interest for each three-digit Rev 1.1 sector included in the correspondence table (in cases where the correspondence table indicated linkages at the four-digit level, these were aggregated up to the three-digit level so that all comparisons at this stage were made at the three-digit level). The size of corresponding three-digit sectors common to a pair of two-digit Rev 1.1 and Rev 2 sectors relative to the total size of the Rev 2 sector were used to assign fractions of that two-digit Rev 2 sector to corresponding Rev 1.1 sectors. To illustrate, suppose a particular two-digit Rev 2 sector was composed of five three-digit Rev 2 sectors that corresponded to five three-digit Rev 1.1 sectors, three of which corresponded to one two-digit Rev 1.1 sector (totaling average gross revenues of 200) and 2 of which corresponded to a second two-digit Rev 1.1 sector (totaling average gross revenues of 300). In this case, 200/500=40% of each variable value from the Rev 2 sector would be allocated to the first Rev 1.1 sector, and 300/500=60% of each variable value from the Rev 2 sector would be allocated to the second Rev 1.1 sector.

The second method, ("count only") used a similar approach. However, instead of calculating the fraction of each two-digit Rev 2 code to assign to each two-digit Rev 1.1 code using an estimate of the size of corresponding three-digit sectors, this method used the count of three-digit sectors common to a pair of two-digit Rev 1.1 and Rev 2 sectors relative to the total number of three-digit sectors that make up the two-digit Rev 2 sector to assign the fractions. In the same example as above, 3/5=60% of each variable value from the two-digit Rev 2 sector would be allocated to the first two-digit Rev 1.1 sector, and 2/5=40% of each variable value would be allocated to the second Rev 1.1 sector. For the first method above, in cases where gross revenues was not available in 2007 in order to make the necessary estimation of sector size, fractions were allocated based on the "count only" method.

The resulting versions of the crosswalk between the two sector classifications are similar, and there is no clear test that can be used to assess which of the two methods provides a more accurate representation of each Rev 1.1-equivalent sector from 1995 to 2016. Visual inspection of trend lines (focusing in particular on the gross revenues variable for Romania over time) seems to suggest that the "count only" method leads to a smoother correspondence across Rev 1.1 and Rev 2. While there remain some jumps in the times series according to both correspondences (which could reflect the similarity in timing of the transition from Rev 1.1 to Rev 2 along with the global financial crisis), the "count only" method appears to minimize these jumps. For this reason, the "gross revenues and count" method was disregarded in favor of moving forward with the "count only" method.

Data Cleaning. After downloading data for as many countries, variables, years, and sectors as possible and constructing a crosswalk to link the data across Rev 1.1 and Rev 2 sector classifications, the next step was to combine the data, clean the data, and select years, countries, and sectors to include in the final analysis. Data that was downloaded according to the Rev 2 sector classification (for 2007 onward) was split fractionally to corresponding Rev 1.1 sectors as per the crosswalk and combined with the data that was downloaded according to the Rev 1.1 sector classification (for prior to 2007). In cases where data was available under both classifications for overlapping years (as was often the case for 2005, 2006, and 2007), the data from the Rev 1.1 classification was prioritized, then filled in with data from the Rev 2 classification when the Rev 1.1 data was missing.

Selection of Years, Countries, and Sectors for Final Analysis. Years, countries, and sectors to include in the final analysis were selected to maximize availability of data and minimize the incidence of missing observations in the resulting panel. Less than 70% of the data across all countries and sectors were available for years prior to 1999 and after 2015, so only the years from 1999 to 2015 inclusive were retained for the analysis sample. The 70% cutoff applies to all countries and all sectors; this percentage would be higher once restricting to the final set of countries and sectors used in the analysis.

Similarly, Malta and Cyprus were dropped from the analysis sample due to low data availability. The remaining 12 countries retained featured relatively high data availability: at least 75% of the desired data were non-missing for the 12 countries retained across all years, and at least 85% of the desired data were non-missing when restricting attention to the years 1999-2015. Again, these percentages become larger once we focus on the final sectors chosen for their sufficient level of data availability.

Finally, because of the importance of estimating pre-trends for Romania, sectors were dropped if data were missing for any of the seven variables in either 1999 or 2000 for Romania. This requirement eliminated 14 sectors from the analysis sample (C10, C11, C12, C13, C14, G51, G52, I60, I61, I62, I64, J65, J66, and K67). An additional four sectors were dropped for reasons of low data availability: considering only 1999-2015 for all seven variables and the 12 retained countries, these were four sectors for which more than 10% of the desired data was missing (D15, D16, D19, and D23).

After restricting the analysis sample to the years 1999-2015, dropping Cyprus and Malta, and restricting to the 48 sectors where data was available for all seven variables for Romania in 1999 and 2000 and where there was a sufficiently high level of non-missing data, remaining missing values were imputed as follows. In cases where there was a missing observation for a particular variable within a time series, the missing observation's value was estimated by a simple average of the value for the most recently available and next available year, including for cases where the gaps were greater than one year. For cases where 1999 and potentially several consecutive years were missing, all of those observations were set to match the value for the earliest available year. Similarly, in the rare situation where there were potentially several consecutive years missing data ending in 2015, all of those variable's values were set to match that variable's value for the latest available year.

The ultimate level of imputation of missing values required was minimal, ranging from 0.8% to 1.9% of desired year by country by sector observations across the seven variables.

Appendix D.2 UN Comtrade Data

Downloading the Data. For for the exports of goods and services, the data was downloaded from the UN Comtrade official website. We downloaded the data for the same set of twelve countries that constitute the pool of donor countries for the SCM exercise using Eurostat data: Bulgaria, the Czech Republic, Estonia, Hungary, Ireland, Latvia, Lithuania, Poland, Portugal, Romania, the Slovak Republic, and Slovenia. For goods exports, we downloaded the data disaggregated by two-digit SITC Rev 1 commodity codes. For service exports, we downloaded the data disaggregated by two-digit (X.X) EBOPS2002 codes (except Financial Services which is just classified as 6).[45] The variable of interest was Trade Value (US$) for Trade Flow=Exports.

Data Cleaning and Selection of Years and Codes for Final Analysis. The dataset for goods exports had the following features:

- There was a high rate of missing data for BG in 1995 and HU, PT in 2015
- We had to remove some codes to make sure that the codes entering the aggregates (high- versus low-intensity of IT inputs use) are consistent across countries. Removing the following sectors leads to a (close to) consistent panel for 1996-2014. 34. Missing: BG (1995, 1998-2000, 2003), LV (1998), RO (1996-1997), SK (2005); 35. Missing: BG (1995-2006), HU (1995-1996, 2001-2002), IE (2010-2011), PT (1997-1998, 2015), RO (1996), SK (1995-1996, 2005); 41. Missing: RO (2003); 52. Missing: HU (1998-1999, 2006, 2009, 2015), IE (2001, 2003, 2005), LV (2009), SK (2007); 93. Missing: CZ (1996-2001, 2003) , EE (1995-1998, 2002), HU (1995-1996, 2015), IE (1996), PL (1996-1997), RO (2006), SK (1995-2003); 95. Missing: BG (1995, 1997-2001, 2003, 2007-2013), HU (1995, 2003, 2009, 2015), IE (1997, 1998, 2000), RO (1998-2005), SI (1996, 1999). Note that in 2008, SK does not have sector 12.

The download of the service export data allowed us to notice that the data is available only from 2000 to 2016. Even during these years, most countries present a lot of missing values across codes. In particular, for Romania the treated NACE Rev 1 sector (72) starts reporting data only in 2005, not allowing us to study the service exports of this sector. The resulting (unbalanced) panel does not allow for an SCM, as we only observe one year before the event in 2001. Moreover, the large extent of missing values delivers noisy estimates of treatment effects.

Creation of Crosswalk between EBOPS2002 Codes and NACE Rev 1.1 Sector Codes. Our categories of high- and low-intensity (of IT usage) sectors were constructed based on Romania's input-output Table for 2000, which is at the two-digit NACE Rev 1 sector level. Hence, we had to construct a mapping between two-digit SITC Rev. 1 commodity codes and two-digit (X.X) EBOPS2002 service codes and the two-digit NACE Rev 1 codes. After commodity or service codes were mapped to NACE Rev 1 codes, we aggregated export values at the NACE Rev 1 level.

[45] EBOPS2002 includes the standard items, memorandum items, and supplementary items. Only standard items were used. For codebook, visit link.